The Optimally Diversified Equity Portfolio in South Africa: An Artificial Intelligence Approach

Aaron E. Block

School of Economic and Business Sciences

Faculty of Commerce, Law and Management

University of Witwatersrand

Johannesburg

South Africa

The Optimally Diversified Equity Portfolio in South Africa: An Artificial Intelligence Approach

Aaron E. Block

473754

Supervisor: Dr. Y Seetharam

School of Economic and Business Sciences

A thesis presented to the School of Economic and Business Sciences, Faculty of Commerce, Law and Management, University of Witwatersrand in fulfilment of the requirements for the degree of Master of Commerce (M.Com) in Business Finance.

January 2017

Declaration:

I, Aaron Block, declare that this research project is my own, unaided work. It has not been submitted before for any other degree or examination at this or any other university.

Aaron Block

Date

Acknowledgements

I would first like to thank my thesis supervisor, Dr. Seetharam, for your unyielding support through the development of this dissertation. The feedback you provided was consistently timeous and invaluable and you remained in contact over various mechanisms whenever I had a question or required clarification on my research. The academically challenging nature of our dialogues persistently motivated me to improve my thought processes and thereby my work.

Secondly, I must express my deep appreciation to my girlfriend and family for providing endless support and encouragement over my years of study and throughout the completion of this dissertation. This, one of my proudest accomplishments, would not have been possible without you.

Finally, I complete the acknowledgements section with a quote by Shakespeare, poetically describing an early method of diversification and the calming effects one may enjoy through its proper application.

"My ventures are not in one bottom trusted,
Nor to one place; nor is my whole estate
Upon the fortune of this present year:
Therefore my merchandise makes me not sad."

(Act 1, Scene 1; Shakespeare: *The Merchant of Venice*, 1905)

Definitions of Terms and Abbreviations

Artificial Intelligence - A branch of computer science focusing on the simulation of intelligent behaviour in computers. This concept relates to the capability of a machine to imitate intelligent human behaviour.

Beta - A measure of the volatility of a security or a portfolio in comparison to the market as a whole. Beta is computed as a variable known as the beta coefficient.

Bloomberg Professional Platform - A popular software platform that combines real-time data on markets, news and research, analytics, communication tools and execution capabilities into one fully integrated solution.

Correlation - A statistical measure that evaluates the degree to which two securities move in relation to one another. Correlation is represented by the variable termed the correlation coefficient and is bound between the values -1 and 1.

Covariance - A measure of the degree to which two variables move in tandem. This is computed by calculating the mean value of the product of the deviations of two variates from their respective means.

Diversification - The act of spreading one's assets so as to minimise the effect of a singular event affecting all assets held negatively.

Efficient Frontier – Refers to a set of optimal portfolios that offer the highest expected return for a defined level of risk, or alternatively, the lowest amount of risk an investor need encounter to receive a given level of expected return. Portfolios that lie below the efficient frontier are deemed sub-optimal while those that lie on or above the efficient frontier are considered superior.

Genetic Programming – This term refers to a set of mathematical techniques that are based on biological principles that mimic concepts such as natural evolution/selection as well as the neurological functioning of the brain.

Idiosyncratic Risk – Also termed unsystematic risk, relates to firm specific risk factors such as key person risks, fraud risks and the like. Idiosyncratic risk is possible to limit through the use of diversification. This stands in contrast to systematic risk which refers to the risk inherent in the equity market as a whole.

Johannesburg Stock Exchange (JSE) - The JSE offers secure, efficient primary and secondary capital markets across a diverse range of securities, supported by post-trade and regulatory services.

Law of Large Numbers - A statistical term used in probability theory that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.

Markowitz Portfolio Theory - A theory on how risk-averse investors can construct portfolios to optimize or maximize expected return based on a given level of market risk, emphasising that an increased amount of risk is an inherent function of obtaining higher rewards.

Neural Network - A computer system modelled on the human brain and nervous system. A Neural Network is characterised by the pattern of connections among various network layers, the number of neurons in each layer, the learning algorithm and the neuron activation functions.

Neurons - Specialized cells in the brain which transmit information across the central nervous system.

Particle Swarm Optimization (PSO) - A population based stochastic optimization technique inspired by social behaviour of bird flocking or fish schooling.

Penny Stocks - Common shares of small public companies that trade at low prices per share. These stocks are generally considered highly speculative and high risk because of their inherent lack of liquidity, large bid-ask spreads, small market capitalisation and limited monitoring and disclosure.

St. Petersburg Paradox – This paradox relates to probability and decision theory in economics. It is based on a particular (theoretical) lottery game that leads to a random variable with an infinite expected payoff but nevertheless seems to be worth only a very small amount to potential participants. The St. Petersburg paradox is a situation where a naive decision criterion which takes only the expected value into account predicts a course of action that presumably no actual person would be willing to take.

Standard Deviation - A measure used to quantify the amount of variation or dispersion of a set of values from its mean. This is equivalent to the square root of variance.

Swarm Intelligence - The collective behaviour of decentralized, self-organized systems that can be natural or artificial. Individual members of the swarm follow simple rules, and although there is no centralized control structure dictating how individual members should behave, interactions between members lead to the emergence of "intelligent" global behaviour, unknown to the individual members.

Systematic Risk – Risks that are assumed to affect the equity market as a whole and thus cannot be diversified, if one assumes the portfolio held by an investor is restricted to the equity asset class.

Traditional Simulation – An approach utilising a random selection of shares over multiple trials, under specific constraints, in an attempt to find the optimally diversified portfolio.

Transaction Costs - Costs incurred when making an economic exchange or equivalently the cost of participating in a market.

Variance - A measurement of how far each number in a data set is from its mean.

The Optimally Diversified Equity Portfolio in South Africa: An Artificial Intelligence Approach

Abstract

Diversification has remained a central tenet in investment theory over multiple decades due to its demonstrated value as a risk mitigation technique. Increasing the number assets in a portfolio, where the magnitude of correlation is relatively slim, increases the amount of diversification while also encountering increased costs in the form of transaction costs, taxes and the like. Thus, it is imperative to solve for the optimal point of diversification to ensure an investor does not encounter unnecessary costs.

This study aims to solve for the point of optimal diversification in an equity portfolio, focusing on the South African environment. This is achieved by employing a framework using both the traditional simulation method as well as more advanced mathematical techniques, namely: genetic programming and particle swarm optimisation. Marked improvements are realised in this study with regards to the methodology and results through the application of advanced mathematical approaches in addition to removing the restriction of equal weightings being applied to each share in the portfolio.

The results revealed that an optimal portfolio can be constructed using up to only 15 shares. Secondly, the genetic programming approach demonstrated increased strength compared to the traditional simulation and particle swarm optimisation approaches, obtaining a greater level of diversification with fewer shares. Finally, although the aim of the study is focused on modelling the relationship between the number of shares in a portfolio and the achievable diversification benefits, it is also established that the portfolios indicated as being optimally diversified achieved market beating returns.

Table of Contents

Contents

List of Tables

List of Figures

Chapter 1: Introduction

Diversification is defined as the act of spreading one's assets so as to minimise the effect of a singular event affecting all assets held negatively. The term also refers to the situation where a company may vary its range of products in order to protect against a downturn in any individual product's industry (Stevenson, 2010). The work by Bernoulli on the St. Petersburg paradox in 1738 (as cited in Rubinstein, 2002) is widely believed to be the first evidence of the study of diversification and the effects. Diversification is utilised as a risk mitigation technique, however it must be recognised that diversifying beyond that which is deemed necessary may lead to suboptimal outcomes. This is due to the reality that by marginally increasing the amount of diversification one will also encounter marginally increasing costs.

This can be observed in both an asset management scenario, as is the focus of this dissertation, as well as in the application of product diversification. In the case of asset management, focusing on an equity portfolio, one can decrease the risk in the portfolio by adding uncorrelated shares to a portfolio, as will be further expounded upon in this study. Increasing the number of shares to a portfolio however leads one to encounter increased costs in the form of trading costs, taxes and the like. In the product diversification example it can be observed that a firm increasing its product set has a positive effect on risk, as if one product category were to be negatively impacted the firm would also be exposed to the second, uncorrelated product category. This strategy may limit the economies of scale that a firm can obtain, resulting in the firm being less efficient.

The above illustrates the importance of discovering the intersection of where marginal benefit of diversification meets the marginal cost. The point where the marginal benefit of diversification meets the marginal cost can be said to be the optimal point of diversification. This study aims to solve for the point of optimal diversification in an equity portfolio focusing on a South African environment. Three mathematical approaches are applied with this intention in mind: a traditional simulation approach, a genetic programming approach and a particle swarm optimisation approach. Each of these approaches will be applied to the data over a set of four separate but related tests as to both solve for the optimal point of diversification as well as further allowing for a comparative inspection as to which mathematical approach attains the best result, thus being deemed the strongest mathematical approach. Each of these approaches and the relevant tests will be elucidated upon in turn however it is valuable to continue by further reviewing the initial literature on diversification, credited to Bernoulli (1954) exploring the St. Petersburg paradox.

The St. Petersburg paradox involves the flipping of a fair coin until it lands on tails. The total number of flips, n, determines the magnitude of the pay-out which is made to equal $\$2^n$. If the coin lands on tails the game ends immediately, conversely if heads is landed on the coin is flipped again – this continues until the coin lands on tails. The question proposed related to what a rational person would be prepared to pay to play this game?

Traditional mathematics at the time suggested that a rational person should be prepared to pay an infinite amount to play this game as the potential pay-out tends towards infinity. This inference is not practical as no individual in reality could be thought to be prepared to pay an infinite amount to play a game where he has small probability of winning an infinite amount in return. Bernoulli (1954) solved this problem by introducing the concept of marginal utility. Marginal utility refers to the concept that the additional unit of return resulting from any small increase in wealth is inversely proportionate to the quantity of wealth previously possessed. In other words, a $1 prize would make a greater positive impact to a person whom only possesses $1, or a similar small amount, as opposed to someone whom may have a million dollars. As such the current quantity of one's wealth should be considered when deciding the worth of a potential pay-out. The St. Petersburg paradox was solved in this way as instances as where the pay-out begins to become extraordinarily large will lead to the subsequent win having less additional utility to the investor; this will continue until the additional pay-out has such a limited impact on an investor's utility that it should not be included when considering what an investor would be prepared to pay to play the game. As such the entry fee tends towards a finite amount that a player would be prepared to pay to play the game (Bernoulli, 1954).

Bernoulli (1954) then applied this to the case where an individual may stand to lose an amount, rather than gain an amount as in the traditional St. Petersburg paradox. The question was rephrased to "what amount should a rational person be prepared to pay to minimise the probability that a certain loss could be incurred?" Using the same principles as mentioned above, it was conceived that if an individual experiences the potential of a loss, they would be willing to pay an increasing amount to insure against this, up unto the point where the marginal cost of insuring an additional unit is equal to the marginal benefit of insuring the same unit. It was thus asserted that it is better to divide goods that are exposed to some small danger (risk) into several portions so as to not risk them altogether (Bernoulli, 1954). The payment necessary to minimise the probability of a loss should continue, however one should not encounter needless additional costs if the additional benefit from dividing the goods is worth less than the cost of doing so. This line of thought formed the foundation for several studies aimed at assessing the optimal amount of diversification by weighing the costs of additional diversification against the risk minimising benefits thereof (Bernoulli, 1954). Therefore, Bernoulli (1954) provided the groundwork of both the mathematical equations that underpin solving for optimal points while taking into account cases of marginality, as well as expounding upon the thought paradigms one should consider when exploring diversification and risk to reward trade-offs.

The concepts of diversification and risk go hand in hand however it is important to not use the terms 'risk' and 'uncertainty' interchangeably as was first highlighted by Knight in 1921 (Knight, 2012). The term 'risk' refers to situations where there are known outcomes that occur from a known probability distribution (Tversky & Fox, 1995). 'Risk' can thereby be quantified by attaching probabilities to each potential risk event based on statistical analysis or experience. 'Uncertainty' however relates to either known or unknown outcomes that occur from an unknown probability distribution (Tversky & Fox, 1995). Diversification is shown to only be useful when the probability of a specific outcome is unknown but where the probability distribution is known or can be approximated. Regarding investments, known or approximated probability distributions are required in order to use various statistical models that could be applied to forecast future share price changes. Relating these definitions to the realm of investment demonstrates that although an investor may be uncertain about the changes of future asset prices, it can be said that he does not face uncertainty but rather that he faces risk – as the distribution of asset prices can be approximated.

Practically, an investor may be uncertain of a specific company's future share price. However, by understanding the probability distribution of all share price changes (all being understood to mean a representative sample), an investor can probabilistically predict what a company's future share price may be. This is accomplished by utilising the probability distribution of a share as an input into a model that relies upon this measure to forecast the next series of prices for the share that still match the same probability distribution as the one input (Markowitz, 1991; Akgiray, 1989). This can be applied repetitively to calculate a series of future prices for each share in the representative sample (Pai & Lin, 2005). Once each share's price in the sample has been forecast, an investor is able to construct a portfolio with the risk-reward characteristics that he desires. This study demonstrates a methodology that allows an investor to employ a set of prices to discover a portfolio that delivers the optimal level of diversification. Furthermore, if an investor elects to hold more risk than is necessary, the above approach will allow him to adjust his portfolio accordingly until a suitable level of diversification is found and utilised.

When an investor faces risks regarding the probability of specific future events and how they could affect his portfolio, he can utilise the strategy of diversification in an attempt to minimise the risk of his portfolio, independent of what the various outcomes may be. This protects the investor from idiosyncratic risks – risks that are attribute specific (Markowitz, 1952b). In the case of equities idiosyncratic risk is usually taken to refer to firm specific (unsystematic) risk factors (such as key person risks, fraud risks and the like). Idiosyncratic risk can also however refer to other risks such as industry specific risks (such as new regulations being enforced that hinder a specific sector) or country specific risks (Markowitz, 1952b). By using the strategy of diversification, which would spread an investors wealth across various assets (or equities as in this study) each exposed to different countries and industries the investor is protecting himself from the effects that idiosyncratic events may have on their portfolio. Lowering the effects of individual risk factors in this fashion is in line with the theory set forth by Markowitz (1952b).

By the same logic, if an investor is sure of all future events or of all future share price movements with 100% certainty, he will in no case utilise diversification or prefer a diversified portfolio (Markowitz, 1991). The investor will, in this instance, place all of their available funds in the security that they know for certain will achieve the maximum returns. Thus, being certain of future outcomes the investor holds a risk free position and in this way he has no need to diversify his portfolio. A risk free position can be described as a position when the known outcome is certain to occur at a point in the future, and thus is independent of the former probability distribution. In the situation where an investor knows future events and future share price movements with certainty he faces neither idiosyncratic risk factors, nor systematic (market wide) risk factors as all possible events and their effects are known with certainty (Markowitz, 1991)..

Furthermore, if two securities are known (with certainty) to achieve maximum but equal returns an investor will, ideally, be indifferent as to which of the securities or which combination of the securities to invest all of their available funds in, *ceteris paribus* (Markowitz, 1952b). The *ceteris paribus* assumption above excludes the considerations of behavioural factors, transaction costs, taxes, liquidity and the like that may be different between the two alternatives.

The foundations for understanding diversification as well as the importance of calculating the optimal point of diversification has been laid out above. The aim of this study is to solve for the minimum number of shares to be included into a portfolio in order to obtain the optimally diversified portfolio within a South African environment. This problem is solved by applying three approaches, namely: the traditional simulation approach, the genetic programming approach and the particle swarm optimisation approach all of which will be explained in further sections of the dissertation. The secondary initiative of this study is to demonstrate which of the above three approaches obtains the best results – the result where either a smaller number of shares is shown to expose the investor to the same amount of risk as a larger portfolio of shares that was calculated by another approach, or conversely the approach that solves for the result that exposes the investor to a smaller amount of risk whilst utilising the same number of shares in the portfolio as was calculated by a different approach. The approach that obtains the best result is said to be the strongest mathematical method with regards to solving the above problem.

Thus this study investigates two research questions. The first refers to the question: "What is the minimum number of shares to include into an equity portfolio, in order to achieve an optimal level of diversification in a South African environment?" The second research question focuses on answering the question: "Of the three selected mathematical approaches, which demonstrates the most strength when solving for the optimally diversified portfolio in a South African environment?" With regards to the first research question, it is hypothesised that that optimally diversified portfolio will consist of between 10 and 30 shares, in agreement with previous literature that will be further explored in the following section. This question is explored in two scenarios, the first being when each share in the portfolio is restricted to be of equal weight, while in the second scenario the weights are allowed to vary between 0 and 1 while also restricting the sum of the weights in the portfolio to equal 1; these restrictions will be explained in more detail in further sections. Although previous literature has analysed this research question by making use of portfolios where shares are restricted to be of equal weightings, this dissertation is one of the first to allow the weights on each of the shares in the portfolio to vary – this has the potential to display further insight into the strength of the optimisation processes and may indeed uncover further diversification benefits due to the more flexible allowances. The hypothesis for the second research question is that the more computationally intensive and complex models of genetic programing and particle swarm optimisation will display increased strength in solving for the optimally diversified portfolio as opposed to the traditional simulation method. The advanced approaches of genetic programming and particle swarm optimisation have not been applied to the South African equity market in prior literature. Additionally, the most recent study with a focus on the number of shares required to construct an optimally diversified portfolio in a South African environment was completed by Neu-Ner and Firer in 1997, as will be further discussed in Chapter 2. This study serves to update and compare the results achieved while including the employment of the contemporary, advanced approaches of genetic programming and particle swarm optimisation

This dissertation will continue in Chapter 2 by exploring the findings and methodology of past literature relating to the aims of this study. The review of past literature is completed in sections, firstly the two types of diversification are elaborated upon as well as the literature relating to the respective concepts. The chapter then continues by examining literature covering the two artificial intelligence approaches, genetic programming and particle swarm optimisation. This section describes the methodology that underpins each artificial intelligence approach and continues by elucidating upon the work of previous authors that employed these approaches along with their respective results and shortcomings. The discussion is then followed by an inspection into behavioural factors that may lead an investor to sub-optimal diversification. Thereafter, Markowitz portfolio theory is elaborated upon. The dissertation continues in Chapter 3 with an elaboration upon the data and methodology utilised. Finally, the analysis and results discovered are discussed in Chapter 4 which is followed by Chapter 5, which provides a conclusion to the dissertation accompanied by a section detailing the considerations and limitations of the current study, a summary of the results as well as an examination of the potential areas where future research could aim to contribute and improve upon current literature.

Chapter 2: Literature Review

In this chapter, the concept of diversification and how one can attain the 'optimal amount' of diversification will be further analysed and discussed in depth, through the compilation of past studies and their respective results. The approaches, strengths and pitfalls of each paper as well as their findings will also be further elucidated upon. The following analysis will discuss the concept of diversification across all three mathematical approaches implemented in this dissertation, namely: the traditional simulation approach, the genetic programming approach and finally the particle swarm optimisation approach. In addition, studies where the above mentioned approaches were applied to scenarios other than investigations into diversification are elaborated upon in order to assess the generalisability of the approaches to both financial and non-financial problems. The chapter then continues with an exploration of the Markowitz portfolio theory and thereafter concludes by expanding upon the behavioural factors that may lead to sub-optimal diversification.

2.1 Optimal Diversification

Diversification as it relates to finance can be demonstrated by two main concepts. The first, and more well-renowned, is when diversification is applied to asset (commonly equity) portfolios to minimise the risk of an investors' portfolio at a given level of return. This concept can be viewed conversely too – diversification can be applied to an asset portfolio to maximise the return of the portfolio at a given level of risk. The second method of diversification relates to product differentiation by a firm. Although this second method of diversification has gone largely undocumented in financial literature, many of the lessons learned from this method can be transferred and applied to the first, more traditional approach. Both methods as well as how they relate to the diversification of financial assets will be discussed below.

2.1.1 Diversification by Assets

The relationship between the risk of a portfolio and the number of securities (usually referring to the number of equities) in that portfolio has received a great deal of attention over the years due to the major implications it has for the structure and very existence of financial intermediaries, as well as for the behaviour of all investors. When an investor decides on the size of the portfolio he will hold, he is establishing a trade-off between the decreased risk due to more effective diversification by increasing the number of shares included in the portfolio – particularly those that have enhanced diversification properties such as exhibiting low correlations with one another, against the increased transaction costs of adding additional securities to the portfolio. Transaction costs can come in the form of potentially decreased returns, as well as monitoring costs which refer to costs borne by an investor that enable him to monitor an increasing number of securities (these can be direct, such as licenses allowing him to buy or sell additional shares as well as indirect, such as the increased attention required to monitor a portfolio comprising of a larger number of securities) amongst others (Elton & Gruber, 1977).

Elton and Gruber (1977) hypothesise that if large portfolios are required to get most of the benefits of diversification then financial intermediaries should exist solely as a method for providing investors with the benefits of diversification at lowered transaction costs. However, if one can obtain most of the benefits of diversification with a portfolio consisting of a relatively small number of securities, then an investor can achieve effective diversification directly – he can maximise the diversification benefits of his portfolio using a small number of securities thus avoiding numerous transaction costs that would be encountered by the need to incorporate a large number of securities and as such will not need a financial intermediary to assist in attaining this benefit. This would mean that financial intermediaries can only justify their existence by their ability to select securities that will exhibit superior (above market) performance.

Whilst several previous papers have made use of a simulation based approach as a method to discover the most efficiently diversified portfolio, and which shares it is made up of, Elton and Gruber (1977) used an analytical expression to model the relationship between risk and the number of securities included into a portfolio. The derived mathematical expression allowed the researchers to utilise factors of the population of shares being investigated (which included the average variance, co-variance, returns and expected returns) and specify the number of shares the portfolio should consist of. The formula would provide an approximate minimal variance that a portfolio with the specified number of shares should yield. The number of shares to be included in a portfolio would then be adjusted and the respective answers analysed and compared to one another. As this approach computes the hypothetical optimally diversified portfolio using a mathematical approach which relies on parameters derived from the population of shares, the approach does not allow the researchers to view which of the shares should be included in the optimal portfolio, but merely how many shares should be utilised.

In this way one can observe the effects on risk when new securities are added into the population of those initially forming part of the portfolio. Elton and Gruber (1977) argue that one need not resort to simulation when relatively well-developed statistical models allow for the determination of the relationship between risk and return across fluctuating market conditions. This further allows one to explicitly observe which factors influence the effect of the size of the portfolio on risk and the relative importance of each factor. The analytical solution should however be approximately equal to a simulated solution, should all possible iterations be simulated in the analysis.

Furthermore Elton and Gruber (1977) noted that the risk measure used by previous authors, as was employed by Evans and Archer (1968), somewhat underestimated the true 'total risk' of a portfolio. Previous studies examined the dispersion of a portfolio's return around its mean return. Elton and Gruber (1977) note that this neglects the probability that the mean return will be different from the market. As such they stated that when analysing the benefits of diversification, the risk (standard deviation or variance) of the selected portfolio should be compared to the risk of an equally weighted portfolio of all the tradable securities in the population (the standard deviation of a portfolio is taken to be a quantity expressing the amount that the members of a group differ from the mean value of the group). In this way, two parts of portfolio risk are captured, the first being the variance of the return on the portfolio from the portfolio's expected return and secondly, the variation caused by the difference between the expected return on the portfolio and the expected return of the population of shares included (Elton & Gruber, 1977). It is assumed that the equally weighted portfolio was utilised as opposed to the value weighted portfolio as this was commonly accepted at that period in time, although this reasoning was not expressly stated in the study.

Elton and Gruber (1977) analysed the weekly returns of a population consisting of 150 shares selected from the New York Stock Exchange (NYSE) and American Stock Exchange over the period of June 1971 to June 1974. It was demonstrated that 51% of a portfolio's standard deviation of returns can be diversified away with between one and ten shares; adding a further ten shares eliminates another 5% and adding a third set of ten shares further eliminates the standard deviation by a further 2%. The analysis assumed equal weightings on all securities included into the portfolio. This approach is said to be optimal when an investor has no information about future returns, variances and co-variances of the shares involved (as in standard probability theory and statistics, covariance refers a measure of how much two variables change together). It is noted that to the extent that these factors can be forecast, risk could be further reduced.

It is hypothesised that utilising a portfolio where the weightings on individual shares are allowed to be adjusted (so that the shares are no longer equally weighted), could yield further risk reduction benefits. Unequal weightings allow the model to be relatively more nimble than the rigid methodology of assigning an equal weighting to each share included in the portfolio. As such, using equal weightings on each share in the portfolio provides an upper limit on the risks an investor faces (Elton & Gruber, 1977). This could not be accomplished previously due to computing limitations at the time of writing the article, however it is expected that the drastic increase in computing power coupled with the methodology that will be utilised in this dissertation, to be explained in the following sections, can be employed to derive optimal weightings of shares within a portfolio. The flexibility of this approach could potentially uncover further positive effects of diversification.

Elton and Gruber (1977) further noted that if transaction costs for a security are strictly proportional to the size of the transaction, then the total amount of transactions would be independent of the number of securities in the portfolio. A fixed component to transaction costs would however modify the conclusion – if transaction costs increase as the number of securities included in a portfolio increases, it is likely that an investor would prefer to hold fewer securities rather than more.

Statman (1987) indicated that a well-diversified portfolio should consist of at least 30 shares. This evidence stood contrary to the industry accepted norm of the time, which was that a well-diversified portfolio need only consist of 10 shares, as was shown to be optimal in previous research (Elton & Gruber, 1977). The approach used by Statman (1987) was a traditional simulation based approach, as is the first approach applied by this dissertation. The traditional simulation approach refers to the utilisation of a random selection of shares over multiple trials, under specific constraints, in an attempt to find the optimally diversified portfolio (Statman, 1987). As the selection of shares is completed at random, the most optimally diversified portfolio found may not be the globally most optimal diversified portfolio as there may be various combinations of shares that were not included in the testing phase (Abraham & Nath, 2000). This problem, referred to as the ‘local optimum trap’, is a generic fault inherent in the simulation approach.

Statman (1987) stated, in line with previous research, that the risk of a stock portfolio depends on the proportions of the individual stocks included in the portfolio as well as their variances and covariance’s respectively (these terms are each explained as well as their formulas provided in the sections that follow). A change in any of these variables will change the risk (variance) of the portfolio. It is however generally true that when stocks are randomly selected and combined in equal weighting into a portfolio, the risk of the portfolio declines as the number of different stocks included into the portfolio increases (Markowitz, 1952b). Furthermore, it should be noted that only the unsystematic risk can be diversified, as systematic risk is assumed to affect the equity market as a whole and thus this risk cannot be diversified if one assumes the portfolio held by an investor is restricted to equity (Statman, 1987). Unsystematic risk refers to firm specific risk while systematic risk refers to risks that affect the entire market – this risk cannot be diversified as mentioned previously (Markowitz, 1952b).

Statman (1987) used the analogy that diversification can be viewed in the light of traditional economics, where one should continue to increase the number of shares in a portfolio until marginal benefits equal marginal costs. Marginal benefit in the context of diversification is the reduction in risk of the portfolio. Marginal costs refer to the transaction costs involved in adding an extra share to a portfolio (Statman, 1987). The transaction costs can be accounted for directly (by subtracting a portion of a portfolio’s return for each additional security added to a portfolio) or they can be statistical, by adding a measure to the variance of a portfolio for each additional security added to the portfolio, or by the introduction of a rule stating that an additional share should not be added to a portfolio if the addition does not reduce the variance of the portfolio by a pre-specified amount (Magill & Constantinides, 1976; Fang, Lai & Wang, 2006).

Statman (1987) mentions the difficulties in attempting to quantify the transaction cost of adding a share to a portfolio directly as these differ drastically across different geographies and exchanges. In some instances transaction costs are related to the nominal amount of the trade (those who trade in larger sizes attract a relative reduction in transaction costs) while in other instances the transaction costs are related to the frequency of the trades (the transaction costs are higher for those that trade more frequently); in some circumstances transaction costs may be a combination of both factors as well as others (such as the market capitalisation of the companies involved, with companies with smaller market capitalisations attracting higher fees) (Statman,1987). Due to these difficulties in quantifying transaction costs, this dissertation makes use of a statistical transaction cost when attempting to create the optimally diversified portfolio, as will be further expounded upon in the sections to follow.

It is expected that initially the marginal benefit of adding a share to a portfolio will far outweigh the marginal cost of doing so, however as more shares are added to a portfolio the marginal costs will begin to increase at a faster rate than the marginal benefits until eventually marginal benefits equal marginal costs. At this point a portfolio can be said to be optimally diversified (Mayshar, 1979; Statman, 1987).

Statman (1987) utilised returns as the common measure necessary to compare marginal benefits to marginal costs. In this way the risk reduction benefits of diversification, in units of expected return, can be determined by a simple comparison of any two portfolios. A 500-stock value-weighted portfolio was used as a benchmark and as such was believed to represent an attainable, fairly diversified portfolio. Although this benchmark portfolio is notably different to the benchmark used by Elton and Gruber (1977), in that it is value weighted while Elton and Gruber (1977) used an equally weighted benchmark; it is assumed that the costs of maintaining a value weighted portfolio and an equally weighted portfolio are equal and thus the findings can be directly compared to one another (Statman, 1987). It is not however necessary to create or utilise a benchmark portfolio in the current research study as a statistical transaction cost is used as will be explained in the methodology section.

Number of Shares Required to Optimally Diversify a Portfolio

There has been a vast amount of additional literature focused on how many shares constitute an optimally diversified portfolio, with many finding differing results. Evans and Archer (1968) examined the rate at which the variation of returns for a randomly selected portfolio is reduced as a function of the number of securities included in the portfolio. They concluded that an optimally diversified portfolio should contain approximately ten shares. Gup (1983) argued that an investor does not need to invest in a large number of industries or shares and that an optimally diversified portfolio can be gained by acquiring between eight or nine shares while Stevenson and Jennings (1984) noted that an optimally diversified portfolio should consist of between eight and sixteen shares; similar findings were made by Reilly (1985) whom asserted that between 12 and 18 shares should be included in an optimally diversified portfolio. Statman (1987) thereafter, counter to previous research, stated that an optimal portfolio should consist of at least 30 shares.

Solnik (1974) analysed the variance of weekly returns and displayed how diversification affects large share portfolios in various countries. It was concluded that the effectiveness of diversification in reducing the risk of a portfolio differs from country to country. This is due to the average covariance relative to the overall variance being different internationally; in some countries share prices tend to move together to a greater extent than they do in other countries. Table 1 below demonstrates their findings:

Table 1: *Changes in the Effectiveness of Diversification by Country*

Percentage of the risk of an individual security that can be eliminated by holding a random portfolio of shares within selected national markets	
USA	73.0%
UK	65.5%
France	67.3%
Germany	56.2%
Italy	60.0%
Belgium	80.0%
Switzerland	56.0%
Netherlands	76.1%
International Shares	89.3%

Note: Adapted from “Why not Diversify Internationally rather Domestically?”, by Solnik, 1974, *Financial Analyst Journal*

Neu-Ner and Firer (1997) wrote the seminal paper of a study on diversification focusing on the Johannesberg Stock Exchange (JSE). The objective of the study was to establish how many randomly selected JSE shares are required to achieve a well-diversified portfolio. This was compared to an equally weighted portfolio of all securities in the population. Neu-Ner and Firer (1997) noted that risks inherent in an investment are often described as being made up of the risks which are common to all assets and thus cannot be diversified (systematic or market risk) and those which are unique to the asset and thus can be eliminated by diversification (firm specific risk). As portfolio size increases, so the risk of the portfolio falls due to the elimination of firm-specific risk. Of importance to investors, is the number of assets in the portfolio beyond which the addition of further assets will not (for all practical purposes) result in further reduction in risk without the costs involved overcoming the benefit.

It is interesting to note that the JSE is unique in several of its' characteristics. It experiences a very high concentration of economic power in very few hands (Neu-Ner & Firer, 1997). It is widely acknowledged that although the JSE lists over 400 shares, the top 40 shares can act as a fair reflection of the South African stock market as a whole as the top 40 companies (as listed by market capitalisation) represent over 80% of the total market capitalisation of all JSE listed companies at any one time. This fact was demonstrated by accessing data from the Bloomberg professional system and calculating the ratio of the value of the JSE Top 40 index to the JSE All Share index. Furthermore the JSE does not allow for short selling shares (Neu-Ner & Firer, 1997; JSE Equities Rules, 2016).

The only study of diversification on the JSE before Neu-Ner and Firer (1997) appears to have been an unpublished pilot study by Bradfield (1993). Bradfield (1993) found that in comparison with the NYSE, approximately five more shares were needed to create a 'completely' diversified portfolio on the JSE. As such it can be said that South African shares have a proportionally higher percentage of undiversifiable risk as the South African economy is somewhat less diversified than that of the U.S. However holding an excessive number of shares increases monitoring costs as well as potentially transaction costs. This highlights the importance of finding the optimal number of shares on the JSE necessary to be included in order to achieve a well-diversified portfolio.

Neu-Ner and Firer (1997) had three objectives in their study. Firstly, using a simulation approach they aimed to determine the relationship between the number of randomly selected JSE listed shares in a portfolio and the risk of the portfolio.The second objective was to find out how many randomly selected shares are required to achieve a relatively well-diversified portoflio. Thirdly, whether the relationship between the number of randomly selected shares from the JSE and the risk an investor faces can be modified through the utilisation of certain selection rules; for example selecting only shares from the list of *Financial Mail* “Top 100” Top performer companies, restriction of the selection of shares to only those that have a relatively high market capitalisation value or the selection of shares based on the value of their beta. The beta of a share is a measure of volatiliy or systematic risk. A share with a beta of 1.5 for example denotes that for each 1% that the market moves, the share (or portfolio of shares) will move 1.5% in the same direction as the market. A beta of 0.8 is taken to reflect that for each 1% move of the market the share (or portfolio of shares) will move 0.8%.

The population studied by Neu-Ner and Firer (1997) included all securities listed on the main board of the JSE during the period June 1993 to June 1996; shares that were listed or de-listed during the study period were excluded. Furthermore the role of dividends when calculating the share returns was not taken into account. This eventually provided a research population of 532 securities. The data studied consisted of the weekly closing prices of these respective shares. Once the data was downloaded from INET (Intelligent Network Share Data Service) into an excel spreadsheet, the weekly returns on the shares were calcuated. Portfolios were thereafter randomly constructed assuming equal investment in each share .

The process of simulation was carried out by randomly selecting *N* shares (between 1 and 532) and then finding the risk associated with that portfolio. Simple random sampling without replacement was utilised. In statistics, simple random sampling refers to selecting a random subset of individuals from a larger population with each individual being chosen randomly and entirely by chance. Furthermore, each individual has the same probability of being selected at any stage throughout the sampling process. Using simple random sampling without replacement ensures that the same security cannot be included into a portfolio more than once, for logical reasons (Singh, 2003).

One thousand portfolios were created for each value of *N* and the average risk over the 1000 trials, where each trial represents a portfolio comprised of *N* shares, was then taken to be the expected risk associated with that portfolio. Risk was measured by three different methods. Firstly, the standard deviation of the weekly returns on the portfolio were calculated; secondly, the variance was calculated (this measure was seen as necessary to calculate, despite that it is derived from the standard deviation, as it allowed the results to be compared to previous studies, such as to the study by Elton and Gruber (1977), whom utilised the variance metric as a proxy to illustrate the risk of a portoflio) and finally the comparison of the standard deviation of the portfolio to the standard deviation of the equally weighted population portfolio was calculated.

Six sets of simulations were run, each employing a seperate sub-population of shares derived from the overall population. Each sub-population was selected from the overall population with the guidence of a distinct rule. The following rules were used to create the various six portfolios: firstly portfolios consisting of shares from the entire sample were created and analysed. Secondly the selection of shares was limited to only those that were in the Financial and Industrial sectors of the JSE. Thirdly the portfolios were simulated using only the shares listed in the *Financial Mail's* "Top 100" top performer companies. Fourth, portfolios were made to consist of only the shares listed as the largest 150 companies as listed in the *Financial Mail's* "Top 150" Market leaders by market capitilisation. Finally shares were seperated according to their beta with shares with a beta above 1.1 in 1993 being classified as high beta while shares with a beta value lower that 0.5 were classified as low beta; thereafter the portfolios were created through random simulation of each sample population as explained above respectively (Neu-Ner & Firer, 1997).

The results showed that using the first set of simulations, where the entire available population of shares was utilised, the expected risk associated with holding a single share can be reduced by 25% by holding two shares and by 50% if a portfolio consisting of six shares is held. Increasing the number of shares in a portfolio to 10 reduces the risk by 60% and increasing the number of shares in a portfolio to 25 decreases the risk by 70%. The maximum reduction of risk by diversification is 80.5% however this can only be reached utilising a portfolio consisting of over 200 shares.

These findings were relatively in line with those of Bradfield (1993), although Bradfield (1993) did not investigate portfolios greater than 50 shares. The percentage of total risk that was diversified in a random 50 share portfolio was 25.2% while Bradfield (1993) found it to be 25.3%. The period studied by Bradfield (1993) spanned from January 1988 to December 1992; the difference between this period and the period studied by Neu-Ner and Firer (1997) may explain the slight differences found as different securities may have been observed and the relationships between the securities could have changed across the periods. Both Bradfield (1993) and Neu-Ner and Firer (1997) found the the average weekly return did not depend on the number of shares included into the portfolio.

When the risks were converted to variance terms they can be compared to the results of Elton and Gruber (1977). Elton and Gruber (1977) found that the ratio of portfolio variance of a 1000 share portfolio to the variance of a single share was 15%; in essence this indicates that a portfolio consisting of 1000 shares reduces the variance experienced by the portfolio by a measure of 85% when compared to a portfolio consisting of a single share. Furthermore, the reasoning above can be perceived to postulate that 85% of risk in the above portfolio is diversifiable. Neu-Ner and Firer (1997) found that 96.6% of the variance of a single share was diversifiable leaving only 3.4% of non-diversifiable market risk. This shows that a more significant reduction in risk was found by Neu-Ner and Firer (1997).

The results found by the simulations run on the Financial and Industrial sector of the JSE showed slightly lower risk than the previous set of simulations where the shares of all industries were included. A portfolio with a single share from the Financial and Industrial sector had an expected risk of 6.15% compared with 6.73% for a share selected at random from the entire sample. This finding continued to grow proportionally as the number of shares in a portfolio increased; utilising shares from the Financial and Industrial sector offered a substantially better risk reduction benefit as an increasing number of shares was added to a portfolio (Neu-Ner & Firer, 1997).

In the next set of simulations run by Neu-Ner and Firer (1997), where the sample population was limited to *Financial Mail's* list of "Top 100" performer companies in June 1993. The aim was to assess the impact of limiting the choice set to the best companies in terms of share price performance over the previous five years. The results showed that portfolios constructed using these shares exhibited much lower risk. Compared to 6.73% of risk associated to holding one share in a portfolio in the first set of simulations, the same portfolio under the above criteria now showed the risk associated with holding one share to be 4.51%. Similar reductions of risk were observed as the number of shares to be included in a portfolio increased. Once again it was found that increasing the number of shares in the portfolio was independent of the returns of the portfolio, although the average return of the respective portfolios was lower when compared with the first set of simulations (Neu-Ner & Firer, 1997). The results when the population of shares were limited to the top 150 largest companies in terms of market capitalisation were found to be in line to the previous set of simulations run (the "Top 100" performers) with the exception that the average return was lower (Neu-Ner & Firer, 1997).

Finally, the study led by Neu-Ner & Firer (1997) analysed the results when the population was restricted based on the sensitivity of the shares to the market (based on the share's beta). When shares were exceptionally sensitive to market movements (had a beta of above 1.1), portfolios exhibited a higher level of risk for each number of shares included in a portfolio. The expected risk of holding a portfolio of one share increased from 6.73% to 8.79%. This finding continued as more shares were included in the portfolio. Interestingly, this increase in risk was not compensated for by increased returns.

Shares with a relatively low sensitivity to market movements (having a beta of below 0.5) showed lower levels of risk; with the risk of holding one share reduced to 5.98%, with similar reductions occuring as the number of shares in a portfolio increased. The benefit to the portfolio was further emphasised as the decrease in risk coincided with higher expected returns of 0.73% per week instead of 0.69% as in the first set of simulations. It was shown once again that the expected returns were independent of the number of shares to be included in a portfolio.

In essence Neu-Ner and Firer's (1997) findings agreed with those of Statman (1987), that a well diversified portfolio should include at least 30 shares. It was further established that significant benefits of diversification could be achieved by holding smaller portfolios; holding a portfolio of ten shares reduces risk by nearly 60% when compared to the average risk of holding a single share. Increasing the number of shares in a portfolio from 10 to 30 reduces risk by a further 12%. The second key finding was that as the number of shares in a portfolio increases, the risk associated with the portfolio decreases too. Lastly, the study found that the average weekly return on the portfolio is independent of the number of shares included in the portfolio.

An additional feature noted by Neu-Ner and Firer (1997), is that the magnitude of benefits an investor can obtain through the employment of diversification strategies is dependant on the type of 'rule', if any, applied to the population of shares. The table below highlights this feature - it can be observed that depending on which selection of shares was utilised to represent the population (based on the sector, past performance and beta of the share), differing amounts of risk could be diversified out of the portfolio. The findings between the different sets of simulations is summarised in the table below:

Table 2: *Results of Simulations Run by Neu-Ner and Firer (1997)*

Selection Rule:	Portfolio Risk	Portfolio Return	Summary
Financial and Industrial	Slightly Lower	No Change	No Change
Top Performers and Market Leaders	Lower Risk with few shares and converges with many	Much Lower	Lower risk with few shares coupled with lower return
Large Beta	Much larger systematic risk irrespective of the number of shares	No Change	Higher Risk
Small Beta	Lower systematic risk no matter the number of shares	Slightly higher	Lower risk with the possibility of a higher return.

Note: Adapted from "The Benefits of Diversification on the JSE", by Neu-Ner and Firer, 1997, *Investment Analysts Journal, 46, 57*

The current study aims to advance the methodology of Neu-Ner and Firer (1997) in the following ways: Firstly, shares that were listed or de-listed over the time period are still included into the sample period as to avoid survivorship bias, which refers to the bias that occurs when only live securities are incorporated (Gregoriou, 2008). Surviorship bias leads to distorted results as it is common that the securities that delist do so due to extremely poor performance; albeit that delisting can also occur for other reasons such as privatisation. If these shares are not assimilated into the population it creates a distorted environment wherein achievable returns, benchmark returns and the effects of diversification are all misrepresented. In addition, diversification generally provides significant benefits in scenarios where a single, or a few, shares perform particularly poorly. These benefits would be overlooked without the inclusion of shares that list/delist over the observed period (Swensen, 2009). Secondly the role of dividends has been included into the return on shares calculated. Thirdly, the best iteration is utilised when observing the risk reduction benefits of a portfolio rather than the average of all iterations as was used in Neu-Ner and Firer (1997). Finally, simulations are run on both portfolios when the constituents are equally weighted as well as on portfolios where the individual weights on shares are allowed to vary within given constraints. These improvements are expected to find further risk reduction benefits when investing on the JSE as well as serving to update the study done by Neu-Ner and Firer (1997) for the current period studied.

This study aims to carry out the analysis of the benefits of diversification while applying the advancements mentioned above as well as across three methodologies, as opposed to the analysis of diversification under various population selection rules as in Neu-Ner and Firer (1997). In addition to the traditional simulation approach, this study employs the methodology of Genetic Programming and Particle Swarm Optimisation; these approaches are discussed in turn in the sections to follow. Furthermore, the increased computing power currently available compared to that which was available at the time of Neu-Ner and Firer's (1997) paper is expected to provide better simulated results than those presented by Neu-Ner and Firer (1997). The methodology of Genetic Programming and Particle Swarm Analysis is expected to further positively effect the diversification benefit for an investor. This is due to the approaches being more advanced and computational in nature, as opposed to the method of chance employed in random simulation. The chapter continues now by exploring the second method of diversification, relating to product differentiation by a firm.

2.1.2 Diversification by Product/Service Offering

The second method of diversification relates to when a firm attempts to differentiate their product (service) offering. Although this line of thought is traditionally kept separate from diversification as it relates to share portfolios, many findings from this second method (such as those that will be discussed below) can be applied to the first. Furthermore it is important to include the effects of product diversification as it relates to firm performance (and share price performance) to discover if any diversification or performance benefits exist to be gained when a firm diversifies its own product offering.

Ansoff (1957) stated that diversification of a product offering can be viewed as a 'growth alternative' for a company. There are several reasons a company may select to utilise this strategy such as to: distribute risk, utilise excess productive capacity, reinvest earnings, obtain top management and the like. In addition Comment and Jarrell (1995) suggested that some firms may diversify their product offering or look to engage in acquisitions as a way to increase the performance of their core business – they could branch out if they are performing poorly in a certain sector.

Furthermore, Ansoff (1957) pointed out that there are certain conditions, which were referred to as contingencies, which make diversification particularly desirable. These are environmental conditions that cannot be predicted, but if they occur will greatly affect sales and thus firm performance. Examples of contingencies are technological breakthroughs, recessions, wars as well as economic and consumer trends. The effectiveness of diversification relies on both external factors (such as how firms and their competitors will perform in addition to contingencies) as well as internal factors (such as the ability to diversify, capital allowances and risk tolerance) (Ansoff, 1957).

Rumelt (1982) demonstrated that the product diversification strategy that had the largest positive impact on firms financially were strategies that entailed the firm diversifying into areas that drew on some common skill or resource that the firm had already acquired or had control of. In this regard, it was concluded that firms would perform better when diversifying if they stay within the bounds of their core competencies or if they are willing to hire experts to be actively involved in the management of the diversified section of the business.

Lang and Stultz (1994) found weak evidence to support that diversified companies perform relatively poorer compared to their specialised counterparts. Morck, Shleifer, and Vishny (1990) found that in some cases the market reacts negatively to unrelated acquisitions, however in other instances a positive reaction is experienced. A positive reaction was generally experienced if it could be shown that the newly diversified firms could easily be dismantled and run separately. This would lead to investors believing that shares in the parent company would now offer similar benefits to investing in two separate specialised companies (Comment & Jarrell, 1995).

Interestingly however, if an investor were concerned with the effectiveness of the diversification of their portfolio, they would be indifferent between purchasing shares of one company that has diversified their own operations and between two companies where both have chosen to specialise and thus not diversify their own operations. Transaction costs may however sway the investor to purchase the share of the company that is diversified themselves so as to only incur these transaction costs once. This would however be offset by the ability of one company to compete with two companies, both specialising in where their core competencies lie – this follows the same principles as examined by Markowitz (1952b).

Examining both methods of diversification reveals a key common warning with regards to diversification. That is: extreme care that must be taken so as to be not overly diversify or what Francis (1986) refers to as having "superfluous diversification". With regards to the diversification of a share portfolio, Ansoff (1957) cautions investors against spreading their portfolio too thin as this will lead to no real gains being felt in any sector. Rumelt (1982) makes a similar warning to firms with regards to product diversification; he states that when a firm desires to diversify their offering the benefits of diversification must be weighed off against the cost of no longer being industry experts. These warnings highlight the importance of finding how many shares make an optimally diversified portfolio as both under- and over-diversification entails costs associated with them, many of which are unnecessary as they could have been avoided. The Literature Review section continues below with a focus on the methodology underlying genetic programming, accompanied by an exploration of instances where the approach has been previously applied.

2.2 Genetic Programming

The following section elaborates on the fundamental techniques and concepts key to understanding the methodology underpinning genetic programming. The section begins by describing the origins of genetic programming and continues by exploring the two theoretical approaches that form the foundation of genetic programming, namely from the neurological perspective and thereafter from the evolutionary perspective. The study continues by expounding upon the mathematical processes and respective constructs employed by genetic programming and thereafter follows by elucidating upon the advantages and disadvantages of the approach. The section then concludes with a thorough review of the implementations and applications completed by previous literature.

The methodology of genetic programming is an applied extension that is founded upon the concepts underlying genetic algorithms that were initially developed by Holland (1962). Genetic programming and a genetic algorithm are dissimilar in their technical aspects as an algorithm provides an answer in a symbolic form as opposed to a neural network or solution given by a genetic program, which represents an answer in a numerical form (Allen & Karjalainen, 1993). The methodology behind both however is extremely similar and as such the terms will be used interchangeably in the section below.

Genetic algorithms refer to a set of mathematical techniques that are based on biological principles that mimic concepts such as natural evolution/selection (Miles & Smith, 2010) as well as the neurological functioning of the brain (Abraham & Nath, 2000), both of which will be further expanded upon below. Within the broad definition of genetic algorithms, a potential solution can be represented as a set of parameters, termed genes. These genes can then be joined together to form strings of values, known as chromosomes. By allowing the genetic algorithm to follow these processes, one can apply a structured architecture with learning and generalisation capabilities which in turn are able to evolve with the aim to compute and optimise real world problems (Abraham & Nath, 2000). The above terms and definitions as well as the process that genetic programming involves is discussed further in the sections below.

Miller (1986) further developed the work of Holland (1962) by developing an adaptive model of economic behaviour. Miller (1986) overcame the critiques against the utilisation of the biological approach, which concentrated on the argument that the models at the time focused on equilibrium conditions as well as the use of assumptions regarding the abilities of individual agents. This was accomplished through the creation of a model that relied on specific biological concepts. This structured methodology led to reduced ambiguity which confounded previous biological based approaches. Through the use of two major assumptions Miller (1986) was able to overcome the previously mentioned concerns aimed at biological models at the time. Firstly, it was assumed that some economic behaviour has close analogues to the biological behaviour of ecosystems and species and secondly, that this behaviour can be effectively modelled and analysed (Miller, 1986). The first tenet was relatively well recognised by economists at the time and can even be found in instances as far back as 1798 (Miller, 1986), however, it was not until specific biological processes, which incorporated both economic and psychological theories of behaviour, were utilised that one could infer important properties of the model by deriving them theoretically. This approach allowed for a more highly efficient adaptive model to be developed as compared to the one initially used by Holland (1962). The approach used by Miller (1986) is further expounded upon below.

Miller (1986) noted that any system with an adaptive process could be described in terms of three major elements: allowable structures, its environment and an adaptive plan. It is observed that in an adaptive system, individual actors interact with the environment through various characteristics that are described by structural forms. These are guided by factors such as purchasing patterns, decision rules and genetic forms among other factors relevant to the problem at hand. At every time period the environment determines the performance of the various structures and thereafter provides a given level of information back to the actor. When using this model, no assumptions are necessary with regards to the stability of the environment nor the level of information presented to the actor – although some degree of feedback is required. Finally, the dynamic interactions between the environment and the structural forms are determined by the adaptive plan which governs changes in the existing set of structures from one time period to the next based on the available information. One therefore only needs to be able to define the environment, structures and the adaptive plan in order to be enabled to use an adaptive model. This is not however simple as it is necessary to make inferences regarding which elements of the structure, and consequently of the environment, are important to the analysis. The elements of choice may include the information on hand, computational ability and available time (Miller, 1986).

Miller (1986) further explains the model by describing that the economic agents follow an adaptive plan which modifies a small set of memorised behaviours based on the relative performance of each behaviour in previous evaluations. The model further assumes that behaviours are composed of individual building blocks and that the individuals reproduce and recombine themselves based on past successful actions. At any given time a behaviour is undertaken, the agent receives a known payoff and thereafter at any given time period the agent maintains a small subset of previous past behaviours. The retained (learned) subset is then used to develop new behaviours through the constructs of crossover and mutation. These allow the model to exploit information gained about the structures in previously run iterations while simultaneously exploring new structures (Miller, 1986). These constructs will be further expound upon in the sections to follow.

Miller (1986) applied the developed adaptive model to a variety of contexts including consumer demand behaviour, decision making under uncertainty, market structures, technological change and economic demography. In all of the examples the adaptive model was shown to be a valuable approach to modelling economic and social phenomenon and the results were demonstrated to be extremely promising. The results indicated that given enough time and proper conditions, agents' behaviour tend towards optimal patterns and that adjustments to new environmental conditions is initially very rapid but quickly slows as the behaviours approach optimal levels. Consequently, Miller (1986) demonstrated that various parameter values may generate testable hypotheses using an adaptive modelling approach, and furthermore, the application to varied problem sets in the paper confirmed the flexibility of the model to various contexts.

Miller (1986) developed upon and advanced the work of Holland (1962). As previously mentioned, developments on the biological approach since the work of Miller (1986), specifically within the framework of genetic algorithms, can be described as being based on two major themes: firstly, those which concentrate on mimicking neurological, cognitive and organisational processes of the human brain and secondly, those which aim to mimic the processes involved in evolutionary biology (Shachmurove, 2002). Each of these two themes and their relation to genetic algorithms will be expanded upon below, commencing with the neurological perspective.

2.2.1 Neurological Framework

The brain consists of billions of functional elements called neurons (specialized cells which transmit information across the central nervous system). An individual neuron consists of a cell body, dendrites and an axon (Gerstner, Kistler, Naud, & Paninski, 2014). A graphical representation of an individual neuron is displayed below:

Figure 1: Representation of a Single Neuron

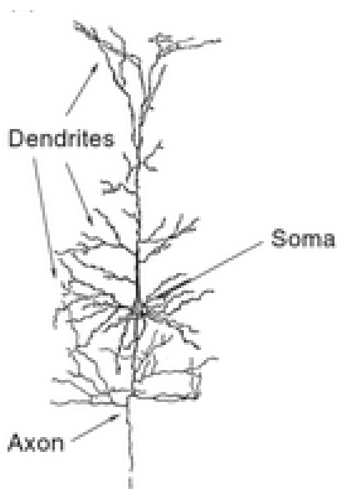


Figure 1: *Drawing by Ramon y Cajal, extracted from:* "Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition", Gerstner, Kistler, Naud and Paninsk, 2014

Each individual neuron receives electrical stimuli from other neurons surrounding it through connectors called dendrites. Dendrites can be thought of as the input device that collects signals from other neurons and transmits them to the soma. The soma refers to the cell body and is equivalent to the central processing unit of a computer, essentially performing the processing step. This step consists of testing whether the total input arriving at the soma exceeds a certain threshold (the action potential) or not. If it does not, then the cell does not 'fire' or send any further information. If however the input exceeds the action potential, an output signal is generated and control of the process is then assumed by the output device, referred to as the axon. The axon then delivers the signal to the surrounding neurons and so the process continues (Gerstner, Kistler, Naud & Paninski, 2014).

The signals that are sent between neurons consist of short electrical pulses. The space between two neurons that is used to transfer signals between two neurons is called the synapse (Shachmurove, 2002). The neuron that sends a signal across the synapse is called the presynaptic cell while the receiving neuron is called the postsynaptic cell (Gerstner, Kistler, Naud & Paninski, 2014). Neurons are embedded in networks of countless other neurons that are used to store information and learn meaningful patterns. This is achieved by strengthening the interconnections between neurons (Shachmurove, 2002; Gerstner, Kistler, Naud & Paninski, 2014). A group of interconnected neurons is displayed graphically below:

Figure 2: Representation of a Group of Neurons

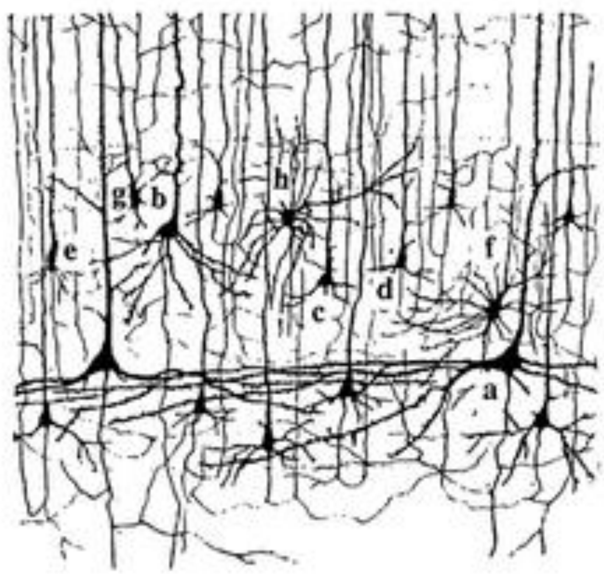


Figure 2: *Drawing by Ramon y Cajal, extracted from:* "Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition", Gerstner, Kistler, Naud and Paninsk, 2014

This process, modelled mathematically, is the basis of genetic algorithms and neural networks (Shachmurove, 2002). The neural network is characterised by the pattern of connections among various network layers, the number of neurons in each layer, the learning algorithm and finally the neuron activation functions, each of which will explained further below. The neural network is essentially a set of input and output units with each connection having an associated weight. The system begins with the learning phase which tweaks and adjusts the weights with the aim of correctly predicting or classifying the output target. It then continues to optimise the output process with utilising various forms of feedback (Thawornwong & Enke, 2004).

2.2.2 Evolutionary Framework

The second approach to the genetic algorithm process is based on natural selection, as first described by Charles Darwin in *"The Origin of Species"* (Darwin, 1859). This follows the idea that over generations, certain individuals of a species will survive due to some trait that they have whilst others, who do not possess the trait, will cease to exist (Gibson & Gibson, 2009). Holland (1962), as previously mentioned, invented the methodology behind genetic algorithms. In traditional genetic algorithms, genetic structures are represented as character strings of a fixed length. Although adequate across an array of applications, this is restrictive when the size or form of the solution cannot be assessed beforehand. Koza (1992) extended the methodology by changing the processes put forward by Holland (1962) to allow variable length representation of genetic structures. Following this ideology, genetic programming can be seen to comprise of three main sections (as described by many authors including Allen & Karjalainen, 1993). The three main processes that make up genetic programming are: search, adaptation and finally optimisation. These can be split into five further components, namely population, evaluation, selection, crossover and mutation. These processes and their respective components are elucidated upon below.

Although a genetic algorithm and genetic programming are dissimilar in their technical aspects (an algorithm provides an answer in a symbolic form as opposed to a neural network or solution given by a genetic program, being represented in a numerical form (Allen & Karjalainen, 1993); they can best be described by expounding upon a similar idea. Kaastra and Boyd (1996) discussed an eight step process to design a neural network, however genetic programming can be most easily understood by elucidating upon the framework laid down by Allen and Karjalainen (1993).

As previously mentioned, genetic programming is inspired by the theory of natural selection in order to allow one to generate new candidate solutions (Miles & Smith, 2010). The process begins with data being input; thereafter a population of randomly generated solution possibilities is formed. The next generation of solutions is created by recombining promising candidates from the first (previous) generation. This process is called crossover which entails randomly selecting two parents from the population; selection is biased towards selecting parents that are relatively fit – that have a seemingly strong explanatory power (Allen & Karjalainen, 1993). The parents are subsequently split at randomly chosen locations and then combined by joining a portion of genetic material from each parent. This results in a new unit in the next generation with a new genetic structure (Allen & Karjalainen, 1993). This effectively refers to the mixing of subtrees in a population. A second operator, known as mutation, is also used in genetic programming. Mutation refers to replacing subtrees with new randomly generated subtrees (Miles & Smith, 2010). This essentially alters each gene with a small probability, introducing a small proportion of randomness into the model. Traditionally crossover was viewed as the more important of the two techniques as it allows for rapid exploration of a search space (Allen & Karjalainen, 1993). This is however somewhat offset as mutation aids to ensure that no point in the search space has a zero probability of being examined. This aims to guarantee that the algorithm does not quickly become trapped in a local optimum (Allen & Karjalainen, 1993).

This new unit is then tested and, if it has a relatively strong fitness level, replaces one of the relatively unfit members of the parent population. This process continues until the generation is complete and will continue to occur to until a termination criterion is satisfied. The final result is a collection of relatively fit solution candidates of which one or more can be applied to the original problem (Holland, 1975).

The exact subset of genetic programming that will be utilised is called a flexible neural tree model. In this model, solution candidates are represented as hierarchical compositions of functions. In these tree like structures the successors of each node afford the arguments for the function to the corresponding node. The terminal nodes (bottom level) refer to the input data. In this way the entire tree is interpreted as a function which can be evaluated by working successively from the bottom layer to the top. Furthermore, the structure of the tree is not specified beforehand but rather is solved for as part of the process within the algorithm (Allen & Karjalainen, 1993).

The flexible neural tree model follows a specific sequence of events. This entails beginning with random structures and corresponding parameters (Chen, Yang, Dong & Abraham, 2005). Secondly, the structure is improved and once an improved structure is found the model fine tunes the parameters (Chen et al., 2005).

This process continues until a satisfactory solution is found (Chen et al., 2005). This is displayed graphically by Figure 3 below:

Figure 3: Representation of a Neural Tree

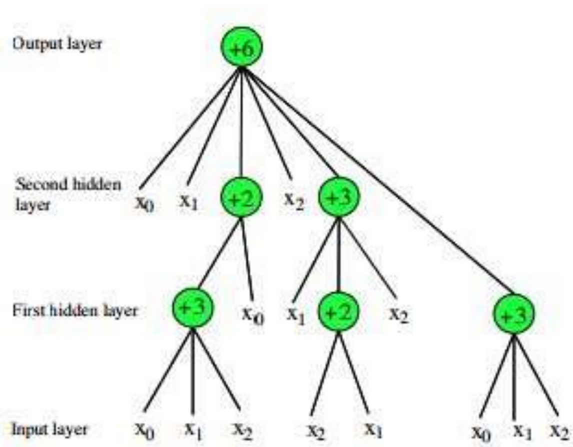


Figure 3: *A typical representation of neural tree with three instruction sets and three input variables x0, x1 and x2* (Chen, Yang & Dong, 2004).

The tree structure optimisation is governed by the following set of processes; firstly, one can change one terminal node (the terminal node refers to the bottom (first) layer of the tree structure) by randomly replacing it with another terminal node. The second option is to change all the terminal nodes in the neural tree and replace each with another terminal node. Thirdly, one can select a random 'leaf' in a hidden layer (as shown in Figure 3) and replace it with a newly generated sub-tree. The final option is termed pruning and refers to selecting a functional node (a node that is a function of other nodes or variables) and replacing it with a terminal node (Chen, Yang & Abraham, 2007). Each of these processes is used in order to optimise the tree structure.

The parameters are optimised by using three techniques. The first technique involves all variables being selected (or combinations thereof) with equal probability of surviving into the next strain of the tree. The variables that have a greater contribution towards forecasting (and thus explanatory) power are then be enhanced and have a relatively higher probability of surviving further generations. Finally, various evolutionary operators are used to govern how the model selects the appropriate variables to include into the model automatically (Chen et al., 2007). As previously described, tree structure optimisation and parameter optimisation continue until a satisfactory solution is found. This process ensures good individual performance of variables as well as for combinations thereof and consequently removes the effects of multicollinearity to a large extent. (Chen et al., 2007).

It is important when implementing a neural algorithm to be mindful of the Bias/Variance dilemma. Bias and variance are two useful concepts in characterising the generalisation behaviour of learning algorithms (German, Bienenstock & Doursat, 1992). Bias refers to the systematic component of the generalisation error while variance refers to the additional error incurred due to over-responsiveness of an algorithm to random fluctuations. Chen, Yang, Dong and Abraham (2005) describe the dilemma by elucidating that members of the population must be both accurate and diverse. This poses a new dilemma whereby when a set of predictors is generated they are required to have reasonably good individual performances and independently distributed predictions for test points (Chen, Yang, Dong & Abraham, 2005). The Bias/Variance dilemma can be stated as follows: models with too few parameters are inaccurate due to a large bias; in other words they lack flexibility as the number of parameters is limited. In addition models with too many parameters are inaccurate because of a large variance. These models are overly sensitive to the sample details and as such changes to the sample will cause vast variations in the results. Finally, the third step involves solving for the appropriate model complexity. This involves the proper architecture and number of parameters (Battiti & Brunato, 2014).

The third step should follow the principle of *Occam's razor.* Occam's razor is often considered one of the fundamental tenets of modern science (Domingos, 1999). In its original form Occam's razor was translated to mean 'Entities should not be multiplied beyond necessity' (Tornay, 1938). In essence Occam's razor states that given two explanations of the data, with all other things being equal, the simpler explanation is preferable. This principle asserts the notion that one should avoid using overly complicated models in favour of using simpler models with similar results (Blumer, Ehrenfeucht, Haussler & Warmuth, 1987).

It is imperative to follow careful experiment procedures to measure the effectiveness of the learning process, a specific caveat to be mindful of is to not test the performance of an algorithm on the same example data that was used for training (Battiti & Brunato, 2014). As mentioned above, the objective of using methods of machine learning is to obtain a system capable of generalisation to new and previously unseen data. If the genetic algorithm is not capable of or efficient in this, it cannot be seen to be ‘learning; but merely ‘memorising’ the set patterns found in the data (Battiti & Brunato, 2014).

Time series analysis is one of the most widely used traditional approaches in finance and economics modelling. In general time series analysis possesses the following traits: data intensity, an unstructured nature, encompasses a high degree of uncertainty and finally contains hidden relationships (Huang, Lai, Nakamori, Wang & Lean, 2007). There are two models that can be used to describe the behaviour of time series data, linear and nonlinear models (Huang et al., 2007). Linear models take a linear approach to time series analysis and are typically applied through Box-Jenkins techniques, Kalman filters, Brown’s theory of exponential smoothing or piecewise regression. Each technique relies on transforming the data from the time series into a linear function that is then used to forecast future values. Nonlinear models take a nonlinear approach to the time series analysis and are typically applied through one of the following techniques: Takens theorem, the Mackey-Glass equation and neural networks (genetic programing and algorithms) (Huang et al., 2007). It has become widely accepted that through the use of intelligent systems, human like expertise can be modelled (Abraham & Nath, 2000).

2.2.3 Advantages and Disadvantages

Genetic programming has various advantages as well as disadvantages. Miller (1986) noted that the framework places small demands on the abilities of economic agents whilst also being inherently dynamic in nature, thus avoiding the need for the system to be in an equilibrium state in order to demonstrate good performance. This allows the model to be less sensitive and consequently increases the ability to find accurate solutions in environments that contain error term assumptions, noise, and chaotic (unpredictable) components (Miller, 1986). Furthermore, the methodology does not make assumptions regarding the nature of the distribution of the data (Shachmurove, 2002). This allows the system to also display good performance in the presence of abnormal distributions of data, of which mainly non-linear relationships is particularly useful (Kaastra & Boyd, 1996; Huang et al., 2007; Shachmurove, 2002).

Shachmurove (2002) highlights that since economic, financial and social systems are complex and subject to human reactions and counter-reactions by different agents or players, it is difficult, if not impossible, to model such a system from first principles that incorporates all potential reactions and counter-reactions. In such systems it is extremely beneficial to utilise models which emulate and simulate the economy or society in question; this is what the neural network methodology is capable of delivering. The genetic methodology has been shown to be powerful for uses such as pattern recognition, classification as well as forecasting (Kaastra & Boyd, 1996; Huang et al., 2007). It has the ability to analyse complex patterns quickly and with a high degree of accuracy (Shachmurove, 2002). In addition genetic programming allows one to search extremely large rule spaces while allowing a multitude of potential rules to be tested in a practical manner (Allen & Karjalainen, 1993). It has been shown that a network can approximate continuous function to any desired accuracy and is still able to perform relatively well with missing or incomplete data (Huang et al., 2007; Shachmurove, 2002). Genetic programming's main advantage however has been shown to lie in the domain of problems that cannot be solved easily, if at all, using classical techniques (Miles & Smith, 2010).

Conversely, genetic programming has a few disadvantages and is subject to various limitations. There is no common framework or structured methodology for choosing data, development, training or verifying results (Shachmurove, 2002). Thus, output quality may be unpredictable. In addition, genetic programming is data dependant and as such any optimisation or prediction performance will be vastly affected by differing input data (Huang et al., 2007). Changes of a model’s constraints and parameters would also lead to a marked difference in the quality of output (Wagner & Brauer, 2007). Furthermore, excessive training times may be needed in certain situations (Abraham & Nath, 2000). Genetic programming is also critiqued due to its ‘black box’ nature which refers to the fact that it is extremely difficult if not impossible for a user to calculate how relations in the hidden layers are estimated and applied (Goonatilake & Treleaven, 1995). In addition, the concept of ‘previous knowledge’ is difficult to incorporate as the model gains its strength based on how this knowledge is represented (Abraham & Nath, 2000). A neural network’s performance is also highly dependent on its structure as the interaction allowed between the various nodes of the network is specified using the structure only (Chen, Yang & Dong, 2004). The results also have two shortcomings: They frequently do not provide analytical solutions that are provided by other models (Miller, 1986) and secondly, genetic programming may be ‘overzealous’ and attempt to over-fit or under-fit data (Gilbert, Krishnaswamy & Pashley, 2000 as cited in Shachmurove, 2002). It is always possible to model a mathematical function that perfectly represents the historical data of a time series but this limits the predictive capability of the model and leaves it with very little, if any, generalisation capacity.

The section above expounded upon the central techniques and concepts that underpin the methodology of genetic programming. The subchapter continued by outlining the two thought paradigms that can be used when analysing the methodology of genetic programming: the neurological approach and the evolutionary approach. The way in which genetic programming overcame the initial critiques brought against biological approaches at the time it was initially discovered were discussed and thereafter the methodology behind genetic programming was described and analysed. It was indicated that genetic programming has become widely accepted and that through the use of intelligent systems, human like expertise can be modelled. The subchapter then concluded with a summary of both the advantages and disadvantages of using genetic programming. This dissertation now continues by further investigating previous literature where the methodology of genetic programming was applied to a variety of concepts. These were examined through the analysis of their specific methodologies and findings respectively.

2.2.4 Implementations of Genetic Programing

The methodology and concepts that underlie genetic programming was examined in detail in section above however it is interesting to note, in this section, the great deal of evidence displaying the effectiveness of genetic programming when applied to financial data sets. As mentioned previously, genetic programming refers to a set of mathematical techniques that are based on biological principles that mimic concepts such as natural evolution/selection (Miles & Smith, 2010) as well as the neurological functioning of the brain (Abraham & Nath, 2000). Genetic programming and a genetic algorithm are dissimilar in their technical aspects, an algorithm provides an answer in a symbolic form as opposed to a neural network or solution given by a genetic program, being represented in a numerical form (Allen & Karjalainen, 1993). However, the methodology that underlies both, as discussed in previously, is extremely similar and as such the terms will be used interchangeably. The following section outlines the results yielded from previous papers where a genetic programming approach was utilised.

Allen and Karjalainen (1993) used a genetic algorithm approach to establish technical trading rules to model the daily price movements of the S&P 500 composite stock index over the period 2 January 1963 to 29 December 1989 (technical analysis refers to the use of past prices, trading volumes, and other backward looking variables to forecast future price changes). Despite much academic scepticism, technical analysis is still widely used in practice (Zhu & Zhou, 2009). The rules found were then compared to a buy-and-hold strategy as well as the benchmark models of a random walk and an autoregressive model. A random walk model can be defined as a model wherein stock price changes have the same distribution and are independent of one another. In this manner, past share prices cannot be used to predict future share prices (Malkiel, 2015). Conversely, an autoregressive model refers to a stochastic process in statistical calculations where future values (share prices) can be estimated by previous values (Roehner, 1995). Furthermore, conventional statistical tests and bootstrapping simulations were carried out to study the robustness of the results.

In the paper by Allen and Karjalainen (1993), the training set (referring to the set of data used to discover potentially predictive relationships as well as the strength and utility thereof) was made up of the daily data over the years 1964 to 1967 while the validation set was made up of the daily data over the years 1968 to 1969. From each trial, one rule was saved and then tested during the years 1970 to 1989. A population size of 500 was selected and the genetic structures were limited to 100 nodes and had a maximum of 10 levels of nodes. Evolutions were then allowed to continue for a maximum of 50 generations or until there was no improvement shown for 25 generations. 100 independent trials were then carried out using the above mentioned parameters, each being initialised from a different random population.

The technical trading rules established were compared to a buy-and-hold strategy and were found to demonstrate positive excess returns; in addition the results indicated a reduced variability of the returns. The excess returns were both statistically and economically significant but they did however experience slight declines in significance after trading costs were taken into account.

Moody and Saffell (2001) made use of genetic algorithms that aimed to optimise portfolios, select asset allocations as well as to develop trading systems. In their work they viewed investment decision making as a stochastic control problem where the investor's ultimate goal is to optimise a relevant measure of performance, such as profit. A direct reinforcement method was utilised, which gave the system the ability to bypass the need to learn a value function (which refers to a mathematical condition that relates each input value to a corresponding output value), which thereby allowed the system to trade a single security as well as to manage a portfolio and allocate assets.

Moody and Saffell (2001) used an innovative method of viewing the investment process, seeing investment performance as depending upon the sequences of interdependent decisions and as such being path dependant. In this way it is essential for a trading system to take into account the current system state, which includes both current market conditions as well as the currently held positions. In addition the authors took transaction costs (such as commissions, bid/ask spreads as well as others) into consideration, noting that arbitrarily frequent trades or large changes in the portfolios' composition led to excessive costs to the portfolio (Moody & Saffell, 2001).

In their analysis they utilised fixed trade sizes while the algorithm adjusted the parameters of the system in order to maximise the Sharpe ratio (a reference to the risk adjusted returns) of the portfolio (Moody & Saffell, 2001. The system was applied with the intention of discovering tradable structures in the intraday trading of the USD/GBP foreign exchange rate as well as to find value investment methods based on the combination of the S&P 500 stock index and treasury bills (applied monthly). It was found that over the test period of 25 years, in both cases the system discovered a predictable trading structure and showed that the method dismisses the need to build better forecasting models in order to achieve increased trading performance (Moody & Saffell, 2001).

Wagner and Brauer (2007) applied a dynamic forecasting version of genetic programming to examine the relationship between US GDP and its various determinants, including military expenditure. In addition, the length of the time series was allowed to be automatically discovered by the genetic algorithm. The FRED database was utilised from the Federal Reserve Bank, from which quarterly data was analysed, ranging over the period 1947 to 2000. Wagner and Brauer (2007) aimed to rather find which variables demonstrated importance with regards to explaining the behaviour of GDP in the past rather than attempting to discover which variables could be used to predict future GDP.

The results achieved by Wagner and Brauer (2007) were compared to a regression based forecasting model. Regression refers to a relatively simple statistical method used to determine the relationship between a dependent variable and one or more independent variables (Gelman & Hill, 2007). The results showed no clear outperformer between the dynamic forecasting approach and the regression based model - it was found that the dynamic forecasting version gave mostly comparable and relatively equivalent results. The models did however each have their own assumptions. The dynamic method was not allowed to make any prior assumptions with regards to the functional form or the time span of data series that was used as an input while the regression based model assumed a linear functional form and that the entire time span represents a single data generation process which was perceived to be mostly unrealistic. Additionally, their results indicated that when the models were used on undifferentiated data the dynamic based model outperformed (Wagner & Brauer, 2007).

Thawornwong and Enke (2004) used neural networks in order to find relevant variables that displayed strength with regards to forecasting the direction of stock returns. The resulting variables were then provided to the neural networks, including probabilistic and feed-forward networks with the aim of predicting the direction of future excess stock returns. The returns found were then compared to a risk-free return on a one month T-bill, as this is viewed as the minimum returns expected from depositing money into a risk free account, as well as three other models namely: the buy-and-hold strategy, the conventional linear regression strategy and the random walk model.

The neural network approach was selected as it demonstrates strong performance in its ability to model non-linear processes without *a priori* assumptions regarding the nature of the generating processes (Hagan, Demuth, Beale & De Jesus, 1996). This is particularly useful in finance and investment as it is common that little is known, but much is assumed, about the nature of the data generating processes that determine asset prices (Burrell & Forlarin, 1997). Furthermore this approach allows the user to input a large amount of data without the requirement of knowing beforehand which sections of the data are useful (Thawornwong & Enke, 2004).

The variables considered included 31 financial and economic indicators that were collected on a monthly basis over the period March 1976 to Dec 1999 (totalling 286 months). The data was then divided into four sliding window periods. Each of these sliding periods was further split into two period sets. The first period set of each sliding period was used for training and validating the forecasting models, while the second period set was reserved for out-of-sample evaluation and comparison of performance among the forecasting models. As an example, sliding period 1 would consist of a training and validation period from March 1976 to October 1992 and then would run an out of sample test on the period ranging from November 1992 to August 1994. The second sliding period slightly overlapped the first (thus deserving the name ‘sliding window’) and ranged from January 1978 to August 1994 in the training and validation period and the out of sample test was run on the period September 1994 to June 1996 (Thawornwong & Enke, 2004).

This analysis was completed with the aim of discovering the predictive effect of the relevant variables over various time periods as well as to establish the robustness of the out of sample forecasting performance. With this in mind the data was mined to determine which of the 31 variables has the most predictive ability (or combination thereof) and then those were used and input into the model. Trend and seasonal components were excluded from the variables considered as the neural networks attempted to learn the trends and use it in the prediction. The results showed that the neural network strategies outperformed buy and hold strategies, linear regression models and the random walk model as they generated higher profits with lower risks (Thawornwong & Enke, 2004).

Liu, Ng and Quek (2007) also highlighted the fact that time series prediction is traditionally handled by linear models such as autoregressive and moving average models. These however are inadequate when dealing with non-linear data. The authors used neural networks to determine the input dimension and time delay to predict the daily stock price of General Motors Corporation. These are viewed as the two critical factors that affect the performance of a neural network.

The input dimension refers to the number of delayed values used for prediction while the time dimension is the time interval between two data points. Liu, Ng and Quek (2007) used a reinforcement learning-based method to simultaneously determine both of these factors. Reinforcement learning refers to a learning scheme wherein the agent learns from the interaction between itself and its environment. The goal of reinforcement learning is to learn an optimal policy from past experience of the agent, thereby maximising the total amount of reward the system receives over the long run. An agent is the learner as well as the decision maker, in that it takes actions and receives feedback from the environment.

When analysing such a problem one must be cognisant that if the dimension selected is too small, the information for prediction of the next point may not be sufficient; if too large redundant information and noise may be brought into the forecast and over fitting could occur. Furthermore if the delay is too small, two adjacent data points may be too highly correlated but if it is too large, useful information in the delay period may be lost. In order to determine the time delay there are two principal methods. The first method determines the time delay at which the value of the autocorrelation of the data becomes zero for the first time. The second indicates the time delay based on the first local minimum of average mutual information. The proper choice of dimension and delay may be effected by both the prediction model and the data (Liu, Ng & Quek, 2007).

The results illustrated by Liu, Ng and Quek (2007) showed that their method, which was based on reinforcement learning, was able to achieve near optimal performance and selected a dimension and delay factor of 1. These were both shown to be the optimal choices. These predictors were not however used to build automated investment strategies, however more recent literature (as highlighted in this study) has included the building of trading strategies (with the output being a buy, hold or sell signal) through the use of genetic programming (Liu, Ng & Quek, 2007).

Jin, Tsang and Li (2009) utilised an evolutionary constraint-guided method algorithm to examine two economic problems. The first was to search for equilibriums for bargaining problems and the second was aimed at reducing the rate of failure in financial prediction problems. The constraint-guided method is capable of handling both hard and soft constraints within optimisation problems. While searching for constraint satisfactory solutions, the method differentiates candidate solutions by assigning them with different fitness values, enabling favoured solutions to be more easily and effectively distinguished from less favoured ones (Jin, Tsang, & Li, 2009).

This approach is based upon the idea that candidate solutions to an optimisation problem (with both hard and soft constraints) can be categorised into three qualitatively different sets: the infeasible, the feasible and the preferable. Hard constraints refer to conditions of the variables that are required to be satisfied for the solution to be valid. Soft constraints however refer to conditions of the variables which are penalised if they are not met based on the extent of them missing the acceptable criteria. Solutions in an infeasible set refer to solutions that do not satisfy all hard constraints while solutions in the feasible set are those that do satisfy all hard constraints, as well as some soft constraints (if not all). The solutions in the preferable set are considered as the best candidate solutions. These solutions satisfy all hard constraints and are ranked according how well they satisfy the soft constraints, in order of their priority (Jin, Tsang & Li, 2009).

An ideal search approach should be capable of not only identifying the feasible set among all possible solutions, but also distinguishing a preferable set from the feasible sets more easily and efficiently. This is the motivation behind the approach used by Jin, Tsang and Li (2009) which accomplished the above through using a carefully designed fitness function which incorporated problem-specific knowledge about hard and soft constraints directly. They compared their constraint-guided method against other computational techniques such as repair methods and penalty methods over both economic problems and found in both instances that their model compared favourably (Jin, Tsang & Li, 2009).

It must be noted however than when modelling bargaining or social construct problems, Arrow's impossibility problem applies (Arrow, 1963). Arrow's impossibility theorem relates to the central difficulty with focusing on individual preferences and the aggregation of those preferences into a general choice rule. If all individuals are in agreement, this may be possible, however, when there is disagreement this becomes vastly more difficult. In the case of an election, Arrow's theorem begins with a finite set of outcomes, being the potential candidates for office, a finite set of voters as well as their individual preferences over potential outcomes. If individual preferences are assumed to be an unrestricted domain and independent of irrelevant alternatives, relating to independence from other influences, then Arrow's impossibility theorem states that it is not possible to derive a complete and consistent social choice rule exclusively from individual preferences, apart from in dictatorships which are characterised by an inability to determine the intensity of preference (Arrow, 1963). This theorem is as applicable to bargaining and political problems as to threats of nuclear war and reaching new international trade agreements (Jin, Tsang & Li, 2009).

Genetic programming has been utilised to forecast one day ahead share prices (Abraham & Nath, 2000) and has proved useful in economic forecasting and market behaviour (Kaastra & Boyd, 1996). There have been numerous other uses for genetic programming which include predicting exchange rates (Álvarez-Díaz & Álvarez, 2005) as well as others further detailed below.

Fernández and Gómez (2007) used neural networks to compute the mean-variance curve (the mean-variance curve is further explained in the section entitled 'Markowitz Portfolio Theory'). The generalised Markowitz model was utilised in their approach as to include bounding and cardinality constraints - this tactic was employed in order to take into account that smaller amounts of money cannot purchase certain shares. This is not considered as a necessary requirement in this study as the barrier can easily be overcome in 'real world' scenarios through the use of contracts for difference (Norman, 2010). Secondly, the amount of capital to be invested in each asset was limited in both its upper and lower bounds. It is acknowledged that the problem is a mixed quadratic and integer programming problem and as such no computational efficient algorithm exists. This is one of the downfalls of the genetic programming methodology, more of which will be described later.

Chen, Yang, Dong and Abraham (2005) aimed to investigate how the seemingly chaotic behaviour of share prices could be represented using a subcategory of genetic algorithm, specifically a flexible neural tree model. The authors aimed to predict the NASDAQ-100 and S&P CNX NIFTY stock indexes. Their results found that the model considered could represent the stock indices extremely accurately.

Freitas, Souza and de Almeida (2009) used an optimal portfolio prediction based neural model in an effort to discover whether the model could take advantage of short-term investment opportunities. Neural network predictors were utilised to predict stocks' returns and thereafter derived a risk measure based on the prediction errors that were built on the same foundation as the mean-variance model. The results from this model were subsequently compared to the returns accomplished through the stock market index as well as on the corresponding mean-variance model.

Freitas, Souza and de Almeida (2009) highlighted the difference between the prediction-based portfolio optimisation model and the mean-variance model by examining three factors. Firstly, in the prediction-based portfolio optimisation model the expected return of each stock is its predicted return, instead of the mean of its time series of returns, as is the case in the mean-variance model. Secondly, the prediction-based portfolio optimisation model calculates the individual risk of each stock and the interactive risk between each pair of stocks from the variance and covariance of the time series of the errors of prediction; the mean-variance model however calculates the individual risk of each stock and the interactive risk between each pair of stocks from the variances and covariance of the time series of returns. Finally, although both models are based on the normal framework, in the prediction-based portfolio optimisation model the normal variable of interest is the error of prediction of the return of the stocks, while, in the mean-variance model, the normal variable of interest is the return of the stocks.

The analysis was performed initially on 82 stocks that participated on the IBOVESPA Brazilian stock index over the period between October 1999 and September 2007. From the population, a subset of 52 stocks was selected that had a long enough time series for training the neural network and calculating the necessary parameters of the portfolio optimisation model. For each stock of the subset the weekly closing prices (sampled on Wednesdays) between 27 October 1999 and 19 September 2007 were used to train the network. On all days where there was missing data the last available closing price was used (Freitas, Souza & de Almeida, 2009). Risk in the portfolio was measured by the RMSE (root mean squared error), the mean absolute percentage error and the hit rates on the returns. The portfolio was then evaluated based on risk and return measures as well as on accumulated return, portfolio change measure and the turnover index (Freitas, Souza & de Almeida, 2009).

To train and test the neural networks involved in the experiments Freitas, Souza and de Almeida (2009) used a sliding window of 168 weeks of the weekly 413 returns available. In total 12 792 training sessions were calculated (413 weekly returns, minus the 168 sliding window of train and test cycles multiplied by 52 stocks). The sliding window contains the training set (163 input-output pairs) and the testing set (1 input-output pair).

The results demonstrated that the model achieved returns of 291% above the mean-variance model for similar levels of risk. Furthermore, the predictive portfolio showed an improved market index tracking capability, achieving returns of 77% above the IBOVESPA market index. These increases appear optically large; this is due to the nature of the initial results, being numerically extremely small. As such relatively small numerical improvements on a relatively small scale are magnified through the demonstration of percentage improvements, these are commonly referred to as base effects. The results also displayed that it was possible to obtain normal prediction errors even when the time series consists of non-normal data (Freitas, Souza & de Almeida, 2009).

Miles and Smith (2010) employed genetic programming to develop trading rules, which were applied to test the Efficient Market Hypothesis (which refers to the belief that the market fairly reflects that value of securities). There is a large amount of research focused on the Efficient Market Hypothesis and studying its various forms, however that is not the focus of this dissertation. The majority of previous research that aimed at testing the Efficient Market Hypothesis was limited to utilising trading rules that returned simple buy-sell signals. Miles and Smith (2010) used a broader approach wherein the framework was consistent with the standard portfolio model. This technique developed trading rules that are defined as the proportion of an investor's total wealth which should be invested into a risky asset. The methodology utilised the average utility of terminal wealth as the fitness function. The technique was developed using data, retrieved from Datastream, on daily stock prices (unadjusted for dividends) from 1985 to 2005 from 24 diverse companies that traded on the (NYSE). To ensure diversity among the companies selected, two shares were selected from each of the 12 industries implied by Fama and French's industry classification scheme (as cited in Miles and Smith, 2010). These categories were namely: consumer durables, consumer non-durables, manufacturing, energy, chemicals, business equipment, telecommunications, utilities, shops, healthcare, finance, and other. Further selection criteria included that companies must have been active in the market for the time period beginning at the start of the last quarter of 1979 and continued trading on the NYSE through to the end of 2005. The algorithm that was developed aimed to find trading rules for the 24 individual stocks. These rules were then applied to out-of-sample data to test the adaptive efficiency of the markets studied.

When markets are efficient, it is said that investors cannot make profits by exploiting publicly available information (Miles & Smith, 2010). Daniel and Titman (1999) introduced a weaker concept of market efficiency called adaptive efficiency: a market is characterised by adaptive efficiency if profit opportunities disappear when they become obvious. The objective of Miles and Smiths' (2010) paper was to test adaptive efficiency of stock markets by conducting a broadly representative study (using data from 24 stocks across a wide spectrum of industries, as mentioned above) of the efficiency of trading rules evolved using a genetic programming methodology. If the rules evolved by genetic programming using in-sample data have a relatively low fitness when applied to new (out-of-sample) data, it is interpreted as evidence of adaptive efficiency.

The study performed the analyses using 21 out-of-sample periods for each of 24 stocks (for a total of 504 out-of-sample periods). A genetic programming methodology is employed to generate portfolio rules to determine the fraction of wealth to be allocated to a risky asset (one of the 24 stocks) and test adaptive efficiency of the markets on the 21 out-of-sample periods. Any of the remaining wealth would be invested into a riskless U.S. treasury asset. A rolling 5-year in-sample period is used as an input to evolve and select trading rules that are then tested on the following (sixth) year. The first half of each five year in sample period is allotted to training and the second half to selection. Furthermore it was assumed that an individual's trading horizon is 60 trading days (Miles and Smith, 2010).

In addition, transaction costs were incorporated into the analysis. When the investor purchases shares, both the cost of the shares and the transaction costs were subtracted from the cash account. This is completed in order to decrease the incidence of retaining rules that over-trade. This is carried out because rules that over-trade are more likely to be overfitting the data. Unrealistically high transaction costs were used in the training and selection periods and thereafter, realistic transaction costs were used in the testing period. For the training and selection periods, a one-way transaction cost of 0.5% with a two-way flat rate of $5 per share of stock was used in. This was deliberately unrealistically high for the reasons mentioned above. For the testing period, the transaction cost structure used by Allen and Karjalainen (1993) for simulating trading in the S&P 500 index was employed. This entails incorporating a one-way transaction cost of 0.25% of the value of the transaction. Allen and Karjalainen (1993) argued that a one-way transaction cost of 0.25% incorporates all costs at realistic levels, including the cost of the market impact (Miles & Smith, 2010).

Each genetic programming experiment conducted as part of the study involved 10 trials, and each trial consisted of 50 generations. In every generation, a population size of 50 000 random potential trading rules was used. Furthermore, the depth of each candidate solution decision tree was limited to 25 levels (Miles & Smith, 2010). The process of running the 10 genetic programming trials and selecting, at most, one rule to be tested in the out-of-sample period consisted of the following steps: firstly, 50 000 random rules were generated and their fitness values evaluated in the training and selection periods. Thereafter, all of the rules that satisfy the criteria were identified and subsequently saved. If more than 50 rules satisfied the criteria, only the rules with the highest fitness functions in the selection periods were saved. Secondly, a probability of being selected was then attached to each rule. The probability corresponded to each rule's fitness function during the training and selection periods, aiming to assign rules with better fitness values a higher probability of reproducing. Thirdly, rules were then selected based on their attached probabilities and then the crossover operator (with probability 95%) or the mutation operator (with probability 5%) was performed. In this way, 50 000 rules were generated for the following generation. The best 50 performing rules were thereafter saved again. If this generation was not the final generation, the second step was reused. If the current generation is generation 50 (the final generation) then the next trial should begin, going back to the first step (unless this is the tenth trial). If this is the tenth trial then the rules created during the 10 trials should be saved and those that do not satisfy the criteria should be discarded. Of the remaining eligible rules the one with the highest fitness function should be selected and its respective performance studied in the testing period.

A population size of 50 000 was at the time substantially larger than other studies (usually using approximately 500 as a population size) (Miles & Smith, 2010). An assumption is made in the analysis, as is common, that the activities of the simulated trader do not have a major impact on the stock price. Their findings showed that in general, the trading rules that the methodology generated do not outperform the simple buy-and-hold strategy. This was therefore understood as proof that of the 24 shares examined, all were adaptively efficient between 1985 and 2005.

The majority of the trading systems designed in academic literature (as the above) produce remarkable results but are typically applied to trading a single asset rather than a group or portfolio of assets. In this respect, risk is relatively ignored because during periods of high volatility, higher returns are expected. Despite this constraint, the section above highlights the robust results achieved by the genetic programming approach over a variety of applications, demonstrating strength of the methodology supporting genetic programming as well as its relatively flexible nature. This chapter now continues by further exploring the second of the artificially intelligent systems made use of in this dissertation, namely particle swarm optimisation.

2.3 Particle Swarm Optimisation

The methodology behind particle swarm optimisation (PSO) was discovered by Kennedy and Eberhart (1995), a social psychologist and electrical engineer respectively. Their initial intent was to simulate the graceful but unpredictable flight of a bird flock (Eberhart & Shi, 2001). This, however, developed into a computational method that has demonstrated the ability to solve difficult problems efficiently and reliably (Poli, 2008).

The roots of PSO are split into two main component methodologies. The first is that of artificial life as it relates to bird flocking, fish schooling and swarming theory in particular. The second relates to the ties that PSO shares with evolutionary computation and thus genetic algorithms, as previously described (Eberhart & Kennedy, 1995).

When swarms solve problems in nature, their abilities are usually attributed to 'swarm intelligence'. The most prolific examples of this are flocks of birds, shoals of fish as well as social insects such as bees and ants (Poli, 2008). The aim of PSO is to identify and thereby take advantage of the 'swarm intelligence' found within nature and apply it to separate scientific and industrial purposes (Poli, 2008).

In keeping with previous literature, the data points used will be referred to as particles rather than as points. While it could be argued that the population members are massless and without volume (and in this way could be called points), it was decided that due to the fact that the members have velocities and accelerations associated with them, to be further elucidated upon below, it is more accurate to refer to them as particles (Eberhart & Kennedy 1995; Kennedy & Eberhart, 1995).

In PSOs, a number of simple entities (the particles) are placed into the parameter space of a problem or a function at random (McCarthy & McCluskey, 2009). Each particle represents a possible solution and as such, evaluates the fitness level at its current location (McCarthy & McCluskey, 2009). Next, the particles each determine their movement through the parameter space by combining some aspect of their own historical fitness value along with one or more members of the swarm; subsequently the particle will move through the parameter space with a velocity determined by the locations and fitness values of the other members of the swarm (along with one or more random variable/s as will be discussed later) (Poli, Kennedy & Blackwell, 2007; Poli, 2008). The next iteration will take place after all the particles in the swarm have moved (Poli, Kennedy & Blackwell, 2007). The members of a swarm that a particle can interact with are called its' neighbourhood while the social neighbourhood of all the particles (or more simply, all the particles in the swarm combined) are referred to as the PSO's social network (Poli, 2008).

It is important to highlight that one particle by itself has almost no power. Kennedy and Eberhart (1995) point out that it does not seem too large a leap of logic to suppose that some of the same rules that underlie animal social behaviour, also underlie parts of human behaviour. It is described that the individual members of a school of fish (or the like) can profit from the discoveries of their individual experience coupled with the previous experience of all other members of the school during the search for food, which benefits all in the population (Wilson, 1975). In the same vein, progress in the strength of PSO is increased by the number of particles involved as well as the interaction (or social informational sharing) between particles (Poli, Kennedy & Blackwell, 2007).

When comparing the behaviour of animals such as fish and birds to that of humans (or when trying to simulate human behaviour through PSO), Kennedy and Eberhart (1995) highlighted one important distinction, namely that of abstractness. In the case of animals, the physical movement of an individual and of the swarm can be adjusted so as to avoid predators or to seek out food as well as to optimise the environment. This is the foundation of PSO as previously mentioned. However, when simulating human behaviour, not only should physical movement be taken into account but also that of cognition or experiential variables as well. The major distinction when considering cognition rather than merely physical movement is that cognitive does not change or move in patterns such as a flock of birds would change direction, but more importantly the concept of collision. This refers to the ability of two people to hold identical attitudes and beliefs without mentally 'colliding' into each other – this is different to animals where the same physical space cannot be held by two individuals at any one point in time. When modelling human behaviour using PSO this is arguably the most important consideration – two particles are now allowed to occupy the same space with regards to cognition, but cannot occupy the same space with regards to physical space (Kennedy & Eberhart, 1995).

Millonas (1994) developed the five basic principles of swarm intelligence. These are namely: First, the proximity principle which relates to that the notion that the population should be able to carry out simple space and time computations. Second is the principle of quality: the population should be able to respond to quality factors in the environment. Third is the principle of diverse response, meaning that the population should not commit its activities along excessively narrow channels.

Fourth is the principle of stability which holds that the population should not change its mode of behaviour every time the environment changes. The final principle is that of adaptability: the population must be able to adapt its behaviour when it is worth the computational price. In addition it was pointed out the principles four and five are the opposite sides 'of the same coin' (Eberhart & Kennedy, 1995).

Initially there were two main types of PSO, global and local. Global PSO follows an algorithm containing six steps:

1. Initialise a population (array) of particles with random starting positions and velocities on *d* dimensions in the problem space
2. For each particle, evaluate the desired optimisation fitness function in *d* variables
3. Compare the particle's fitness evaluation with the particles *pbest* (*pbest* is a generic term that refers to the best fitness value that the particle has achieved thus far). If the particles value is better than the stored *pbest* then set *pbest* value equal to the current value and the location equal to the current location, else do not change *pbest*
4. Compare the particles' fitness evaluation with the population's overall previous best fitness valuation (referred to as *gbest*). If the current value is better than *gbest* then set *gbest* equal to the current value and the location equal to the particles array index, else do not change *gbest*
5. Change the velocity, acceleration and position of each particle according to equations 1 and 2 below respectively:
   a. $V_{id} = V_{id[t-1]} + c_1*rand()*(p_{id}-x_{id}) + c_2*rand()*(p_{gd}-x_{id})$ *(1)*
   b. $x_{id} = x_{id[t-1]} + V_{id}$ *(2)*
6. Repeat step 2 until a criterion is met (the termination criteria can be time, a sufficiently good fitness function or a maximum number of iterations).

This process was set out initially in Eberhart and Kennedy (1995). In equation 1, c_1 and c_2 refer to the weighting of the stochastic acceleration terms that pull each particle toward the *pbest* and *gbest* positions respectively. Low values allow particles to roam far from the target regions before tugging back while high values result in an abrupt movement toward or past target regions (Eberhart & Shi, 2001). In essence, the acceleration coefficients, c_1 and c_2, when combined with random numbers control the random search effect of the cognitive (local) and social (global) components of velocity. Thus the relative nature implies the exploratory nature of the particles. It has been found that the general best values for c_1 and c_2 are 2.0 for almost all applications (Eberhart & Shi, 2001; Poli, 2008).

V_{id} refers to the particle's velocity and $V_{id[t-1]}$ denotes the particle's previous velocity. The particle's cumulated velocity on each dimension is subject to a constraint of *Vmax* which refers to the sum of the accelerations. If the sum of the accelerations is larger than *Vmax*, then *Vmax* is limited to the value input by the user. This is an important parameter as if *Vmax* is too high, particles might fly past good solutions however if *Vmax* is too small particles may become stuck in local optima (Eberhart & Shi, 2001). X_{id} refers to the particle's current position, $X_{id[t-1]}$ refers to the particle's previous position while p_{id} refers to the particle's previous best position and p_{gd} refers to the previous best position of the best particle in the population. *Rand()* refers to a random number generator, generating number of between 0 and a maximum number input by the user (Poli, 2008).

The local version of PSO works in the same way as the global version, however, with the exception that now in addition to *pbest* (which refers to the best fitness value that the particle has achieved thus far) each particle also keeps track of *lbest* in place of *gbest* (Eberhart & Shi, 2001); *lbest* refers to the best value of one (or more) of the neighbouring particles' fitness values (Eberhart & Kennedy, 1995). In the local model, the number of particles that will be included in a certain particle's neighbourhood is specified by the user (Eberhart & Kennedy, 1995). It has been shown that a neighbourhood of 15 percent of the population size provides near optimal performance in the majority of PSO applications (Eberhart & Shi, 2001).

Conceptually, a particle's velocity can also be viewed as simple nostalgia – the individual will tend to return to the place that most satisfied it in the past; the *pbest* value can be seen to represent autobiographical memory, as each individual remembers its own experience while *gbest* can be seen to represent publicised knowledge, or a group norm or standard that individuals seek to attain (Kennedy & Eberhart, 1995).

Since the inception of PSO there has been a large amount of changes, revisions and updates made to the algorithms in order to optimise PSO for specific application areas. Suganthan (1999) suggested that the *Ibest* topology seemed superior for exploring the search space whilst *gbest* converged faster – thus it would be best to begin the search with an *Ibest* lattice and slowly increase the size of the neighbourhood until the population was fully connected by the end of a run. Liang and Suganthan (2005) created a subpopulation of size *n* and occasionally randomised all of the connections. Clerc (2006) developed a parameter free PSO algorithm called TRIBES, established upon concepts inherent in the development of human tribes, where details of the topology evolve over time in response to performance feedback (in this method good tribes can benefit from the removal of their weakest member whilst bad tribes could benefit from the addition of the same new member). There have been numerous other proposals for PSO (for an in depth analysis see Poli, Kennedy and Blackwell (2007) however, this is not the primary focus of this dissertation and therefore will not be discussed in detail.

McCarthy and McCluskey (2009) used a PSO algorithm for the cost optimum design of reinforced concrete beams. They made use of this approach while incorporating multiple constraints. This made the optimisation a multi-variable problem as it furthermore took into account factors such as bending and shear experienced in the beams as well as the reinforcement conditions that are required to resist these forces. It was found that PSO performed relatively well when used for this purpose.

Multiple factors regarding PSO were highlighted by McCarthy and McCluskey (2009), firstly, that the size of the population has a high influence on the performance of the algorithm. If a high number of particles are employed it allows the algorithm to search a greater area in each iteration leading to a greater chance of finding a global optimum. This advantage is however offset because as the swarm size is increased, so too is the required time to find an appropriate solution. Secondly, if $c_1 > c_2$ (as defined in equation 1) particles tend to wander excessively while if the converse is true, the particles will tend to converge rapidly, increasing the possibility of becoming trapped in a local minima (McCarthy & McCluskey, 2009).

As mentioned previously, PSO was established by Kennedy and Eberhart (1995) with the initial intent to simulate the flight patterns of a bird flock. This advanced into a methodology that applies the concepts of swarm theory to efficiently evaluate a computational problem. The methodology behind PSO was examined above however it is interesting to note that PSO has shown promise in a wide array of applications.

Poli (2008) used a mechanical process to divide the literature applying PSO into various categories. Through a search phrase and key terms analysis Poli (2008) generated a graphical representation of key terms that are mentioned in previous literature. This was then used to determine categories wherein past literature had applied the PSO methodology. An example of the graphical representation is shown in Figure 4. Each node in the figure below is representative of a key term, for example “Network”, “Neural” or “Particle”. The lines joining each node demonstrate that the terms have been used in conjunction with one another in the same paper. The closer a node is to the centre of the graph represents the number of times a term has been used over various studies, with the more central nodes illustrating a term has been used more often that the nodes towards the outskirts of the image.

Figure 4: Representation of Key Terms Relevant to PSO Applications

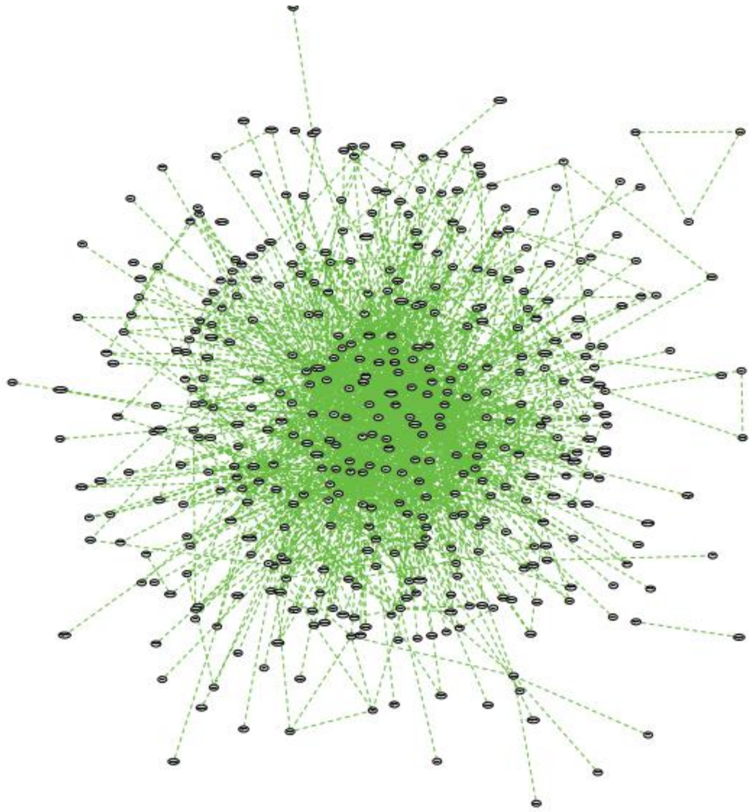


Figure 4*: Example of a graphical representation of key terms relevant to PSO applications and their relationships as demonstrated in:* "Analysis of the Publications on the Applications of Particle Swarm Optimisation", by Poli, 2008, *Journal of Artificial Evolution and Applications, 3, 1-10*

Poli (2008) used the analysis above to determine categories where the application of PSO had been documented in the past. The categories highlighted wherein PSO showed promise included the application to biomedical problems, communication networks, security and military efforts as well as, as in this dissertation, to financial and investment problems (Poli, 2008). The study however did not make reference to which previous studies were included in the analysis above.

It was demonstrated by Poli (2008), that the extent of literature that focuses on the application of PSO to computational problem sets, with documented results, is extremely limited. The vast majority of the literature, thus far, has focused on the theoretical variations, improvements and assumptions of PSO. Poli (2008) indicated that their methodology found 1100 publications using PSO. Of these publications, 350 were proposals for improvements and specialisations of PSO. Around 700 papers could be classified as applications of PSO however the majority of these still focus on the customisation or extension of the methodology to better suit an application of interest.

Furthermore, of the 700 papers, only 55 were journal articles. The remaining findings were those of conference proceedings and other miscellaneos entries. As such it can be observed that the domain of PSO is still in its infancy but the attention focused on the approach is increasing exponentially each year; this notion is further elaborated upon below.

The interest in PSO has surged since the paper by Kennedy and Eberhart (1995) first brought PSO to light. Through the completion of an extensive study, partly outlined above, Poli (2008) created the figure below (Figure 5) demonstrating the increasing attention being focused on PSO. Figure 5 reiterates that the domain of PSO is still in its early stages. Additionally it highlights the notion that the attention being paid to PSO is increasing exponentially each year. This dissertation aims to further the domain by a comparison to the traditional simulation approach as well as to the genetic programming approach with regards to aiming to minimise the variance of a portfolio of shares.

Figure 5: Number of Papers Published per year on PSO

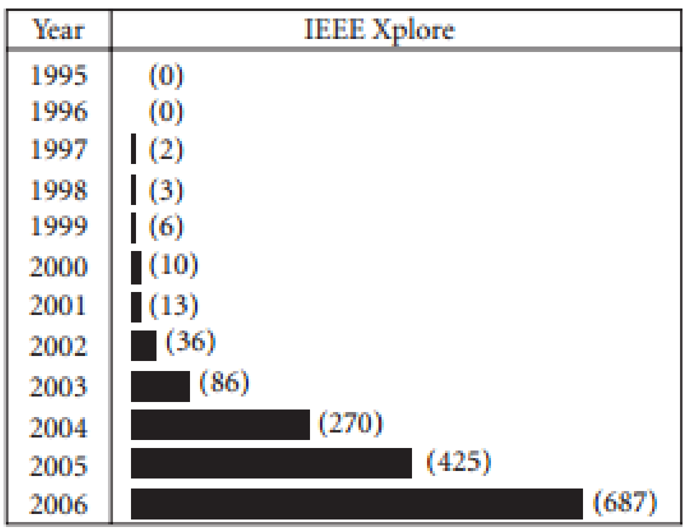


Figure 5: *Display showing the number of papers published per year on PSO (Poli, 2008) since its inception in 1995.*

Papers that directly compare PSO to other optimisation techniques are extremely limited (Zhu, Wang, Wang & Chen, 2011) and this dissertation aims to reduce that gap in the literature, especially with regards to solving financial problems. In the few studies that have aimed to compare PSO with other optimisation techniques, PSO has been shown to achieve results comparable or superior to state of the art solvers (Zhu, Wang, Wang & Chen, 2011). Furthermore, when examining real world problems, where users may have limited computation time and limited precision in estimating instance parameters, many analytical methods have been shown to not be particularly suitable. This is due to the idea that as the problems become larger and more complex, the computing power and time required to find an adequate solution becomes exponentially larger. PSO however has demonstrated the ability to find high quality solutions in a reasonable amount of time, with reasonable computing power (Zhu, Wang, Wang & Chen, 2011). Eberhart and Kennedy (1995) also found that PSO may perform better than genetic programming as the interaction between group members enhances rather than detracts from progress towards the solution. This outperformance was thought to be due to the methodology where the interaction of the particles in PSO will increase the power of the performance while in genetic programming, if two tree structures (as displayed earlier in Figure 3) are different and both have high fitness evaluations, a recombination of them may not yield a better overall result.

Kendall and Su (2005) applied PSO to the problem of constructing an optimal risky portfolio. The optimal risky portfolio, as described earlier, denotes the particular selection of securities and their respective weights that results in the maximum reward-to-variability ratio. The PSO algorithm was developed and tested across multiple restricted and unrestricted portfolios. A restricted portfolio refers to a portfolio where the weightings on each share is constrained between 0 and 1, as such no short selling is allowed. An unrestricted portfolio denotes a portfolio where the short selling of shares is permitted - the weightings on each share is constrained between -1 and 1. At all times, in both restricted and unrestricted portfolios, the total sum of the weights across all shares is ensured to be equal to 1.

In the study above, PSO was tested and compared against the traditional Excel Solver (Kendall & Su, 2005). Initially the data set included 7 share indexes, each from a different country. In further tests, the data set was expanded to include a random selection of shares from the Hang Seng index over the period 5 October 1998 to 2 October 2003. Tests were then run when the number of available shares varied between 5, 7, 12 and 30 shares in turn. Kendall and Su (2005) found that PSO delivered favourable results compared to the traditional Excel Solver method. The performance of the PSO approach however severely decreased in efficiency when the number of available shares increased above 15. This was noted to be chiefly due to larger portfolios requiring exponentially longer run test times. This was noted as a limitation of the study and an aim for future research (Kendall & Su, 2005).

PSO and genetic programming share two main similarities. In both methodologies the system is initialised with a population of random solutions and works methodically through steps, governed by parameters in order to improve on the initial random solution (Eberhart & Kennedy, 1995; Eberhart & Shi, 2001). In addition both systems have been shown to display adequate strength and remain effective in noisy environments (Poli, 2007). It was even shown that in some cases noise improved the strength of PSO as to allow it not to become trapped in local optima (Poli, 2007).

There are however fundamental differences between PSO and genetic programming. In PSO, once the system is initialised with a population of random solutions, each potential solution is assigned a randomised velocity and the potential solutions are then flown through the problem space (as described above) (Eberhart & Shi, 2001). Thereafter, each particle keeps track of its' coordinates in the problem space which are associated with the best solution thus far (Eberhart & Shi, 2001). This is notably different from the methodology of genetic programming as outlined in the earlier section of the dissertation.

In summary, PSO shines for its simplicity and for the ease in which it can be adapted to best suit different systems, multiple different application domains as well as the ease with which it can by hybridised with other techniques (Poli, 2008). PSO combines the evolutionary methodology behind genetic programming with the paradigm of artificial life based on swarm intelligence. The sub-chapter above described the methodology and foundations that the PSO approach is built on. Despite research comparing PSO to other optimisation techniques still being fairly limited, the studies that have done so thus far were expounded upon above in addition to the studies that focused on the variety of other areas that particle swarm intelligence has been applied too. Their respective methodologies were explored, their results illustrated and finally a comparison of the similarities and differences between PSO and genetic programming was completed.

The sections thus far in the Literature Review section have highlighted the importance of finding the optimal number of shares to include into a diversified portfolio, as well as the costs of failing to do so. This establishes the significance of fully grasping the relationship between the risk of a portfolio and the number of shares the portfolio consists of. As such, an exploration of the Markowitz Portfolio Theory is carried out in the sub-section below in order to further elaborate on this relationship.

2.4 Markowitz Portfolio Theory

Markowitz stated that the process of selecting a portfolio can be divided into two stages (Markowitz, 1952b). The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage begins with relevant beliefs about future performances of securities and ends with the choice of the portfolio. Markowitz aimed to explore the second stage of portfolio selection (Markowitz, 1952b).

Throughout his analysis, Markowitz assumed that an investor views increasing returns positively and thus should aim to maximise returns while viewing increasing variance of a portfolio negatively and thus should aim to minimise a portfolio's variance (Markowitz, 1952b). Furthermore, it is noted that should an investor know the future returns of a selection of shares with certainty he would in no case utilise diversification but would instead choose to place all funds available in the security with the greatest return.

It is further noted that although adding securities to a portfolio should generally decrease the risk of the portfolio, the 'Law of Large Numbers' does not fully apply to a portfolio of securities (Pearson, 1925). The 'Law of Large Numbers' is a statistical term used in probability theory that refers to the finding that the average of the results obtained from a large number of trials should approximate the expected value and will tend closer to the expected value as an increasing number of trials are performed. The 'Law of Large Numbers' assumes that each trial run is independent of the results achieved previously. This is not the case with a portfolio of securities as the returns from securities may be inter-correlated and if two shares are perfectly correlated to one another, adding both to a portfolio would not reduce the amount of variance inherent in the portfolio. As such the optimal selection of shares would entail selecting a portfolio of shares, each with low joint risk. Furthermore, diversification cannot eliminate all risk as market risk will continue to remain (only idiosyncratic risk can be diversified away, as previously explained).

Independently, Roy (1952) concentrated on a specific portfolio that an investor should select based on minimising the upper bound of the chance of a dread (negative black swan) event. This was then applied to a portfolio of *n* assets for either speculative gain or to maximise the income yielded (Roy, 1952). Although Roy (1952) has not received the same recognition as the 1952 article by Markowitz, Markowitz (1999) noted that Roy (1952) can claim an equal share of the honour of being called the father of modern portfolio theory.

Markowitz (1952b) laid the foundations of modern portfolio theory by outlining the general ‘rule’ that investors do not get rewarded for non-systematic risk and thus should diversify their holdings accordingly. The ‘Law of Large Numbers’ (Bauer, 1996) was applied to share price expectations – that an investor’s actual return should trend towards the expected return of the market as more securities are added to his portfolio (as non-systematic risks, such as black swan events, affecting a single company would only have a limited effect on a well-diversified portfolio; compared to if the investor held only the one share that the event negatively affected). This was applied under the constructs mentioned previously regarding the ‘Law of Large Numbers’.

Markowitz (1952b) discovered numerous notable findings. Firstly, diversification can reduce (and potentially minimise) the risk of a portfolio, nevertheless it cannot eliminate all possible risk. This is due to the fact that only diversifiable risk (or non-systematic risk) can be diversified away.

Secondly, a portfolio with a maximum expected return does not necessarily imply that it has minimal variance. Thus, the return alone should not be the deciding factor when selecting a portfolio but rather the ratio of return to risk. This leads to the derivation of Figure 6, below:

Figure 6: Markowitz Efficient Frontier

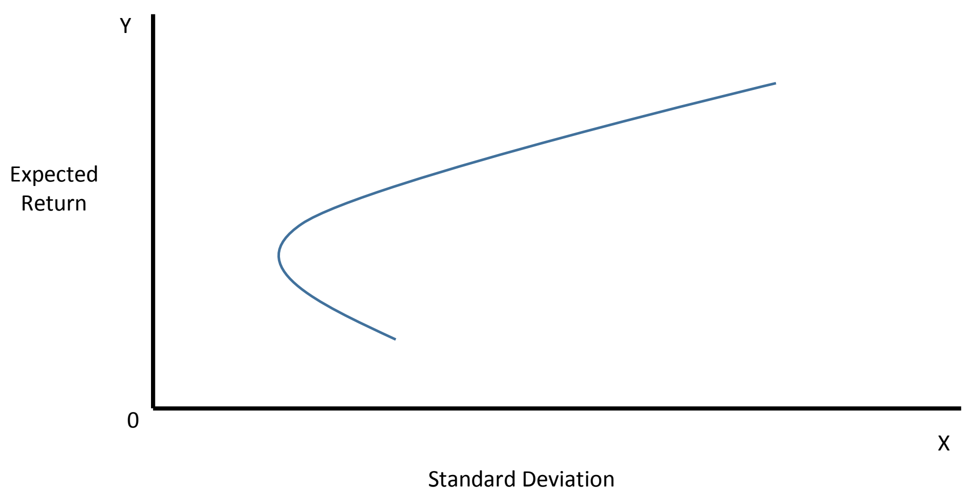


Figure 6: *The graph showing the relationship between the expected return and the standard deviation of a portfolio – as put forward in Markowitz (1952b).*

Figure 6 displays that for each level of return there is an associated minimal level of standard deviation. Furthermore, the dual problem states that for each level of risk (standard deviation) there is not a point that can offer a greater return. The line displayed is known as the efficient frontier. At any point along the efficient frontier an investor can be viewed as owning an optimally diversified portfolio as he cannot realise any additional returns without assuming additional standard deviation (risk). Rephrased, the investor now achieves minimum variance on their portfolio on a given expected return. Any points above the efficient frontier are preferable to any points vertically below them as the investor would be increasing return while assuming the same level of risk. Points below the efficient frontier are sub-efficient as an investor is assuming a lower return than is possible when assuming the given level of risk. This also points to the notion that an investor can increase his expected return by increasing a portfolio's variance and thus assuming more risk.

The final conclusion of Markowitz (1952b) is that the expected return – variance figure (as pictured above) does not imply that any and all diversification will pay off – but rather that one should diversify for the right reasons. This is to take into account share price covariance, industries and other company factors. It is highlighted, however, that if two shares of equal variance are purchased, the resulting portfolio will have less variance than either of the two shares held separately due to a portion of the non-systematic risk being diversified away, assuming non-perfect correlation between the two shares.

Markowitz (1991) mentions how an optimising investor would behave with regards to portfolio diversification. The statement that if an investor knew all of the future share returns with certainty he would invest in one share only (the one with the highest return) and in no case would he prefer a diversified portfolio, is reiterated in this article. Thus, the existence of risk and uncertainty is imperative to the study of diversification. It is acknowledged, as previously explained, that risk cannot be thought of as equivalent to uncertainty.

This dissertation focuses not on optimising the risk and return trade off as displayed by the Markowitz efficient frontier, exhibited in Figure 2, but rather on optimising the number of shares to the portfolio variance exchange (Markowitz, 1952b). In order to observe this a number of formulas must be utilised. Firstly, once the data series of the share prices to be studied are retrieved the share price returns must be calculated respectively. This is completed as follows (assuming monthly closing share prices): The closing price of the current month, minus the closing price of the previous month plus the dividends accrued for the month, all divided by the closing price of the previous month.

Assuming the accrued dividends are taken through share price adjustments, the formula for a share prices' monthly return can be seen below:

$$R_t = \frac{P_t - P_{t-1}}{P_t} \tag{3}$$

In the above equation, R represents the return of a share at time t, P_t represents the price of the share at time t while P_{t-1} represents the price of the share the previous period. The result is the return which an investor would have made or lost if he invested funds into a security at the beginning of a month, collected the dividends accrued to him throughout the month and sold the security at the end of the month. A loss is represented by a negative return. The analysis in this dissertation, as in Markowitz (1959), assumes that a dollar of capital gain one might receive is exactly equivalent a dollar of dividends received; as such no tax implications are taken into account (Markowitz, 1959).

The portfolio return can then be calculated by the summation of weighting the respective shares and multiplying their weight by their return respectively. This is completed using the formula below:

$$R_p = w_1 * r_1 + \ w_2 * r_2 + \ \dots \ w_x * r_x \tag{4}$$

Or alternatively

$$E(r_p) = \sum_{i=0}^{n} W_i * E(r_i) \tag{5}$$

In equation 4, R_p represents the portfolio return while $w_1 * r_1$ represents the weight on the first share multiplied by its respective return, $w_2 * r_2$ represents the weight on the second share multiplied by its respective return and so on. In order to calculate the return on the overall portfolio of shares, the weighted return formula was used.

Similarly, if one forecasts possible future share prices rather than looking retrospectively at past share prices, equation 5 should be utilised. In equation 5 above, $E(r_p)$ refers to the expected return on the portfolio of shares. This is once again equal to the summation of the weight on each share multiplied by the expected return on each share. In the methodology employed in this study, the sum of the weights on each share is always made to be equal to one; furthermore the weights on all shares must be non-negative. This restricts the analysis to an environment where an investor is not allowed to short sell a share, but may assign either a 0 weighting or a weighting of larger than 0 but smaller or equal to 1 to any share.

It can be expected therefore that the return of a portfolio should lie between the potential return that could be realised through the highest returning share and the potential return that could be realised by the lowest returning share – any combination of shares cannot yield returns greater than the single share maximum or less than the single share minimum (Markowitz, 1959).

It could then be assumed that similar constraints could be expected when analysing the variance of a portfolio, however this is proven not to be the case. Computing the variance of a portfolio requires not only the weighted variances of the respective shares to be taken into account, but also how the shares may covary. When analysing a portfolio of shares the best measure of variability of the return series is the standard deviation of the portfolio (Markowitz, 1959). The standard deviation on each share can be calculated in the classical statistical sense, as seen below:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \qquad (6)$$

In the equation above σ represents a single share's standard deviation. x_i refers to the current observation in the series while $\bar{x}$ refers to the mean or average of the series. The sum of these squared differences, divided by *n-1* (*n* referring to the total number of observations in the series) equates to the standard deviation of the share. The square of this measure equates to the share's variance.

The standard deviation for each share as well as the variance (variance being equal to the standard deviation squared) for each share in a portfolio can be worked out in the classical statistical manner; however when calculating the standard deviation or variance of a portfolio the process must take into account three factors: firstly the standard deviation of each share, secondly the correlation between each pair of shares and finally, the weights to be assigned to each share. Thus one should not simply utilise the average standard deviation of the individual shares included in a portfolio to be the standard deviation of the portfolio, as it is necessary to account for how shares may covary, or move together (Markowitz, 1959).

One share's covariance with another can be calculated as follows:

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \tag{7}$$

This formula demonstrated that the covariance between two securities is the sum of the product of the deviation of the first security from its mean, multiplied by the deviation of the second security from its mean, divided by the number of observations in the series subtracted by 1. This measure allows insight into how the two securities move together in a portfolio. If the deviations of the securities move in the same direction (both positive or both negative), one can conclude that they move in generally the same direction and as such there is limited diversification benefit when combining the first security with the second. If they move in opposite directions however, one should expect the volatility of the resulting portfolio to be lower than either of the securities' volatility on their own. Additionally, if the securities are completely unrelated, in theory the covariance between the two tends to be zero (Elton, Gruber, Brown & Goetzmann, 2009).

A related concept to covariance is correlation. Correlation is calculated as the formula below:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}} \tag{8}$$

In this formula r is the correlation coefficient, x is the first series observation and y is the second series observation. The correlation coefficient refers to the strength of the relationship between two shares and is a measure that will lie between -1 and +1. If the value is -1 it can be said that the securities move in perfectly opposite directions, thus if security *A* moves in a positive direction by one unit, security *B* will move in the opposite direction, also by one unit. If the value is +1 the securites move in the same direction perfectly. In this case, if security *A* moves in a positive direction by one unti, security *B* will move in the same direction by the same magnitude. A value of 0.5 means that if the first security moves in a positive direction of one unit, the second security will move in the same direction by a unit of 0.5 and so on. The hypothesis from this is that the optimal benefit of diversification will be obtained on two securities if they move in perfectly opposite directions, thus all moves will offset oneanother.

To move between between the correlation and covariance between two shares, the following formula can be used:

$$r_{(x,y)} = \frac{COV(x,y)}{S_x S_y} \qquad (9)$$

The $COV(x, y)$ refers to the covariance between share x and share y while S_x refers to the standard deviation of x and the same is true for S_y respectively.

The general formula to calculate the variance of a portfolio can be seen below:

$$\hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^{M} X_i^2 \sigma_{\varepsilon i}^2 + \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} X_i X_j \gamma_{\varepsilon ij} \qquad (10)$$

In equation 10 (above), $\hat{V}$ represents the variance of a portfolio (which is also equal to $\hat{\sigma}_p^2$, the standard deviation of the portfolio squared). *M* represents the portfolio, being the collection of *M* shares, each with their own weight. The first summation represents the contribution of risk that each stock adds to the portfolio (sum of the square of the participation (weight of each share) X_i, multiplied by the variance of the share price). The second group of summations represents the contribution of the covariance between shares *i* and *j* (covariance being $\gamma_{\varepsilon ij}$) multiplied by their respective weightings (Markowitz, 1952b). The variance of a portfolio is a measure of how the actual returns of the group of securities fluctuate. It is expected that the lower the correlation between securities, the greater the benefit of diversification will be. This occurs because the addition of shares to a portfolio with a low correlation to shares already in the portfolio will push the overall portfolio variance lower.

The return and variance equation are optimised under the following constraints:

$$\sum_{i=1}^{M} Xi = 1 \qquad (11)$$

$$Xi \geq 0, i = 1, \dots, M \qquad (12)$$

Equation 11 ensures that the sum of the weights of the potential shares is equal to 1. In other words all resources are utilised in the portfolio choice. Equation 12 ensures that all weights can only range between 0 and 1. This restricts the portfolio choice to one that does not include short selling shares nor allowing leverage (where the sum of the weights of the shares could equal to larger than 1, as one is borrowing money to increase positions taken on certain shares) (Freitas, Souza & de Almeida, 2009). The restriction on short selling is fitting to the current study as the JSE, the exchange on which the study is performed, does not allow for the short selling of shares (JSE Equities Rules, 2016). The aim of this dissertation is thus to find the optimum number of shares needed in order to minimise the risk inherent in a portfolio.

All else being equal, it can be said that the higher the correlations among security returns, the greater the standard deviation of the portfolio as a whole. In other words, the more the returns on individual securities tend to move in the same direction together, the less do variations in other individual securities ‘cancel out’ these movements by moving in the opposite direction, this leads to a greater variability (standard deviation) in the portfolio as a whole. If however individual securities tend to move in opposite directions, the inclusion of both to a portfolio should decrease the variance of the portfolio. It is also noted that diversification is extremely powerful when shares are uncorrelated. Even when large portfolios are considered, if the shares included are highly correlated, diversification can have only a limited impact on risk reduction. It can thus be said that a security adds much or little to the variability of a portfolio based on the sum of all of its covariances rather than on its own variance (Markowitz, 1959).

Generally, one finds that as the number of shares included in a portfolio increases, the variance of the portfolio decreases at a decreasing rate. This is displayed in Table 3 below:

Table 3: R*elation between the Number of Securities in a Portfolio to the Expected Portfolio Variance*

Number of Securities	Expected Portfolio Variance
1	46.619
2	26.839
4	16.948
6	13.651
8	12.003
10	11.014
12	10.354
14	9.883
16	9.530
18	9.256
20	9.036
25	8.640
30	8.376
35	8.188
40	8.047
45	7.937
50	7.849
75	7.585
100	7.453
125	7.374
150	7.321
175	7.284
200	7.255
250	7.216
300	7.190
350	7.171
400	7.157
450	7.146
500	7.137
600	7.124
700	7.114
800	7.107
900	7.102
1000	7.097
Infinity	7.058

Note: The Effects of Diversification, Adapted from: "Modern Portfolio Theory and Investment Analysis", by Elton, Gruber, Brown and Goetzman, 2009

Table 3 shows the relationship between the number of shares in a portfolio and the portfolio's variance. This data was analysed from U.S. equities – this included monthly share price data on all shares listed on the NYSE in 1975. As described previously, the variance of a portfolio cannot be completely diversified away but merely optimised as to remove the unsystematic risk inherent when purchasing individual shares. The systematic risk in the above sample can thus be seen to be approximately equal (as it tends to) to the variance number of 7.058. It can be observed that the greatest difference, when aiming to minimise portfolio variance, is felt when between 10 and 12 shares are included in the portfolio (in this sample). Thereafter the effect of adding another share to the portfolio becomes decreasingly significant. This can be seen graphically in Figure 7 below:

Figure 7: Graphical Representation of the Relation between the Number of Shares in a Portfolio to the Portfolio Variance

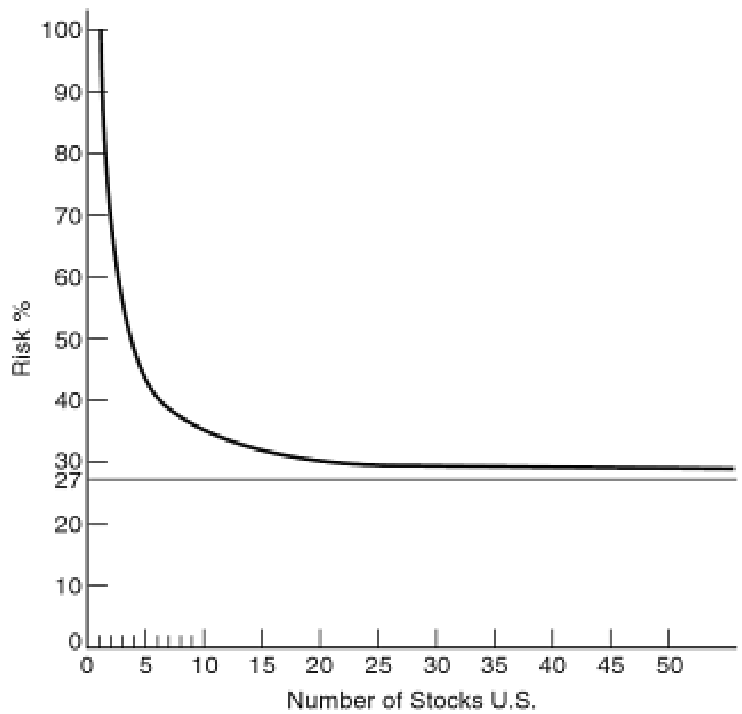


Figure 7: *Graph displaying the Number of Stocks vs Risk (variance) on the sample shown in Table 3.* Adapted from: "Modern Portfolio Theory and Investment Analysis", by Elton, Gruber, Brown and Goetzman, 2009.

The above section allows one an in depth understanding of the Markowitz Portfolio Theory. In addition the importance of discovering the number of shares to include to a portfolio in order to achieve optimal diversification is clearly demonstrated. The numerous findings of Markowitz: that diversification can reduce but never eliminate risk, a portfolio with a maximum expected return does not necessarily imply that it has a minimal variance and finally that not all diversification will pay off, are outlined and thereafter the formulas utilised to analytically describe the findings are presented. The theory of the efficient frontier is explored and thereafter graphically displayed. Finally the graph of the effect that adding securities to a portfolio has on the variance of that portfolio is shown.

Despite the significance highlighted above of discovering the optimal number of shares to include into a diversified portfolio there are instances when investors have been found to be holding sub-optimally diversified portfolios. This can be partly attributed to inherent irrational beliefs and behaviours implemented by investors. This chapter continues with an inspection into the various behavioural factors that can lead to sub-optimal diversification.

2.5 Behavioural Factors Leading to Sub-Optimal Diversification

It is interesting to note that when studying investors' actual portfolios, a common anomaly tends to appear – that investors are in fact not 'properly diversified' (Statman, 1987). The concept of being 'properly diversified' refers an occurrence whereby the investor avoids superfluous diversification – that is, diversifying across shares without much added benefit of the portfolio's risk being decreased (Statman, 1987). Superfluous diversification is balanced against the investor being overly focused in his investments, such as being overly weighted in particular categories such as in certain industries, geographies, sectors and the like.

Jacob (1974) made the assertion that an investor can decrease the number of shares needed to properly diversify their holdings by choosing them judiciously. Attention must be paid to the industries, geographies, sectors and the like that the investor is investing in so as to actively seek a portfolio that is not too heavily focused on one particular factor. Furthermore, it was acknowledged that perhaps the reason that investors tend to be under-diversified is that their share portfolio only makes up a small proportion of their wealth holdings; with the remainder of their wealth being invested in bonds, real estate and the like (Jacob, 1974). However, King and Leape (1998) found that after completing a wide survey, investors appeared not to be properly diversified even after accounting for other assets in the investors' portfolios. Transaction costs were then included into their model but they concluded that the under-diversification anomaly still remained.

Shiller (1990) questioned American and Japanese investors and required them to value each other's markets as well as their own. Significant differences were found, with both investors valuing their own market higher than the others which led to sub-optimal diversification. This could be a display of national pride but also purely because it is natural to overweight factors that we have more knowledge of.

This reinforces the results found by French and Poterba (1993) that in spite of the increasing proof that international diversification tends to pay off, the majority of investors still hold the bulk of their holdings in their domestic country. Notably these findings were shown to be the result of investor choices rather than of institutional constraints. They hypothesise that this could be due to the homecoming bias – the empirical anomaly in finance that states that domestic investors may prefer to invest in domestic rather than international securities. This behaviour occurs despite the fact that it may lead to the selection of suboptimal portfolios which may further lead investors to lose out on the benefits that international diversification may offer (French and Poterba, 1993). The table below shows explicitly that investors tend to hold more domestic assets as compared to non-domestic assets leading them to be under-diversified.

Table 4: *Portfolio Distribution Achieved by Investors Across Geographic Locations*

Equity Portfolio Weights: British, Japanese, U.S. Investors

	Portfolio Weight			Adj. Market Value
	U.S.	Japan	U.K.	
U.S.	.938	.0131	.059	$2941.3
Japan	.031	.9811	.048	1632.9
U.K.	.011	.0019	.820	849.8
France	.005	.0013	.032	265.4
Germany	.005	.0013	.035	235.8
Canada	.010	.0012	.006	233.5

Note: Estimates correspond to portfolio holdings in December, 1989. They are based on the authors' tabulations using data from the *U.S. Treasury Bulletin* and Michael Howell and Angela Cozzini (1990). Adjusted market values exclude intercorporate cross-holdings from total market value, and correspond to June 1990 values.

Note: Adapted from Equity Portfolio Weights: British, Japanese, U.S. Investors, adapted from "Investor Diversification and International Equity Markets", by French and Poterba, 1993

The table above illustrates the cross matrix of the portfolio weightings held by investors based in various countries. The left vertical axis shows the countries whose holdings were examined while the horizontal axis demonstrates the respective weightings held by each country. It can be seen, for example, that of the U.S. based investors analysed, their average holdings are 93% based in the U.S. In addition 1.31% of their holdings are focused in the Japanese markets and 5.9% are held in U.K. equities. The additional countries holdings (Japan, U.K., France, Germany and Canada) can be interpreted in a similar fashion. These notably do not sum to 100%, however, it is noted in the study that the figures are estimations and do not account for factors such as intercorporate cross-holdings (French and Poterba, 1993).

Further explanation of the sub-optimal diversification anomaly is provided by Kahneman and Tversky, who are widely viewed as pioneers in the field of behavioural economics largely due to their 1979 paper on prospect theory (Kahneman & Tversky, 1979). Prospect theory describes the hyperbolic discounting curve that is a mental model utilised by people which runs contrary to the traditional theories of utility theory. Prospect theory essentially states that people display risk aversion in choices involving sure gains and display risk seeking behaviours when facing choices involving sure losses (Kahneman & Tversky, 1979). In this way prospect theory can play a major function in explaining why investors may be over or under diversified. If an investor is facing a large probability of an extreme loss to his portfolio, instead of diversifying the portfolio as would traditionally be thought to be rational, the investor would leverage up his position on a selection of very few shares based on the small probability that this decision could lead to large, market beating gains. Similarly, in the situation when the investor is faced with a decision involving an almost sure gain, instead of leveraging up the position in a bid to further enhance gains (albeit while assuming increasing amounts of risk) the investor would prefer to over-diversify his portfolio in an attempt to refrain from losing any value, irrespective of the outcome of an event.

Kahneman and Tversky (1984) further illustrated that sub-optimal diversification decisions can be partly attributed to the way in which diversification is commonly referred to. Kahneman and Tversky (1984) revealed that a decision is preferred if it is structured as the full protection of one event rather than the partial risk reduction of the overall risk, despite that they may be mathematically equal. Diversification is typically referred to as being a tool that can be utilised to reduce (and optimally minimise) the overall risk of a portfolio. In this way the decision to employ diversification is presented as a relatively less appealing option compared to the instance wherein specific strategies could be said to fully eliminate certain risks afflicting a portfolio.

There are numerous additional behavioural factors that are understood to potentially effect diversification decisions. These include the self-attribution bias, overconfidence and over-optimism (thinking the securities one selects will outperform), anchoring, availability, interference theory, herding, confirmation bias and mental accounting. Each factor/bias is discussed in turn in the section below.

Self-attribution bias refers to situations in which people overestimate the degree to which they are responsible for their own success (Baker & Nofsinger, 2010). This bias can occur if an investor had selected a portfolio of shares that performed excessively well but is not properly diversified. This scenario may lead an investor to attribute their portfolios performance to their share selection skills rather than 'luck' or other potential variables. In this case it is unlikely that an investor will elect to adjust their current portfolio to a more diversified one as they may hold the belief that they do not need to diversify their portfolio to the same extent as other investors due to their superior share selecting abilities.

The self-attribution bias can routinely lead to an overconfident belief system. Overconfidence refers to an instance in which one believes they are right more often than they realistically are (Baker & Nofsinger, 2010). This behavioural bias works in a similar manner to the behavioural bias of self-attribution - if an investor wrongly (or rightly) believes that the high returns on their portfolio are due exclusively to their personal share selections, they may become overconfident in their abilities. This may in turn lead to an investor downplaying the benefits of diversification as the belief is held that their portfolio selection ability is better than the markets, referring to all other investors. In this way overconfidence can lead to a sub-optimally diversified portfolio.

A related behavioural factor is that of over-optimism which refers to the tendency of over exaggerating one's own abilities while holding an overly hopeful outlook (Tversky & Kahneman, 1990). This could lead to a decrease in the diversification of an investor's portfolio as an investor may hold exaggerated ideas of how the economy and selected shares will perform. Consequently this could lead to an underweighting of the potential negative outcomes of their portfolios which may in turn lead to a downplaying of the benefits of diversification, resulting in an under-diversified portfolio.

The fourth behavioural bias is termed anchoring. Anchoring refers to the occurrence when initial information (whether correct or incorrect) impairs ones judgement (Tversky & Kahneman, 1990). Anchoring can lead to 'stickiness' of an ideological system. For example, if an investor believes that the optimally diversified portfolio consists of a minimum of 30 shares, they may disregard current research that states that the optimal portfolio is made up of less (or more) shares. In the same light, if an investor believes diversification is unnecessary, literature stating the opposite may not affect the investor's behaviour.

Anchoring is a behavioural bias that can influence an investor in tandem with, or separate from, the availability heuristic. The availability heuristic refers to the behaviour that current and future probability outcomes anticipated by an investor are dependant on past experience (Tversky & Kahneman, 1973; Tversky & Kahneman, 1990). Information that is more memorable (and thus more easily available to recall) will lead to the possible incorrect conclusion that a secondary event is more or less likely, based on the availability of occurrences of similar previous events. This bias can lead, as described with anchoring, to an individual basing current and future behaviour on past experience. This process can occur both consciously and unconsciously.

The availability heuristic poses multiple threats to academic literature that may indicate that an investor's current behaviour could be suboptimal – as the impact of the literature could be drastically underweighted due to an investors recent experiences. For example, if an investor has experienced a large degree of success by employing diversification in a portfolio consisting of 30 shares, they may be unwilling to modify their strategy even in the face of new literature that may demonstrate the behaviour as being suboptimal. Similarly, if an investor has experienced success devoid of diversification in the past, they may be unwilling to adjust to a strategy that includes diversification. Conversely, this bias could enhance the impact of literature as in the case where an investor encountered an extremely negative loss as a result of not possessing a diversified portfolio, such as those experienced during the 2008 financial crises. This could lead to an investor being aggressively inclined to favour the selection of a portfolio that includes diversification post the undesirable event. This predisposition is due, in part, to their previous negative experience which may be potentially reinforced by the finding that people are more averse to losses than they are pleased with gains (Kahneman & Tversky, 1979).

Linked to the availability heuristic is interference theory. Interference theory relates to the phenomenon that past memories can have an influence when learning new material (Nevid, 2013). There are two categories of interference theory: retroactive and proactive. Retroactive interference refers to when interference affects a memory after it is learned but before it is consolidated and recalled (Nevid, 2013). This is commonly observed in the situation when learning more information leads to remembering less past information. Of more interest in this study is proactive interference. Proactive interference refers to an occasion when new material attempting to be learned is influenced by old, previously learned material (Nevid, 2013). In this way, new memories can interact and thus alter older memories. This can affect investor choice as selective memory can alter investment decisions as these are dependent on the type of memory that is recalled (for example losses when holding less than 30 shares). If a memory is readily available then the effect on investor behaviour will be compounded, in line with the availability heuristic described above. Proactive interference could similarly create a barrier to learning and adapting to new information that was previously thought to be incorrect.

Investor decisions relating to diversification can also be affected by a behavioural bias known as herding or groupthink. Herding or groupthink refers to an occurrence when an individual in a group is reluctant to challenge the conventional wisdom of the group, even when presented with contradicting evidence (Wilcox, 2010). This could impact on the behaviour of investors as if they are involved in an environment that is specifically aimed against or in favour of diversification, it is highly likely that the investor will adopt similar beliefs and behaviours as those around them. This would directly impact on the level of diversification obtained by an investor but notably excludes the case when an investor may follow a contrarian strategy. A contrarian strategy is identified by the buying and selling of securities that stand in opposition to current market sentiment, for example buying shares that the majority of investors are selling due to their belief that the price will progressively move lower.

Additionally, confirmation bias can affect the decisions that an investor may make with regards to diversification. Confirmation bias refers to the human tendency to interpret or remember information that confirms ones' own current preconceptions (Miller, Vandome, & McBrewster, 2009). In other words, the seeking out of information which agrees with one's ideas, values and beliefs while simultaneously disregarding evidence that points towards a contradictory belief (Nickerson, 1998). It can be easily appreciated as to how this can detrimentally affect an investor's portfolio; investors will tend to disregard new evidence that disproves their current investment models or supports a new method that illustrates how to achieve optimum diversification. This evidence would, by way of the confirmation bias, either be ignored or underweighted implying that a large amount of time may be necessary for new research to be accepted as the norm in the investment industry.

The final behavioural factor discussed in relation to the sub-optimal diversification anomaly is that of mental accounting. Mental accounting was recognised by Black (1982, as cited in Statman, 1987) who reasoned that people tend to invest their money in different ‘mental pockets’. Each of these pockets has an associated risk preference to it. This notion supports the under-diversification anomaly (sub-optimal portfolio diversification) as investors do not tend to value an asset based on how it can contribute to their portfolio but rather as a standalone investment decision. This could result in a portfolio that may be comprised of positively performing assets that are, for example, concentrated on one industry – leading to the overall portfolio exhibiting under-diversified. In this way, Black (1982) made the assertion that a lack of diversification does not imply a lack of education or of investor sophistication but rather, could be due to the ingrained human bias to view financial decisions in different mental pockets.

The discussion above highlighted evidence displaying the numerous occurrences where investors’ portfolios have demonstrated sub-optimal diversification properties and additionally several behavioural factors that the anomaly could be attributed to were investigated. It is important to note the aforementioned behavioural biases that could affect investor behaviour. As indicated earlier, genetic programming and particle swarm optimisation algorithms are designed to constantly learn from data. These approaches have proven to be efficient in recognising patterns in a data set (for example, determining if a share needs to be sold based on a predefined rule set). It is possible however, that a human element could inadvertently be included into the algorithm (as the programmer of the algorithm sets predefined rules). For example, literature has highlighted a behavioural anomaly - that people are more averse to losses than they are enthusiastic over gains (Markowitz, 1952a; Fischer & Jordan, 1995; Kahneman & Tversky, 1984). This bias could lead to suboptimal investment decisions if included by way of a predefined rule introduced by the investor/programmer. This bias could also be inadvertently included if the investor/programmer interprets the output signal incorrectly or even disregards the output signal intentionally due to a behavioural bias. Therefore, it is necessary to ensure that the algorithmic process is scrutinised to ensure that behavioural biases do not creep in. Providing these are avoided, the various studies described previously offer strong evidence in support of why genetic programming and particle swarm optimisation can be used to solve for optimal portfolios of assets as well as in potentially forecasting the returns of these portfolios.

2.6 Summary

In Chapter 2 above, the past literature relating to diversification is thoroughly examined. Both types of diversification, as it relates to assets and to products, were inspected for insight and the study continued with a review of how previous authors utilised and implemented the traditional method as well as the two artificial intelligence approaches namely: genetic programming and particle swarm optimisation. The studies were carried out using various methodologies, while the majority also differed over the sample period used; this took the form of using a different time period and (or) analysing differing stock markets. Despite these differences, the studies were able to illustrate the vast differences in the number of shares indicated as necessary to obtain an optimally diversified portfolio. The method in which the current dissertation aims to improve on previous literature was discussed but will however be further elaborated on in subsequent chapters. The optimally diversified portfolio was established to consist of from as few as 5 shares to as many as a minimum 30. The comparison analysis uncovered key insights with relation to how the number of shares needed to obtain an optimally diversified portfolio has changed over time, stock markets and the like. The chapter continued with an elaboration on the Markowitz portfolio theory and finally, the behavioural factors that may impact an investor's selection of assets were highlighted. The dissertation now continues by expanding on the data utilised in the current study as well as further elaboration regarding the methodology behind the two artificial intelligence approaches (genetic programming and particle swarm optimisation) and the way in which each was employed to answer the research questions posed in this research.

Chapter 3: Data and Methodology

3.1 Data

The data employed was sourced through the Bloomberg Professional Platform. The data utilised was the monthly closing share price of each company that was listed at any time on the JSE Main Board over the period from December 1994 – December 2014 (this includes shares that may have been delisted over the period). The share prices were taken as on the 25th of each month starting from the 25th December 1994 and ending on the 25th of December 2014. The day of the 25th of each month was selected, instead of the last day of each month, as to avoid the potential bias of synchronicity of price movements and distortions due to month end trading. On months where the 25th fell on a non-trading day (as in the case of a Saturday or Sunday as well as on a public holiday) the closing price from the closest previous trading day was utilised.

Once the complete list of respective companies had been established, three adjustments were made to their closing share prices (before calculating the return series), in an attempt to obtain the cleanest data set possible and properly account for corporate actions. Firstly, normal cash adjustments were made. This included adjusting the historical share prices for any interim dividends and special dividends, interest on capital as well as partnership distributions. Secondly, abnormal cash adjustments were made. This entailed adjusting the historical share prices to reflect liquidation, capital gains, rights redemptions as well as proceeds or warrants. Finally, capital changes were taken into account which included adjusting the historical share prices to reflect spin-offs, share splits and/or consolidations as well as any rights offerings.

Over and above the three adjustments made to the data mentioned above, there were three further adjustments made in order to clean the data. Firstly, companies whom were on the list but however had zero observations over the entire 20 year period were removed from the data set. Thereafter, companies whom had share price below 100 cents at any point over the 20 year period were removed from the data set. This was due to the sporadic and random nature of these ‘penny stock’ type shares (as the share prices can be drastically changed, in terms of return, on a day to day basis due to single small trades on the company). Lastly, if a share was listed for part of the time period and was subsequently delisted at any stage (or was listed after the beginning of the subset period) the return series was made to equal 0% for the periods that the company was unlisted. Furthermore, if a share did not trade over a monthly period, a 0% return was used. By including companies that were listed (delisted) over the full time period, survivorship bias is avoided.

The preliminary data set, before cleaning, consisted of 984 companies. The three initial adjustments mentioned above were completed (which included normal and abnormal cash adjustments as well as incorporating capital change adjustments). Thereafter, the three additional adjustments were made which caused numerous firms to drop out of the sample. Once the data was cleaned as described in the processes above 356 companies remained.

The remaining population, consisting of 356 companies, were split across 11 sectors in the proportions shown in Figure 8. The below chart is constructed by analysing the composition of the population of companies based on the number of companies in each sector. It can be observed that the majority of shares were concentrated in the financial sector, representing 28% of the population. Furthermore, 73% of the population can be represented through the combination of the top 4 weighted industries. This representation grows to 83.43% when the next heaviest weighted industry is included into the analysis. The utilities sector is represented by merely 1 share, which equates to 0.28% of the population; as such it is shown to round to 0% in the figure below. When analysing the strength of each mathematical approach in solving for the minimum variance portfolio utilising an optimal number of shares, the above analysis will be included as a technique allowing one to assess if the population is adequately represented through the optimally diversified portfolio selected by each mathematical approach. This will be further explained in the methodology section to follow.

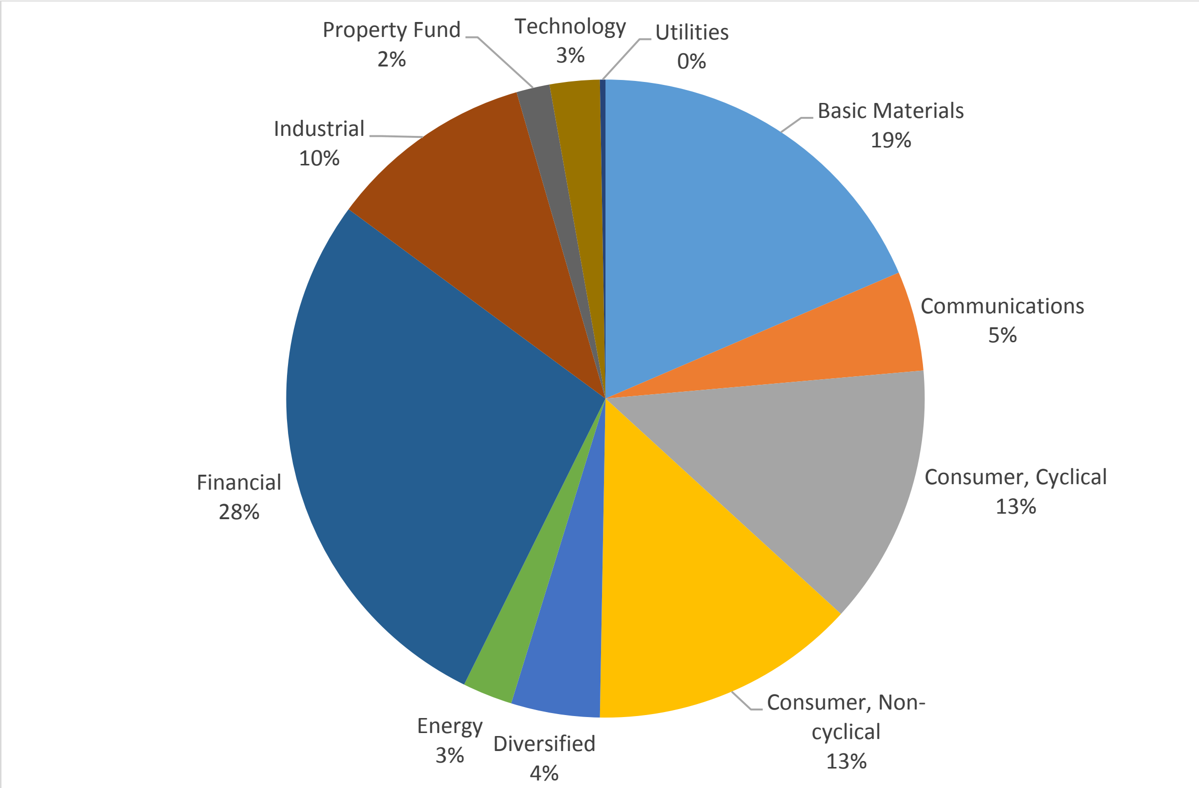


Figure 8: Population Split by Industry

After completing the above adjustments, the return series of the remaining shares was calculated using the standard return formula:

$$R_t = \left[\frac{P_t - P_{t-1}}{P_{t-1}}\right] * 100 \qquad (13)$$

This is a repeat of equation 3 but expressed as a percentage. R_t is the return at time t, P_t is the closing share price at time t and P_{t-1} is the closing share price at time t-1. In this instance, P_t is the current month's closing price and P_{t-1} is the previous month's closing price. The return series thus ran from 25 January 1995 – 25 December 2014. The return series is used in place of the absolute share price values, as in previous literature, in order to avoid the standard deviation of a share being larger purely due to higher share prices and lower for those with lower share prices respectively. Utilising the share price returns overcomes this and allows the standard deviation of shares to be compared like for like. It was thus on the return series that the analysis was performed. As mentioned above, shares that were not listed or did not trade over any specific month to month period were allocated a 0% return for the respective month.

3.2 Methodology

The traditional simulation approach, the Genetic Programming approach and the Particle Swarm Optimisation approach are each applied to the full 20 year data set in turn. The analysis of each approach is applied using four tests – in essence, four tests are run for each mathematical approach; thus 12 tests were run in total. Thereafter, tests to assess the stability of the results were run on the equally weighted pre-specified portfolios, to analyse the variability of the outputs using each approach. Stability testing gives one further insights as to the strength of the application of the 3 mathematical approaches. Through stability testing one can assess whether a method converges to the optimal answer or whether the answer is found by pure chance – which would lead to a large variability in the output with little stability of the results. This will be further expounded upon below.

The first test to be run with each approach solves for the minimum variance portfolio at each number of share, ranging from 1 to 30, while maintaining equal weightings. In this way the portfolio is set so as to ensure that when the portfolio is restricted to consist of only 2 shares, each share will carry a weight of 0.5. When the portfolio is restricted to consist of 3 shares, each share will carry a weighting of 0.333 and so on. This simulation continues under the constraints previously mentioned.

The second test solves for the minimal variance portfolio with no restrictions on the number of shares that can be utilised in any given trial but while still maintaining equal weightings on each share selected. In this test the opportunity set available for each iteration in the simulation is greatly increased as now the portfolio opportunities are not limited to a contain a specific number of shares. In one iteration a portfolio of 4 shares may be selected (with a weighting of 25% on each) while in the next iteration a portfolio of 10 shares may be selected (with a weighting of 10% on each share).

This test however carries the following additional constraints. The portfolio must contain at least one share – this constraint ensures that the optimisation process does not solve for a portfolio containing 0 shares and therefore yielding a minimum portfolio variance of 0. Furthermore, a statistical transaction cost is taken into account to prevent the optimisation process from consistently adding shares to the portfolio with each additional share having an extremely limited impact. This was achieved by ensuring that the best solution found (the portfolio yielding the minimum variance) is better than the previously solved for best solution by at least a measure of 0.01%, or else it is not used.

Test three is similar to the first test but with the added requirement of equal weightings being applied to each share being waived. This test consists of solving for the minimal variance portfolio for each number of share, again ranging from 1 to 30, this time however removing the restriction of equal weightings on each share in the portfolio. This again increases the number of potential portfolios that can be constructed exponentially as in this test a portfolio consisting of 2 shares can be constructed in numerous different ways – for example a 75% allocation to the one share and a 25% allocation to the other, an 80% allocation to one share and a 20% to the other and so on. In order to somewhat limit the opportunity set that the optimisation process can solve for, the weightings on each share were constricted to 3 decimal places and were allowed to vary by a minimum measure of 0.01.

The final test consists of solving for the minimal variance portfolio with no restrictions on the number of shares that can be utilised, whilst no longer requiring equal weightings on shares to be used. This test carries the same restrictions as the second test, namely that the portfolio must contain at least one share; a statistical transaction cost is taken into account to prevent the optimisation process from consistently adding shares to the portfolio with each additional share having an extremely limited impact - this was again achieved by ensuring that the best solution found (the portfolio yielding the minimum variance) is better than the previously solved for best solution by at least a measure of 0.01%.

Each of the four tests were applied using each of the three approaches (the traditional simulation approach, genetic programming approach and particle swarm optimisation approach) in turn. The tests that construct a portfolio at each number of share (tests 1 and 3), ranging from 1 to 30, allow the traditional simulation methodology utilised by Neu-Ner and Firer (1997) to be compared to the Genetic Programming and Particle Swam Optimisation approaches in order to discover if they can provide additional benefits to diversification through a decreased portfolio variance.

The minimum variance reached through each mathematical approach is thereafter compared in order to find out if the more advanced methods, being genetic programming and particle swarm optimisation, offer additional significant value in further reducing portfolio variance compared to the more traditional simulation method. The simulations that solve for the minimum variance portfolio without solving for a particular number of shares (tests 2 and 4) allows for the strength of the mathematical models to be more transparently compared to one another as one can accurately view which method solves for the minimum number of shares to include in the portfolio that yields the minimum variance in a relatively less constricted environment – one with exponentially more options to select in each iteration. The time allowance for each simulation is set to 10 minutes and the sum of the weightings in the portfolio is set to equal 1 at all times while restricting short selling – where weights on a share may be below 0.

The 10 minute time allowance was selected through the application of initial exploratory testing. A sample of tests were run, applying each mathematical approach sequentially, with time allowances of 2, 5, 10, 15, 30 and 45 minutes in turn. Across each of the tests it was demonstrated that the majority of the diversification benefit obtained by an investor was accumulated within the first 10 minutes. Below this point (shorter than 10 minutes) the investor continued to experience noticeable additional diversification benefits as the run time approached the 10 minute mark, demonstrated by a lower portfolio variance. Post this point, incremental diversification benefits to the investor were not significant and in numerous cases were unchanged to those achieved at or before the 10 minute mark. Figures displaying this occurrence are observed and discussed further in Chapter 4.

In summation, this dissertation involves testing the data in four ways. Firstly the minimum variance portfolio is solved for iteratively at each number of shares to be included in the portfolio, ranging from 1 to 30 systematically. This first test carries the additional restriction that each share included in the portfolio is equally weighted. The results of each approach are tested against the hypothesis - the optimal number of shares to include in the minimum variance portfolio should lie between 10 and 30. The second hypothesis is also tested by comparing the minimum variance found using the more advanced mathematical models to the results obtained through the more traditional simulation approach. An additional strength test is carried out on each approach using the results of the test above. The portfolios that are restricted to consist of 5, 15 and 25 shares respectively are simulated not once, but five times. The repetition of simulations is an additional test that is run which provides additional insight to further asses the strength of the respective mathematical approaches. A strong mathematical approach will display results where the minimum variance on each simulation should be extremely similar, if not equivalent. Slight differences could occur if the portfolio selected consists of slightly different constituents that previous or future simulations – such differences when using a strong mathematical approach should prove to be marginal. If these differences are substantial, one can assert that the mathematical approach being used does not display strength in the calculation process.

Following which the second test will be run. This test solves for the minimum variance portfolio while maintaining equal weightings on the shares included in the portfolio; this test however carries a notable difference to the first. This second test no longer pre-specifies the number of shares that should be included into the portfolio. This exponentially increases the potential combinations of shares that can be tested within any given simulation. For example, in the first run of the simulation a portfolio can be constructed to contain 5 shares, in the next run the portfolio may contain 40 shares and so on. The inclusion of this test allows one to more optically view the performance, and thus the strength of each of the three approaches. The results of this test are examined through the use of multiple techniques. Firstly the minimum variance portfolio is examined both with regards to the resulting variance of the portfolio, as well as with regards to the number of shares that were utilised to achieve this minimum variance. If the number of shares utilised lies between 1 and 30, the result is compared back to the corresponding portfolio constructed in the first test. If the portfolio variance solved for is close to or better than the one solved for in test 1, the approach can be said to display strength in the calculation process. Again, these results could be slightly different to those obtained in test 1 due to different underlying shares being selected to be included into the portfolio. Furthermore, the restriction on time allowed for each simulation combined with the exponentially increased number of potential simulations that could occur provide a more rigorous strength test for the mathematical approaches.

The third test is similar to the first in that it calculates the minimum variance portfolio iteratively at each number of shares (ranging from 1 to 30), however, this test no longer requires the shares in the portfolio to be equally weighted. The potential portfolios grow exponentially in this approach as previously a finite amount of portfolios could be constructed, however, the potential possibilities are now far more abundant. For example, the population consisted of 356 companies, to be further elucidated upon in a further section; when constructing a portfolio restricted to 2 shares of equal weightings the number of potential portfolios is 126 736 calculated by the square of the number of individual shares in the population. When the weightings on each share in a portfolio are allowed to vary, even when assuming a restriction of 3 decimal places on each weighting, the number of potential combinations in a portfolio constructed of merely two shares grows exponentially. This allows one to further examine the strength of each approach through a comparison to test one, both by examining the performance of each approach in calculating the minimum variance portfolio as well as allowing one to examine the stability of the series – the Markowitz theory (discussed previously in Chapter 1) illustrates that as more shares are added to a portfolio, so the resultant portfolio should display a lower minimum variance.

The fourth and final test is similar to the second test in that the variance portfolio will be found overall without the number of shares to be included in the portfolio being pre-specified. Although, in this test the restriction of equal weightings on each share is removed. The results of this test are analysed with the same methodology applied in test 2, with the comparison now being carried out against test 3, where the number of shares to be included in the portfolio were pre-specified while allowing the weightings of each share to vary. In all simulations, the sum of the weightings on each share in the portfolio is required to be equal to 1.

The above analysis allows one to answer the two key research questions in this study, as highlighted in the previous section. Firstly, "what is the minimum number of shares that can be included into a portfolio in order to achieve the optimal level of diversification?" This research question is answered using two approaches, the first and more traditional requires the weightings on each share in the portfolio to be equal and sum to 1. The second approach however allows the weighting on each share to vary – this constitutes a novel addition to previous literature. The weights on each share, as mentioned previously, must vary between 0 and 1 and in addition must sum to 1. Thereafter, the second research question is addressed: "Will the more advanced mathematical models of Genetic Programming and Particle Swarm Optimisation reveal new findings compared to the simulation based approach of previous literature?" The comparison of results achieved when applying each mathematical approach in turn, namely: the traditional simulation approach, the Genetic Programming approach and the particle swarm optimisation approach constitutes a further addition to previous literature as comparative studies in this regard are relatively limited as indicated in the literature review section.

Markowitz portfolio theory states that, as one adds shares to a portfolio, the portfolio's variance should decrease. The decrease in variance should initially be significant but should gradually begin to be increasingly marginal as more shares are added past a point – this point is said to be the optimal point of diversification (Markowitz, 1959). The first hypothesis is that this optimal point portfolio is expected to consist of between 10 and 30 shares, in agreement with previous literature. Secondly, it is hypothesised that allowing the weights on each share to vary could discover previously hidden additional benefits of diversification due to the exponentially increased opportunity set that the algorithms can construct. The final hypothesis is that the advanced mathematical models of Genetic Programming and Particle Swarm Optimisation will yield better results (meaning a lower portfolio variance) than the more traditional simulated approach. This will add to the relatively small amount of literature that compares the Genetic Programming approach to the Particle Swarm Optimisation approach, and further compares both back to the traditional simulation approach that was utilised in numerous papers in the past.

As previously mentioned, once all the simulations have been run, using each of the three mathematical approaches in turn, not only will the results be compared but stability testing will also be carried out. This is completed by analysing the first and third tests and entails running the same simulation five times on the sets when the portfolio is restricted to contain 5, 15 and 25 shares respectively. This is carried out on both test 1, where the weights on each share are restricted to be equal to one another, as well as on test 3, where the weightings on each share are allowed to vary. Each of the three mathematical approaches are applied in turn. A stable optimisation process should yield the outcomes showing little variance of the results between simulations when the same constraints are utilised. Finally, when the mathematical approaches are applied to test 2 and test 4, their respective portfolios will be analysed and compared with regards to inspecting the diversification over potential industries included in the population portfolio.

In order to more efficiently calculate the portfolio variance at each level of share as well as when the overall portfolio variance is considered, matrix multiplication is utilised as per the equation below:

$$\sigma_p^2 = \begin{pmatrix} w_1\sigma_1 & w_2\sigma_2 & w_3\sigma_3 & w_4\sigma_4 \end{pmatrix} \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix} \begin{pmatrix} w_1\sigma_1 \\ w_2\sigma_2 \\ w_3\sigma_3 \\ w_4\sigma_4 \end{pmatrix} \qquad (14)$$

Equation 14 is the matrix multiplication extension of equation 10. As in equation 10, $\hat{\sigma}_p^2$ represents the portfolio variance while *w* represents the weight on each share and is the standard deviation of the respective share. *P* in equation 14 (equivalent to $\gamma_{\varepsilon ij}$ in equation 10) represents the correlation between two shares, as noted by the subscripts associated beneath *P*. Using the matrix multiplication approach allows one to more efficiently compute the minimum variance portfolio when the number of shares to be included to the portfolio is pre-specified (ranging from 1 to 30), as well as when solving for the overall minimum variance portfolio (when the number of shares to be included to the portfolio is not pre-specified) – the same formula can be simply applied in this case as, all shares not being used can still be included in the equation by assigning them a zero weighting. Thus a new equation is not required to be evaluated for each additional share included into the optimal portfolio, but rather the three approaches need only to adjust the weightings of the shares for the same formula, under the respective restrictions.

The study aims to find the optimally diversified portfolio, in other words the portfolio with the least variance, by using the minimum (or pre-specified) number of shares – as such the study is retrospective or backward-looking. There is potential in future studies to build in a predictive element, and as such develop a forward looking ability rather than the retrospective approach used in this dissertation, which could effectively indicate to the investor how many shares are expected to provide the optimally diversified portfolio in the future and furthermore which shares this portfolio should comprise of, however developing a forward looking simulation is not the aim of this dissertation. The final constraints that are applied to all of the simulations run are that, firstly, the possible weightings on each share will range between zero and one. This thus excludes the possibility of short selling a share as well as the possibility of purchasing a share with leverage (investing more capital than one has on hand). Furthermore, the sum of the weightings of the utilised shares must always be equal to one – thus all of one's capital must be invested in the market at any one time, assuming that the investors' portfolio consists of only equity and no other assets.

The way in which the full methodology and analysis was carried out in Microsoft Excel is explained thoroughly in a worked example found in Appendix 1. Appendix 1 demonstrates the technique by which the correlation matrix was calculated and thereafter the process in which the portfolio variance is calculated using a matrix multiplication approach combined with binary factors is elucidated upon. Finally, the way in which the various constraints were incorporated into each simulation respectively are revealed.

3.2.1 Implementation of Approaches

The traditional simulation approach as well as the Genetic Programming approach were implemented using the Palisade Student software package. This is a package that can be utilised within Microsoft Excel for building and running of mathematical models. In both approaches the default parameters were utilised. When implementing the traditional simulation approach the 'Risk Optimizer' plugin was used. When specifying the model definitions both the optimisation goal and the criteria to optimise was set to minimum while the constraints were added as per the constraints previously mentioned. Thereafter the settings were input. All default settings were used while the termination was set for each test respectively under the criteria previously stated, namely that of the time of the simulation allowed to run for a maximum of 10 minutes while in tests 2 and 4 the best solution had to in addition improve from the previous best solution by a factor of at least 0.01%. Furthermore, the simulation settings were set to run using an automatic number of iterations and the sampling type used was set to a Monte Carlo with a Mersenne Twister (default) over the Latin Hypercube methodology. These two sampling methods are similar yet contain some differences. Latin Hypercube sampling is a method to ensure that each probability distribution in a model is evenly sampled. This method was used in the past when the accessible computing speed was low as it allowed one to gain a stable output with a smaller number of samples than a simple Monte Carlo Simulation (Glasserman, 2004).

Latin hypercube sampling is a stratified sampling method and shows small advantages over the Monte Carlo simulation methods when the number of variables in a simulation is small, as well as the number of distributions and number of simulations to be run. When there are multiple variables and distributions and a large number of simulations can be run, Latin hypercube samples lose its advantages over the Monte Carlo method. Furthermore, the Latin hypercube sampling method requires a large amount of computing memory as each possible distribution must be sampled before the simulation begins. For example, if a data set has 100 distributions and 5 000 samples are to be run, one would need to generate, shuffle and store 500 000 random numbers prior to beginning the simulation (Glasserman, 2004). The Monte Carlo methodology, however, relies on repeated random sampling to run through the search space to obtain the optimal solution for a mathematical problem (Thomopoulos, 2012). Although this methodology does not ensure that each distribution and sample will be used in at least one iteration, it allows for a larger number of simulations to be run quicker (McLeish, 2011).

In the Genetic Programming simulations, the default settings were also utilised, aside from the constraints previously mentioned. In the model definition, the optimisation goal is set to the minimum while the adjustable cell ranges and constraints are set as previously stated. The sole setting altered from the default settings relates again to the time allowed for the simulation to be run, being set to 10 minutes. Tests 2 and 4 encountered an additional adjustment from the default settings, notably that the best minimum variance portfolio is required to be better than the previous best minimum variance portfolio by a factor of at least 0.01% or else the previous best solution is kept.

For the implementation of the Particle Swarm Optimization the software package provided by XLOptimizer was utilised. In a similar fashion to the Palisade Software, XLOptimizer allows for various mathematical models to be built and run within Microsoft Excel. Ensuring both packages are compatible and work as a plugin, aids to guarantee that neither approach is given preference or computational power advantages with regards to the speed of the optimisation.

Once again the default parameters were used, although these are slightly more numerous and require further in depth exploration. Firstly, the respective adjustable variables are added to the model and their type adjusted to binary (as previously discussed). The constraints are then added, similarly to those added to the Palisade software. Thereafter the specific scenario was chosen, namely Simple PSO, in order to run the Particle Swarm Optimisation analysis. The population variable, which depends on the complexity of the problem and usually range between 10 and 30 was selected at 20, which is the default value. Thereafter, a cognitive and social parameter are confirmed as well as the inertia weight and parameter gamma. The cognitive parameter controls the weight of the pull of each particle to the best position previously achieved by the same particle, this is referred to as *pbest* earlier in the dissertation. The default value is 2, which is the value that was selected. The social parameter controls the weight of the pull of each particle to the best position previously achieved by any particle of the whole swarm, this is referred to as *gbest* in the PSO section above. The default parameter is 2, which is what was selected. The inertia parameter controls the acceleration of each particle, again the default of 0.8 was selected. Finally, the gamma parameter controls the initial maximum velocities of the particles as the fraction of the design space (per variable) that can be travelled in one time step. The default value of 0.4 was selected. The initial placement of particles was also selected to be randomised. The final constraint added, as implemented in the previous approaches, was to limit the amount of time that the process can run for to be equal to 10 minutes. Furthermore, in the Particle Swarm Optimisation method, a large penalty function was utilised in order to ensure that hard constraints, such as the sum of all of the weights of the included shares equal 1, were adhered too. The penalty function is set at 100,000*X, 100 times stricter than the default 1,000*X.

3.3 Summary

This chapter provided further detail on the data utilised in the current study. This included an elaboration on the platforms employed to source the data as well as a discussion regarding the various adjustments completed in an attempt to obtain the cleanest data set possible. Following which the industry split of the resultant population of shares was examined. The industry split of the population portfolio will be compared to the industry split of the portfolios deemed optimally diversified by each mathematical approach in Chapter 4. This provides an additional technique allowing one to assess whether the resultant population is adequately represented by the portfolios deemed optimally diversified by each mathematical approach.

The chapter continued by expounding upon the formulas and the methodology employed in the current study. This included elaborating on the four tests applied across each of the three mathematical approaches as well as describing the stability testing to be carried out thereafter. A clarification on the software involved, the respective settings applied and parameters employed was then provided. The dissertation now continues in Chapter 4 which outlines the outcomes achieved when the above methodology was implemented.

Chapter 4: Analysis and Results

This section encompasses the evaluation of the results obtained through the application of the methodology described in Chapter 3. As described previously, each mathematical approach is applied to the data by means of employing 4 tests. This is completed with the aim of answering the two research questions. The first research question relates to: “what is the minimum number of shares that can be included into an equity portfolio, in order to achieve the optimal level of diversification in a South African environment?” A sub-question to this regards investigating whether allowing the weights on each share to vary uncovers previously undetected diversification benefits. The second research question focuses on enquiring as to: “which of the three selected mathematical approaches demonstrates the most strength when solving for the optimally diversified portfolio in a South African environment?” The second research question aims to ascertain whether the more computationally intensive and complex models of genetic programing and particle swarm optimisation will display increased strength in solving for the optimally diversified portfolio as opposed to the traditional simulation method.

Chapter 4 aims to answer these research questions through the computation and analysis of the results achieved by each test on each mathematical approach in turn. These results relate to solving for the minimum number of shares that can be employed to construct an optimal portfolio. With this in mind, Chapter 4 commences with a demonstration of the results achieved by the traditional simulation approach and continues with the results of the genetic programming and particle swarm optimisation approaches. The section initiates by presenting the results achieved by test 1 and test 3 across each mathematical approach. Test 1 and test 3 relate to the instance when the number of shares the portfolio should consist of was expressly pre-specified for each simulation, ranging from 2-30 shares, as previously described. Additionally, stability testing is carried out on each approach in turn. The chapter continues with an analysis of the results achieved by test 2 and test 4 across each mathematical approach; these tests relate to the scenarios when the number of shares the portfolio should consist of was not pre-specified, while ensuring that the sum of the weightings on each share included into the optimal portfolio should equal 1.

The chapter concludes with a section detailing direct comparisons of the results. This is initiated by a comparison between the results achieved by the simulation approach when the weightings on each share were required to be equal against the results achieved when unequal weightings on each share was permissible. This is followed by a similar comparison, focused on the results achieved by the genetic programming approach. These two comparisons allow for the first research question and corresponding sub-question to be decisively answered as one is enabled to discover the minimum number of shares necessary to optimally diversify one's portfolio, contrasted across equally weighted and unequally weighted portfolios.

The chapter continues with a comparison between the results achieved by the simulation approach with equal weightings and the results achieved by the genetic programming approach with equal weightings. This is followed by a similar analysis comparing the results achieved by both approaches when the weights on each share were allowed to vary. This permits the second research question to be answered. The chapter thereafter incorporates a discussion on the unconstrained portfolios – relating to the simulations where the number of shares to be included in the portfolio was not pre-specified. This is completed across approaches, both when equal weightings were required on each share as well as when weightings were allowed to vary. In this section all implications of the current study are addressed and expanded upon. Finally, the chapter concludes with an evaluation of the diversification benefits of the portfolios deemed optimal across each mathematical approach where the number of shares in the portfolio was pre-specified in each simulation, ranging from 2-30 as previously discussed. This examination differs from the analysis previously described as the diversification benefits in this section relate not to the minimum standard deviation achieved by the portfolio, but rather investigates how the portfolio is spread across various industries.

4.1 Constrained Analysis (Pre-Selected Number of Shares)

4.1.1 Simulation Results

4.1.1a Equally Weighted Portfolio

The first simulation consisted of solving for the minimum variance portfolio for each number of shares (ranging from 1-30) using the traditional simulation approach. This portfolio is restricted to be equally weighted and for the sum of the weights on each share to always equal 1. The results from this first set of simulations are shown in Appendix 2 and can be observed graphically in figure 9 (below). From the graph below it can be seen that although the trend is unstable (in certain instances adding a share to the portfolio in fact increased monthly variance) the anticipated downward trend is observed. It is expected that the stability of the trend can be enhanced by allowing the simulation to run for longer (albeit potentially excessive) amounts of time. This is due to the notion that the simulation approach does not ensure that every available combination of securities will be evaluated, which would become increasingly more difficult as the number of shares added to the portfolio increases.

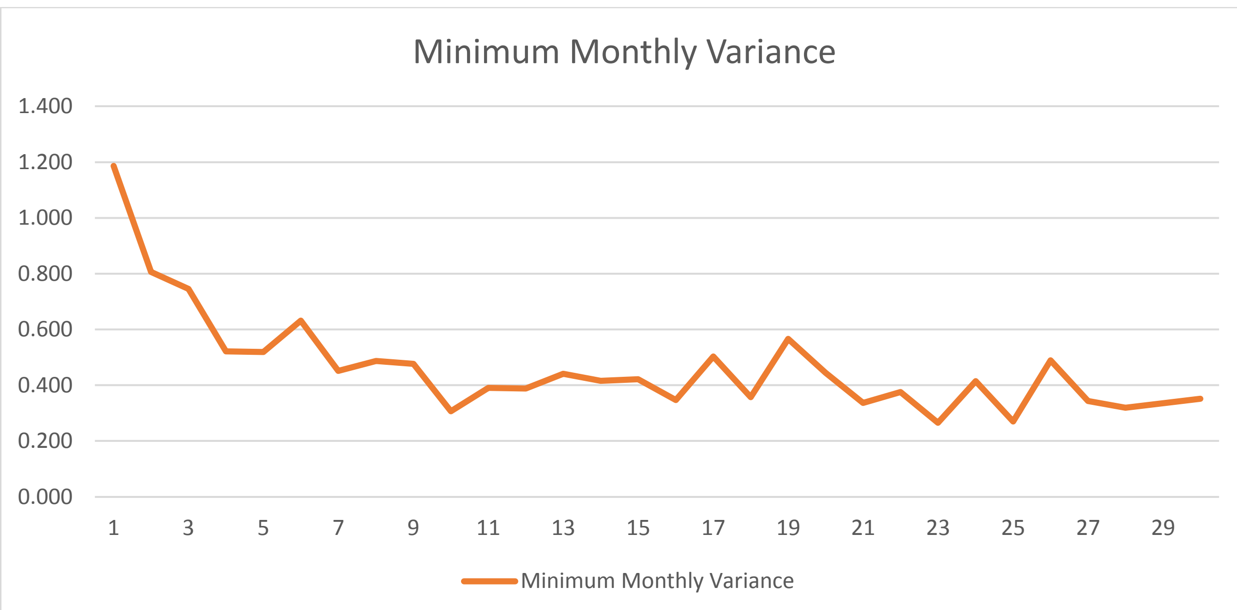


Figure 9: Graphical representation of the results achieved from the first simulation

In keeping with the methodology employed by previous literature, the monthly variance observations are annualised. The trend of the annualised variance is demonstrated below:

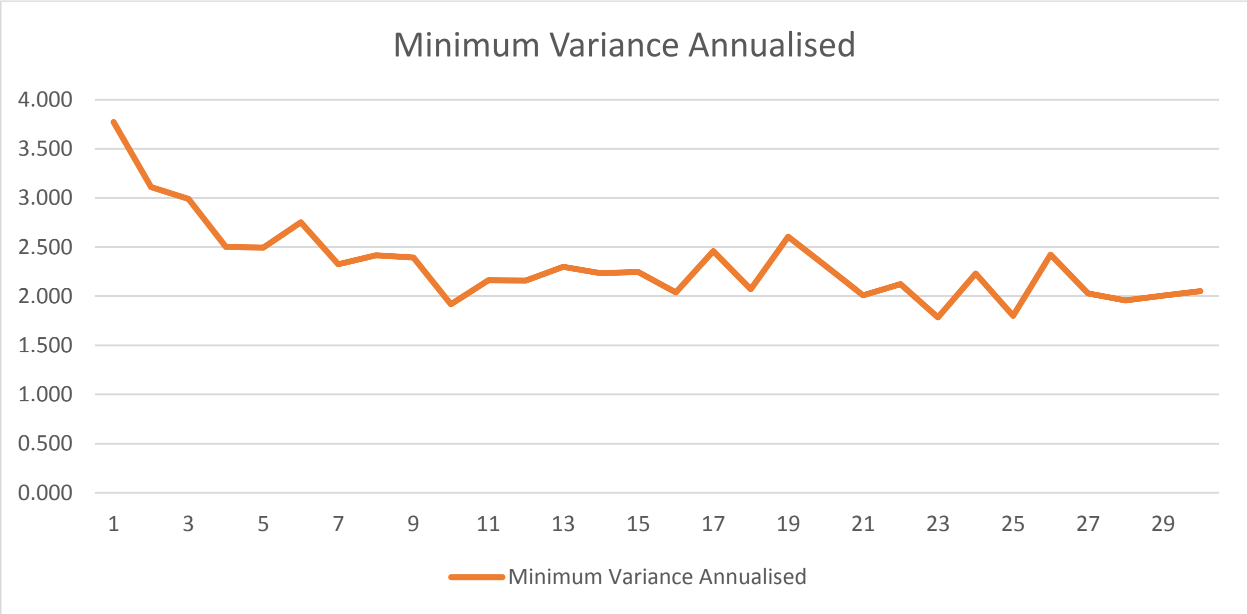


Figure 10: Annualised variance trend of the first test – traditional simulated approach with equal weightings.

In the minimum monthly variance graph it can be seen that optimal portfolio consists of merely 10 shares as after this point the additional benefit of diversification is negligible and in some instances in fact decreases. In the minimum annualised variance graph the point of maximal benefit of diversification indeed also lies at approximately 10 shares. Thereafter, one can further analyse each simulation. Due to the above, the ninth simulation is of particular interest – in this simulation the portfolio was restricted to consist of 10 shares.

In this simulation a total of 4 762 runs were completed. The first portfolio that was generated that satisfied all constraints was discovered on the 2nd run at a time of 46 seconds. This gave the minimum variance portfolio to be equal to 9.282. The minimal portfolio variance was found at the run time of 7 minutes 49 seconds, on trial number 3 652. The progress graph of how the portfolio variance was progressively minimised in this simulation can be observed in Figure 11. The vertical axis on the graph below refers to the portfolio variance achieved while the horizontal axis refers to the simulation number. The graph thus shows that the first successful run achieved a portfolio variance of 9.282. Within 58 seconds, the 84th run achieved a minimum portfolio variance of 2.07. This shows an improvement of approximately 348%. The final best simulation, which was the 3

652th simulation achieved at a time of 7 min 49 seconds, solved for a minimum portfolio variance of 0.307. This is a further improvement of 1.763 or a further 574%.

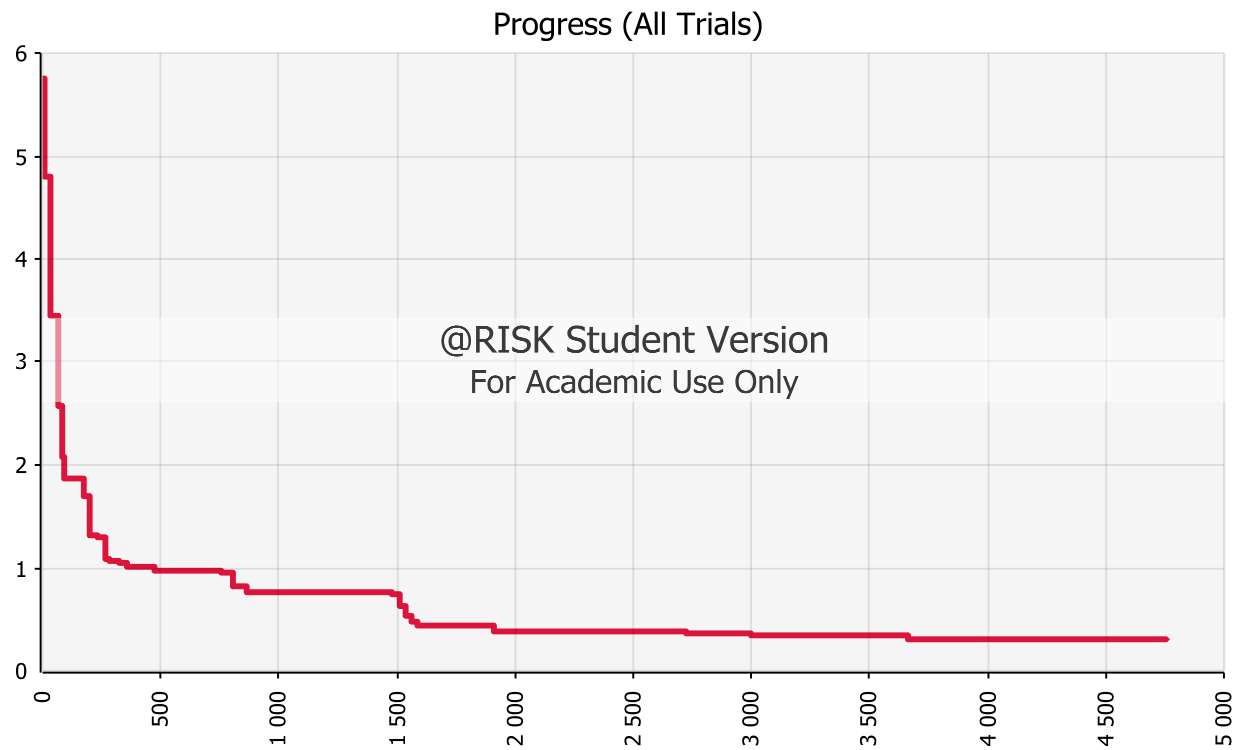


Figure 11: Progression of simulation when the portfolio is restricted to consist of 10 shares

The summary table is shown below:

Table 5: *Summary table of the Simulation displayed graphically in Figure 11*

Results	
Valid Trials	4761
Total Trials	4762
Best Value Found	0,307131749
+ soft constraint penalties	0,00
= result	0,307131749
Best Trial Number	3652
Time to Find Best Value	0:07:49
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Table 5 provides a summary of the full simulation. Here it is shown that 4 762 trials were run, 4 761 of which were valid (99.98%). The best value (minimal portfolio variance) found was 0.307 as demonstrated above. Furthermore the total optimisation time allowed was for 10 minutes. These results demonstrate the extent that the pre-specified, equally weighted traditional simulation approach can optimise the diversification benefits available to an investor within a relatively short amount of time. The extent to which the level of diversification achieved is significant is discussed by application of comparative analytics, expounded upon in the sections to follow.

An additional assessment is carried out in order to test the variability of the method's results. In this case the same simulation is run 5 times in order to further analyse how the outcome changes each time. This is applied to the 5, 15 and 25 share portfolio. The results are as below:

An equally weighted portfolio of 5 shares, with the simulation being run 5 times, yielded the following results:

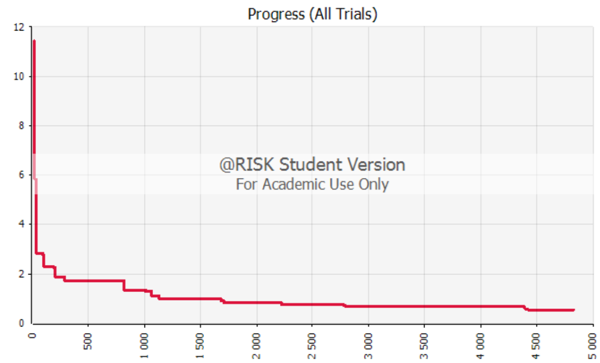
Table 6: *Variability of Variance Test when the Portfolio is restricted to consist of 5 Shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Monthly Variance	0.519	0.452	0.548	0.454	0.452
Annualised Variance	2.495	2.328	2.563	2.333	2.328
Total number of runs	4878	4832	4831	3891	4951
Run when best found	1704	2168	4416	1912	4547
Time when best found	3:57	5:00	9:17	4:56	9:18

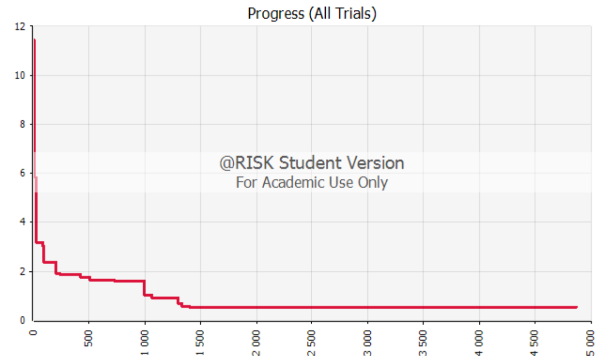
Table 6 shows that the monthly variance over 5 runs ranges between 0.452 – 0.519 which equates to a maximum deviation of 0.067 or a variance of 0.002061. The total number of runs completed averaged 4 676 with 4 out of the 5 simulations falling within 120 runs from each other. Thus far this points towards a relatively stable method. This is somewhat contrasted when observing the run when the best observation was found as well as the time when it was found. The best run fell between 1 704 (this was within the first 35% of observations) and 4 547 (this was within the last 9% of observations). The quickest time where the optimal solution was found was at 3 minutes 57 seconds while the longest, taking almost the full allocated time, occurred at 9 minutes 18 seconds. The 3rd and 5th simulations utilise significantly more time (almost double) to solve for the optimal solutions compared to the other 3 simulations. This was primarily due to a relatively small improvement being made to the portfolio near the end of the simulation – this can be observed through the analysis of the progress across each of the above simulations.

The graphs for each of the 5 simulations is shown in the figure below:

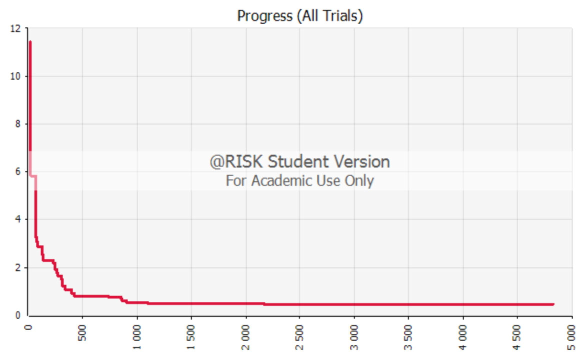
Simulation 1



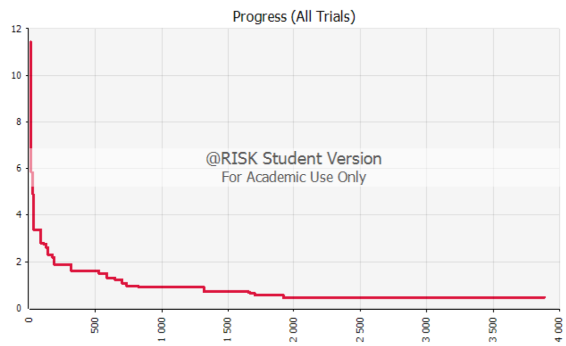
Simulation 2



Simulation 3



Simulation 4



Simulation 5

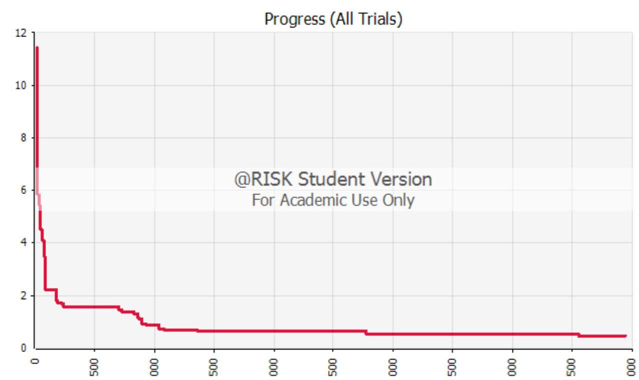


Figure 12: Graphs illustrating the variance minimisation progress achieved over the above simulations

The same exercise aiming to test the stability of the series was completed when the portfolio was given the constraint to consist of 15 shares, the results are displayed below:

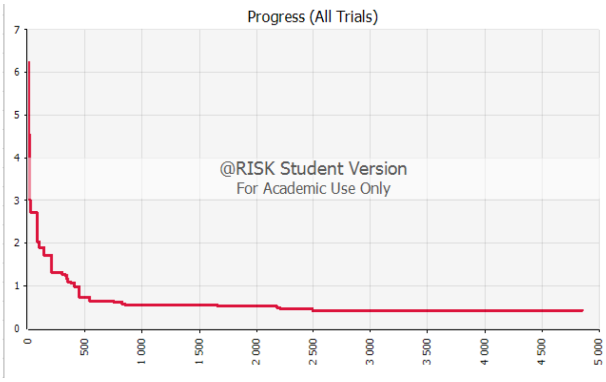
Table 7: *Variability of Variance test when the Portfolio is restricted to consist of 15 shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Monthly Variance	0.421	0.275	0.395	0.420	0.362
Annualised Variance	2.248	1.817	2.177	2.245	2.084
Total number of runs	4859	4851	4883	4863	4569
Run when best found	2483	4700	3799	4472	4455
Time when best found	5:30	9:44	7:54	9:19	9:47

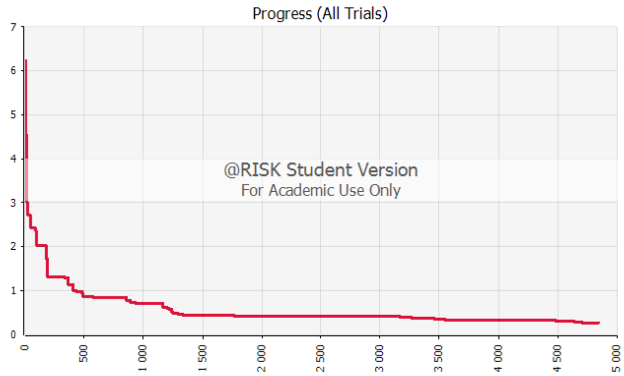
Here it is demonstrated that the monthly variance ranges between 0.275 and 0.421; this equates to a maximum spread of only 0.146 which equates to a variance of 0.00368. The average total number of runs completed was equal to 4 805. Four of the five simulations fell under 78 cumulative total runs of the average. Once again this points towards a stable methodology. The run and time when the best simulation was found again shows slightly more variability. The run when the best result was found ranged between 2 483 (this represents completing 51% of the total number of runs completed) and 4 700 (which represented 96.8% of the simulation being completed). The fastest time in which the best solution was found was at 5:30 minutes with the longest being at 9:44 minutes.

The figure depicting each of the 5 simulations is observed below:

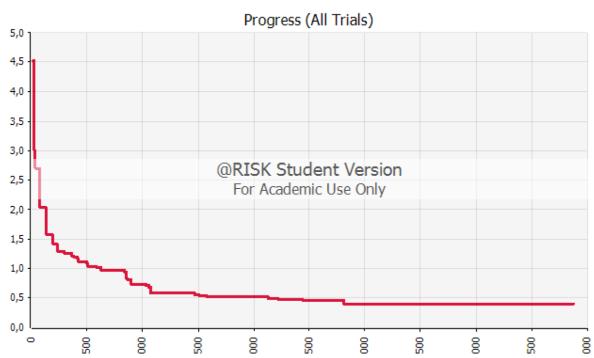
Simulation 1



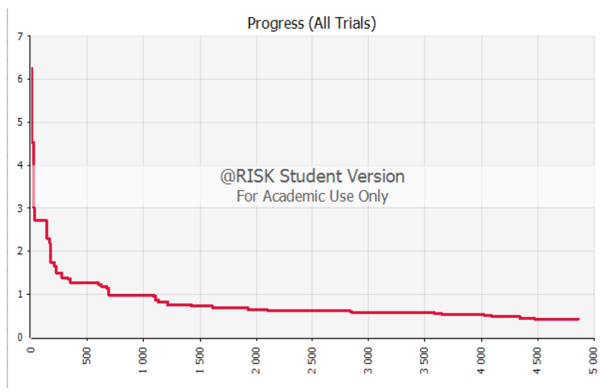
Simulation 2



Simulation 3



Simulation 4



Simulation 5

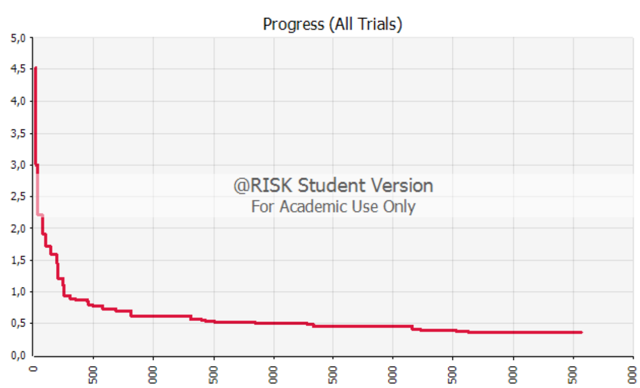


Figure 13: Graphs illustrating the variance minimisation progress achieved over the above simulations

This stability test was then completed with the portfolio consisting of 25 shares. The results can be seen below:

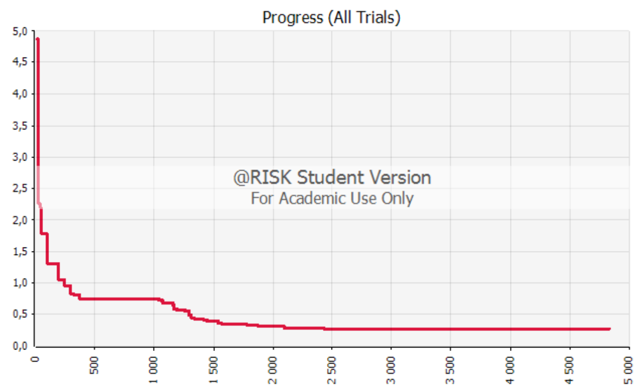
Table 8: *Variability of Variance test when the Portfolio is restricted to consist of 25 Shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Variance	0.270	0.368	0.326	0.352	0.342
Annualised	1.801	2.102	1.979	2.054	2.026
Total number of runs	4837	4751	4730	4744	4445
Run when best found	2484	3884	4299	4159	4087
Time when best found	5:31	8:16	9:03	8:46	9:11

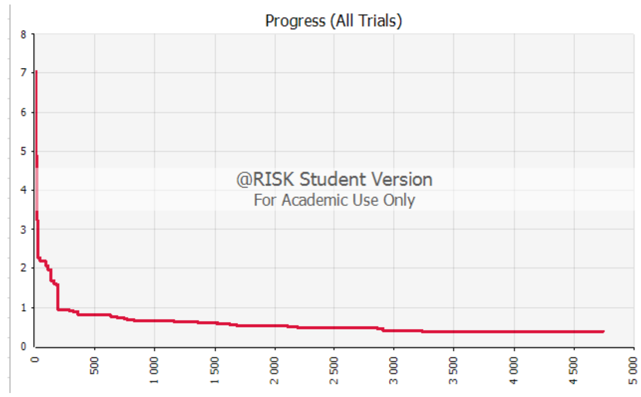
The variance of the portfolio ranged between 0.27 and 0.368 showing a range of 0.098. This equates to a variance of variance of 0.001413 over the 5 simulations. The average of the number of total runs completed was 4 701.4 with 3 out of the 5 simulations being completed within 50 runs of the average. The best solution was found between 2 484 runs (51% of the simulation) and 4 299 runs (with 90% of the simulation being completed). The time to find the best solution ranged between 5:31 minutes and 9:03 minutes.

The graphs of each of the 5 simulations is shown in the figure below:

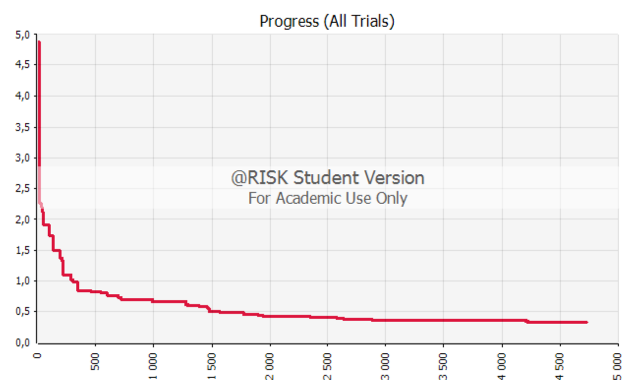
Simulation 1



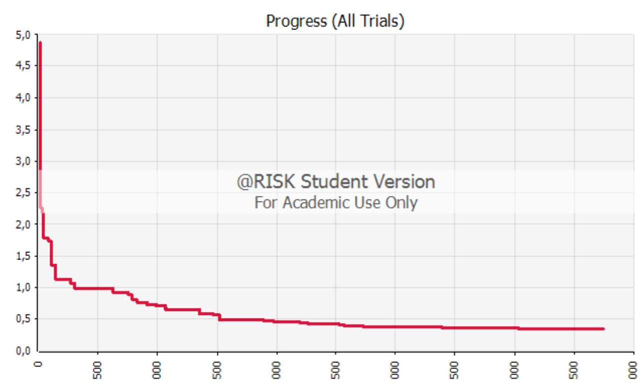
Simulation 2



Simulation 3



Simulation 4



Simulation 5

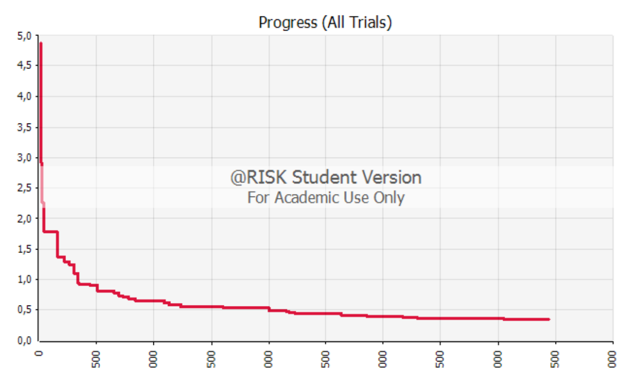


Figure 14: Graphs illustrating the variance minimisation progress achieved over the above simulations

4.1.1b Unequally Weighted Portfolio

The simulation of the portfolio's that allowed the weights on each share to vary, while maintaining the summation of the weights equal to one, was then generated. The table of the simulated results are shown in Appendix 2 while the graphical representation can be seen in **Figure 1**Figure 15 (below). This graph shows that the series has become further unstable as the trend of the line does not show evidence of slowly and smoothly decreasing but is rather relatively violent in its moves. This can be attributed to the increased quantum on the number of potential possibilities the portfolio may consist of. Nevertheless the graph still displays that the optimal number of shares to have included into the minimal variance portfolio lies between 9 and 15 shares.

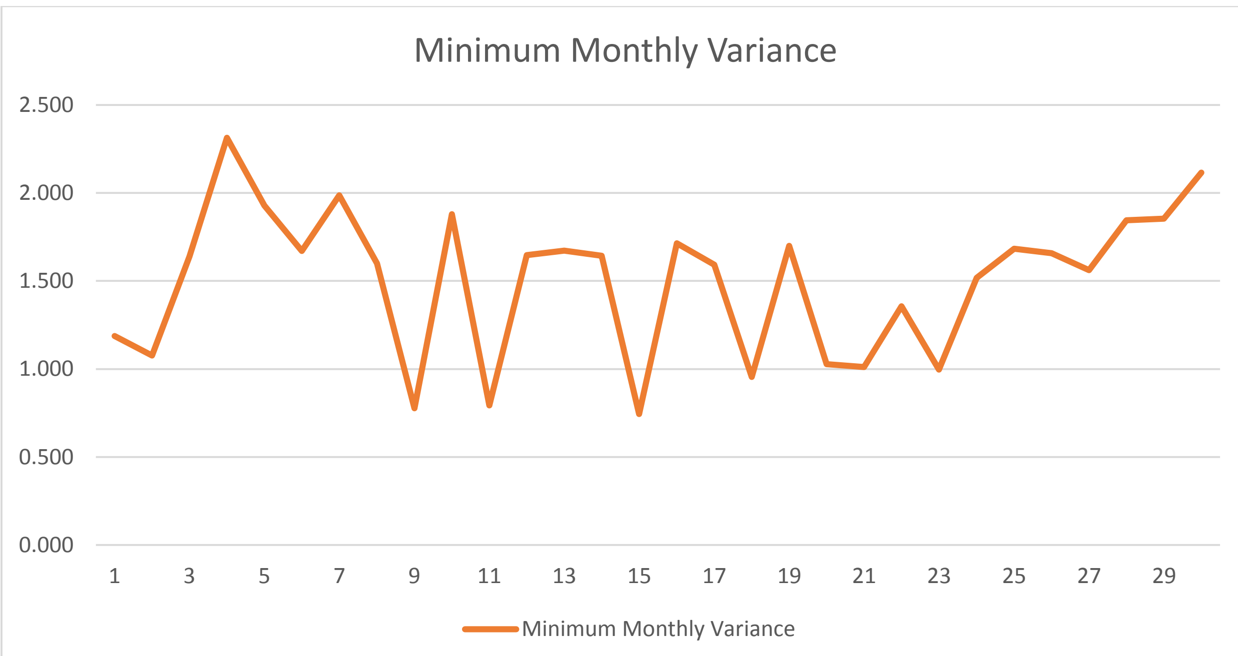


Figure 15: Graphical representation of the results achieved through the traditional simulation approach with variable weights

The annualised variance graph is displayed below. This provides evidence reinforcing the views obtained by utilising the monthly graph and table. Namely, the optimal number of shares lies between 9 and 15, albeit in an unstable series. The detailed results of each trial can be individually analysed. Of particular interest is the 15th observation in this series.

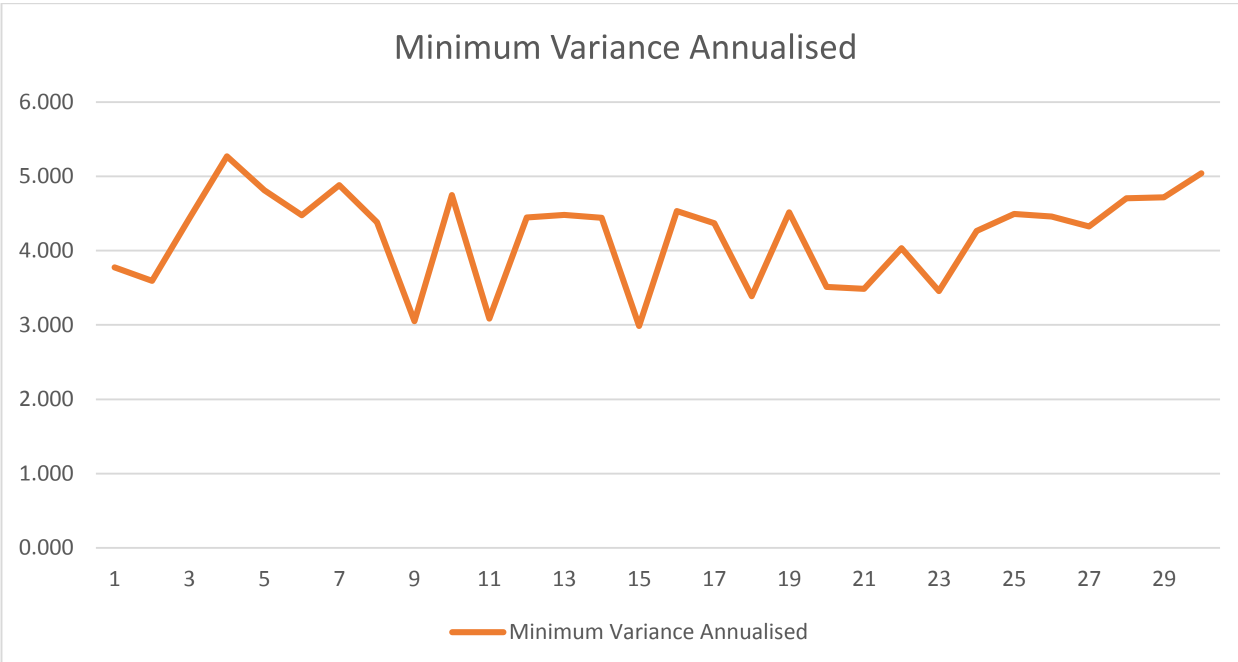


Figure 16: Adjusts the data in Figure 15 to be annualised

In total, 4 484 simulations were run in this test. The first valid scenario where all constraints were satisfied occurred at trial number 553 and at a time of 1 minute and 42 seconds. This achieved a minimum portfolio variance of 0.832. Although the number of trials was larger and the time to the first successful trial was significantly longer in this set of simulations, the first result found was also significantly lower than the result on the initial equally weighted portfolio. The minimal variance portfolio that was solved for when the portfolio was limited to select 15 shares was found at a time of 8 minutes 55 seconds. This was at trial number 3 983 and the minimal portfolio variance solved for was 0.744. The progression of the optimisation is displayed graphically below:

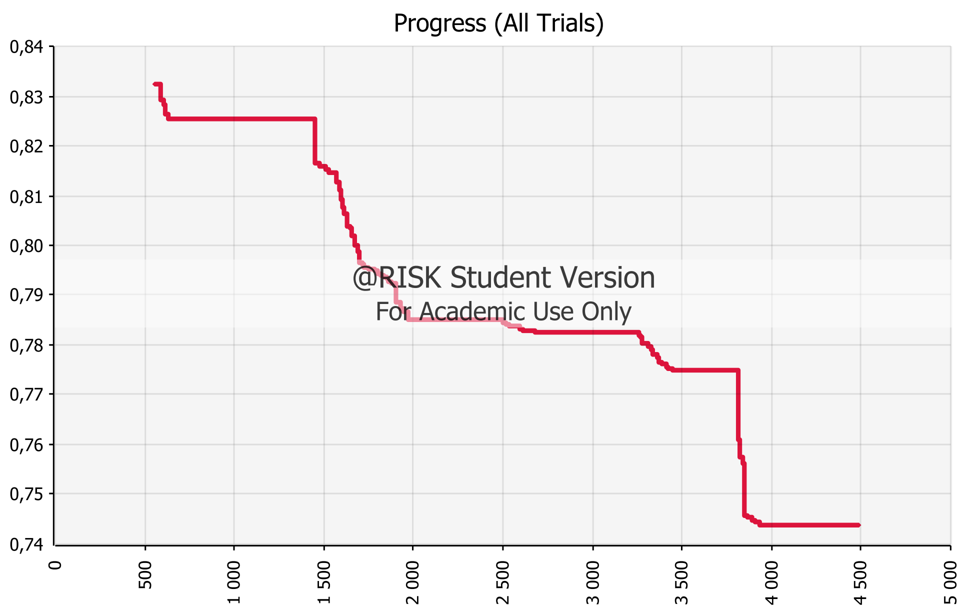


Figure 17: Progression of simulation when the portfolio is restricted to consist of 15 shares

The table summarising the simulation is presented below:

Table 9: *Summary table of the simulation displayed graphically in Figure 17*

Results	
Valid Trials	865
Total Trials	4484
Best Value Found	0,743779012
+ soft constraint penalties	0,00
= result	0,743779012
Best Trial Number	3983
Time to Find Best Value	0:08:55
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Here it can be seen that the valid trials were far fewer than in the equal weighted simulation, just 19.29% of all trials ran satisfied all constraints. The minimum portfolio variance value found was 0.744. Once again, the simulation was halted after 10 minutes.

Section 4.1.1 commenced by outlining the results achieved by the pre-specified equally weighted simulation approach. It was demonstrated that the optimally diversified portfolio consisted of 10 shares, exhibiting a variance of 0.307. The section followed with the analysis of the results achieved the pre-specified unequally weighted simulation approach. The optimally diversified portfolio using this approach employed 15 shares, exhibiting a variance of 0.744 and a less stable series. Instances where shares were added to the portfolio and a higher overall variance was obtained are assumed to be primarily due to permitting variable weightings on each share. This resulted in an increased quantum of potential possible portfolios leading to a smaller possibility of the globally optimally diversified portfolio being discovered, as described previously. These results indicate an improvement to the previous literature of Neu-Ner and Firer (1997) and that of Statman (1987) whom ascerted that an optimal portoflio should consist of at least 30 shares. The chapter continues with the analysis focused on the results achieved by the genetic programming approach when the number of shares a portfolio should consist of was pre-specified, as previously explained.

4.1.2 Genetic Programming Results (Pre-Selected Number of Shares)

4.1.2a Equally Weighted Portfolio

The Genetic Programming approach was thereafter applied to the data to solve for an equally weighted portfolio for each number of shares to be included respectively. A table of the results can be found in Appendix 2 and are displayed graphically in Figure 18 below. From the graph it can be seen that the series is far more stable than the simulation tests. It can therefore by hypothesised that the use of genetic algorithms allowed the analysis to explore more of the search space, giving it a higher probability of finding a global minimum variance portfolio for each portfolio consisting of increasing number of shares. Once again one can visually observe that the majority of the benefit of diversification can be obtained by holding between 10 and 15 shares. This implies that an investor will diversify his portfolio to include between 10 and 15 shares and thereafter will face ever decreasing diversification advantages to have an additional share included to their portfolio. This result is slightly different to the finding previously which displayed the majority of the diversification benefit can be obtained by holding a portfolio of between 9 to 15 shares. The unstable nature of the series inherent in the previous simulation is believed to be the underlying reason for the minor difference in the result.

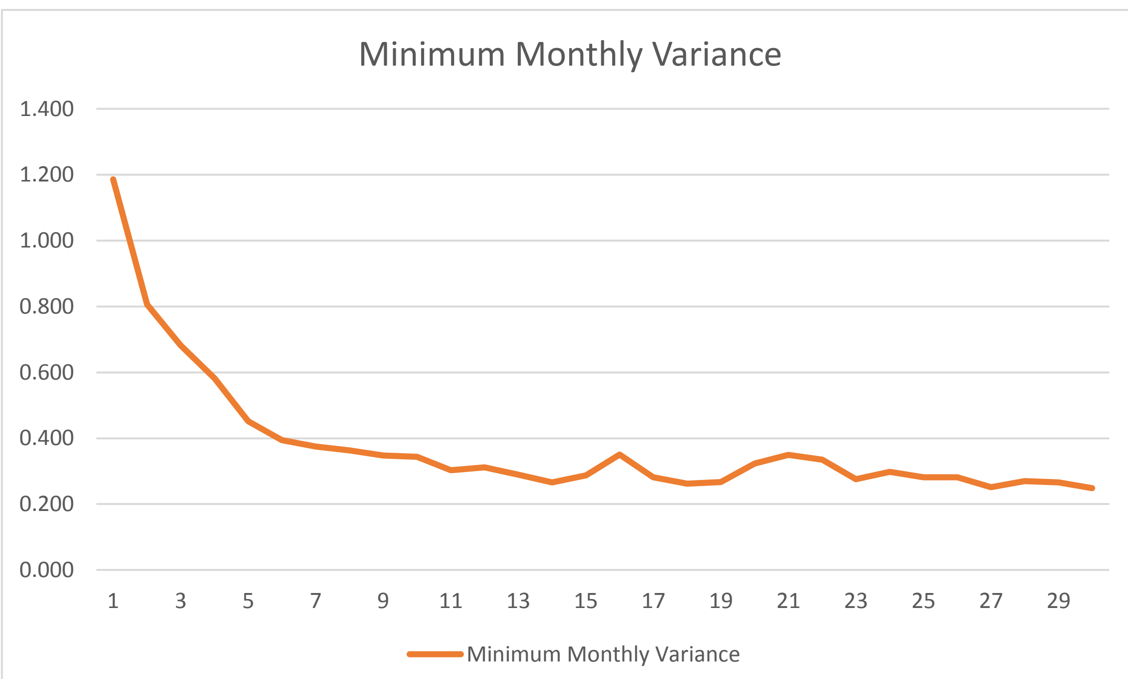


Figure 18: Graphical representation of the results achieved with a genetic programming simulation with equally weighted shares.

The results were once again annualised and the graph of the annualised series is shown below. The annualised minimum portfolio variance graph demonstrates equivalent results to the monthly minimum portfolio variance graph - the optimal diversification benefit can be obtained by including between 9 and 15 shares into a portfolio. As previously shown, each individual simulation can be analysed, here the 13th simulation will be further analysed (this equates to the portfolio that consists of 14 shares).

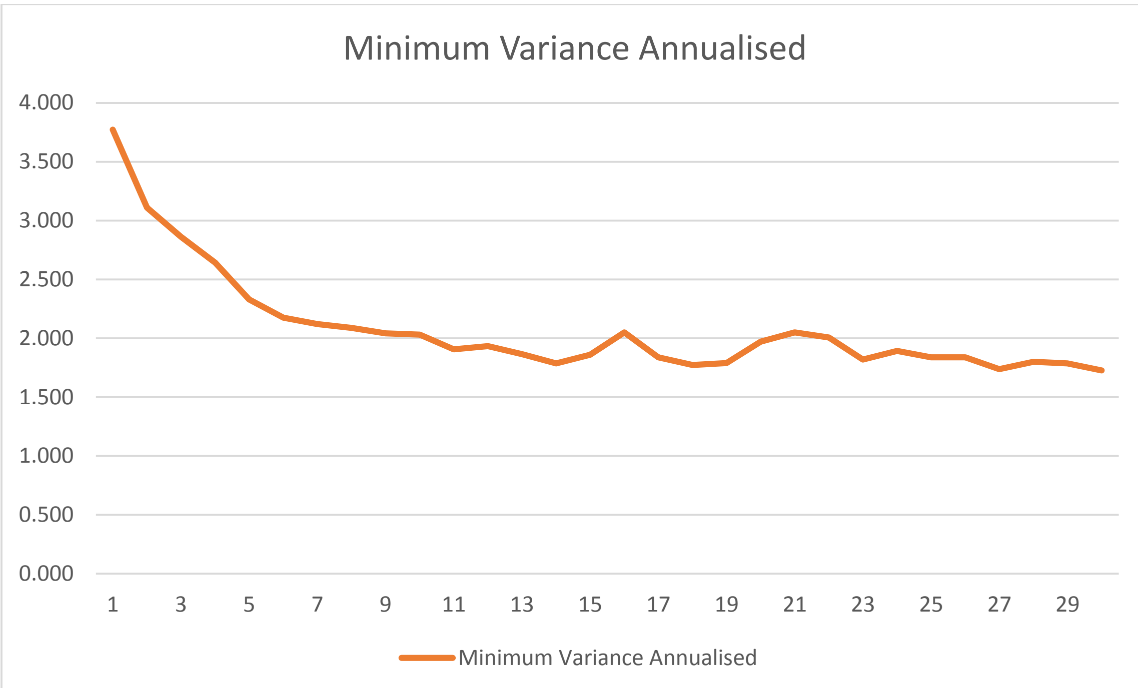


Figure 19: Adjusts the data in Figure 18 to be annualised

In this simulation a total of 11 696 runs were completed. The first portfolio that was generated that satisfied all constraints was discovered on the 2nd run at a time of 46 seconds. This gave the minimum variance portfolio to be equal to 7.152. The overall minimal portfolio variance was found at the run time of 8 minutes 5 seconds, on trial number 9 234. The progress graph of how the portfolio variance was progressively minimised in this simulation can be seen below:

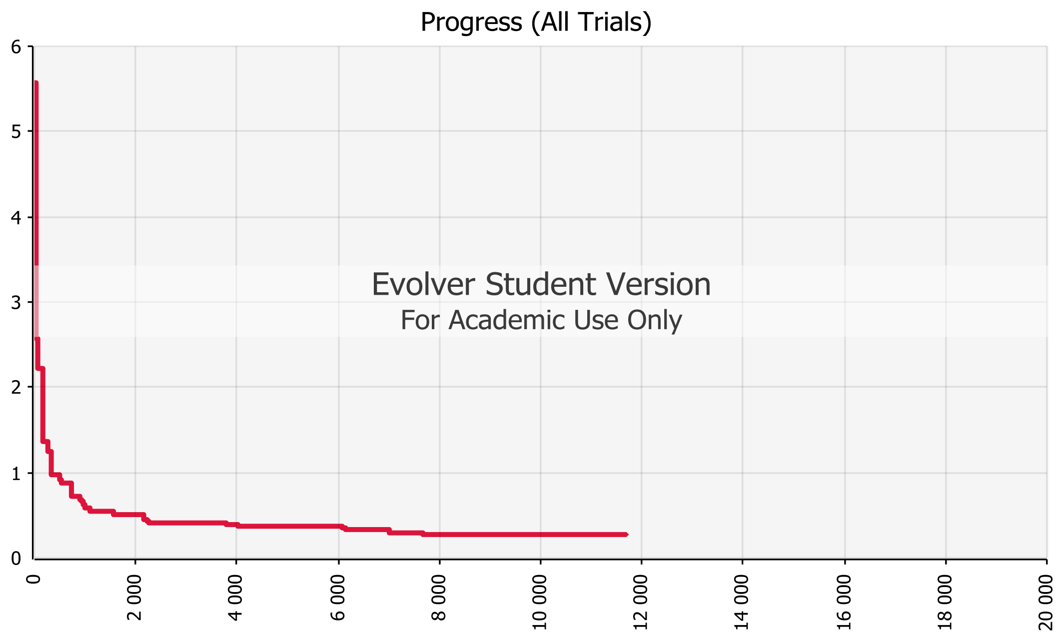


Figure 20: Progression of simulation when the portfolio is restricted to consist of 14 shares

Figure 20 shows that the first successful run achieved a portfolio variance of 7.152. Within a minute the best optimal solution was found to be 0.71 the 320th run. This shows an improvement of approximately 694%. The final best simulation, which was the 9 234th simulation achieved at a time of 8 min 5 seconds, solved for a minimum portfolio variance of 0.266. This is an improvement on the initial value of 6.182 or an improvement of 737%.

The summary table is shown below:

Table 10: *Summary table of the simulation displayed graphically in Figure 20*

Results	
Valid Trials	11695
Total Trials	11696
Best Value Found	0,26631385
+ soft constraint penalties	0,00
= result	0,26631385
Best Trial Number	9234
Time to Find Best Value	0:08:05
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Table 10 (above) provides a summary of the full simulation. Here it is shown that 11 696 trials were run, 11 695 of which were valid (99.9%). The best value (minimal portfolio variance) found was 0.266 as demonstrated above. Furthermore the total optimisation time allowed was for 10 minutes. These results demonstrate the degree to which a pre-specified, equally weighted genetic programming approach can provide an investor with enhanced diversification benefit within a relatively short amount of time. This is further explored in section 4.4 which directly compares the results achieved by each approach.

An additional test was then completed in order to calculate the variability of the results. Five simulations were run in turn on each of the portfolios constructed of 5, 15 and 25 shares respectively in an attempt to view the stability of the output. The results are demonstrated below.

The results from the below table show the simulations run when the portfolio was restricted to consist of 5 shares.

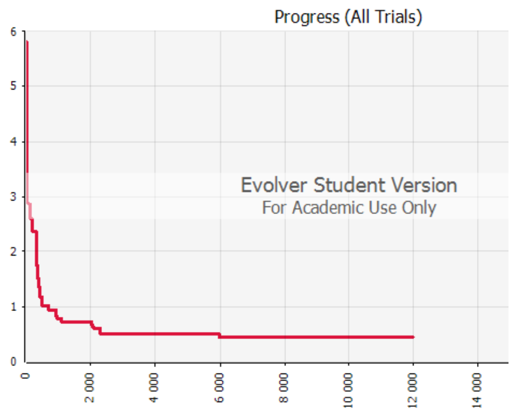
Table 11: *Variability of Variance test when the portfolio is restricted to consist of 5 Shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Monthly Variance	0.452	0.452	0.493	0.452	0.452
Annualised Variance	2.328	2.328	2.431	2.328	2.328
Total number of runs	11980	10129	10054	10254	10048
Run when best found	8894	3658	5122	2167	8799
Time when best found	7:42	4:00	5:19	2:29	8:52

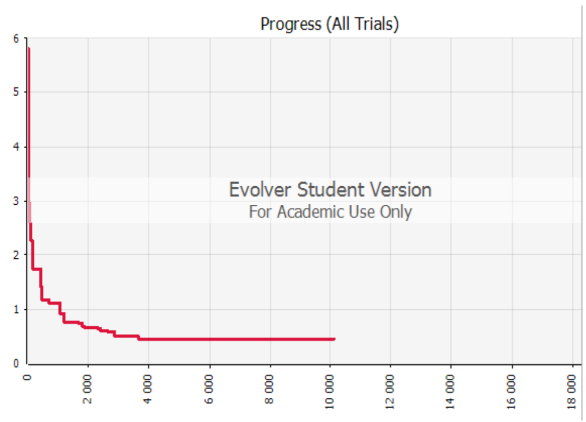
The variance of the portfolio ranged between 0.452 and 0.493 showing a range of 0.041. This equates to a variance of variance of just 0.000333 over the 5 simulations. The stability of the series is shown to be strong as 4 of the 5 simulations yielded the equivalent monthly variance. The average of the number of total runs completed was 10.493. This was more variable with the simulations ranging between 239 runs and 1 487 runs away from the average. The best solution was found between 2 167 runs (21.1% of the simulation) and 8 894 runs (with 74.24% of the simulation being run). The time to find the best solution ranged between 2:29 minutes and 8:52 minutes.

The graphs of each of the 5 simulations is shown in the figure below:

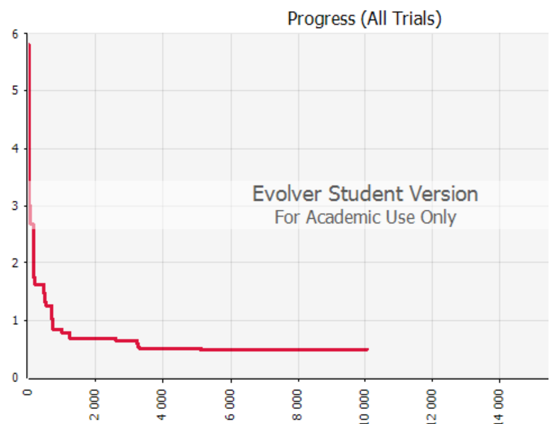
Simulation 1



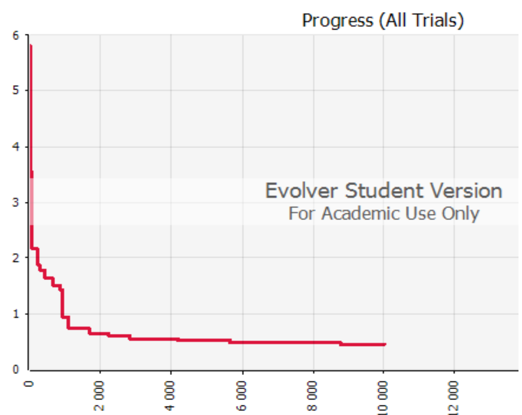
Simulation 2



Simulation 3



Simulation 4



Simulation 5

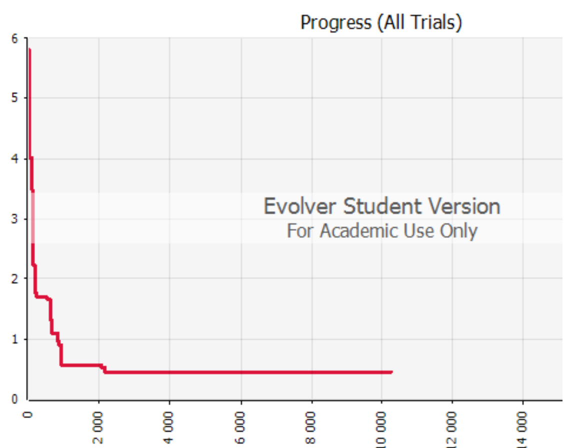


Figure 21: Graphs illustrating the variance minimisation progress achieved over the above simulations

This test was then applied to portfolios that were restricted to consist of only 15 shares, the results of which are show in the table below:

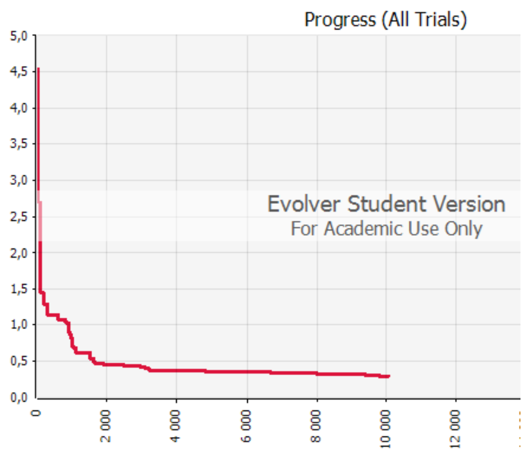
Table 12: *Variability of Variance test when the Portfolio is restricted to consist of 15 Shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Monthly Variance	0.288	0.275	0.299	0.283	0.276
Annualised Variance	1.859	1.818	1.896	1.842	1.820
Total number of runs	10054	10893	11538	11407	11248
Run when best found	9804	5321	6336	9266	6559
Time when best found	9:49	5:12	5:44	8:14	6:11

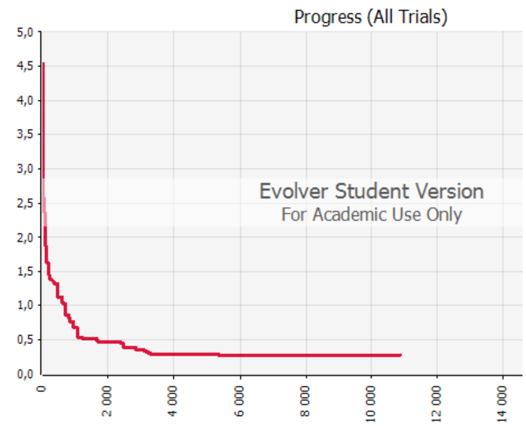
The variance of the portfolio ranged between 0.275 and 0.288 showing a range of 0.003. This equates to a variance of variance of 0.0000992 (approximately zero) over the 5 simulations. Although this points towards extreme stability, the average of the number of total runs completed was 10 493. This illustrates that the simulations varied in the total number of runs completed by between 974 and 135. The best solution was found between 5 321 runs (48.85% of the simulation) and 9 804 runs (with 97.51% of the simulation being completed). The time to find the best solution ranged between 5:12 minutes and 9:49 minutes.

The figure displaying the graphs of each of the 5 simulations is shown below:

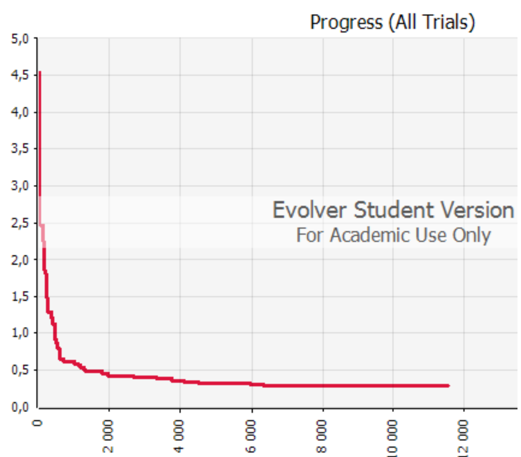
Simulation 1:



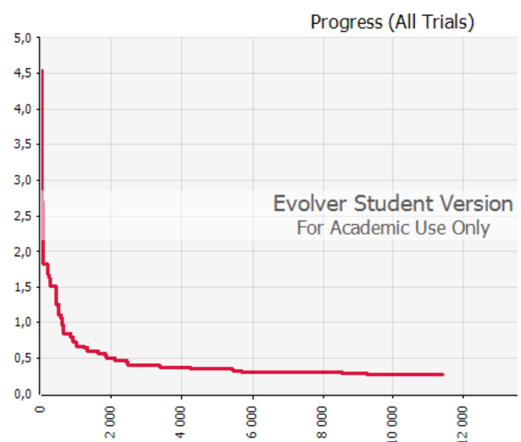
Simulation 2



Simulation 3:



Simulation 4



Simulation 5:

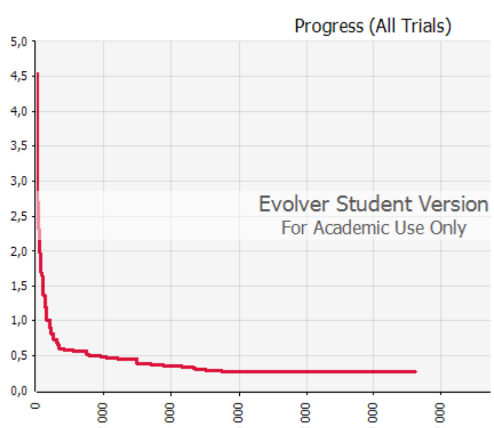


Figure 22: Graphs illustrating the variance minimisation progress achieved over the above simulations

This stability test was then run once again, this time on the portfolio restricted to consist of 25 shares. The results of this test are shown in the table below:

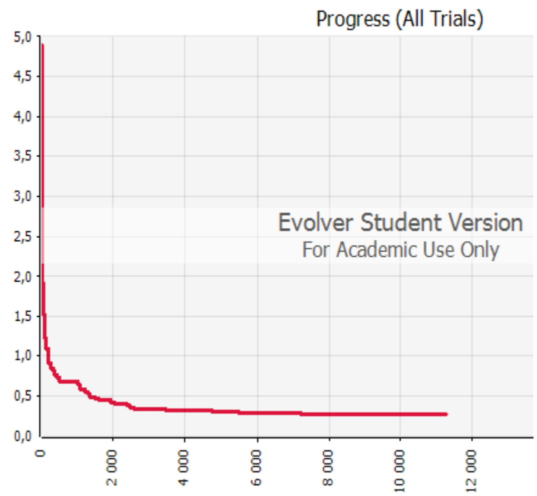
Table 13: *Variability of Variance test when the Portfolio is restricted to consist of 25 Shares*

Simulation	**1**	**2**	**3**	**4**	**5**
Monthly Variance	0.282	0.353	0.234	0.236	0.257
Annualised Variance	1.839	2.058	1.677	1.684	1.758
Total number of runs	11291	10897	11544	11355	11455
Run when best found	8966	9420	10693	9717	8645
Time when best found	8:09	8:48	9:15	8:33	7:37

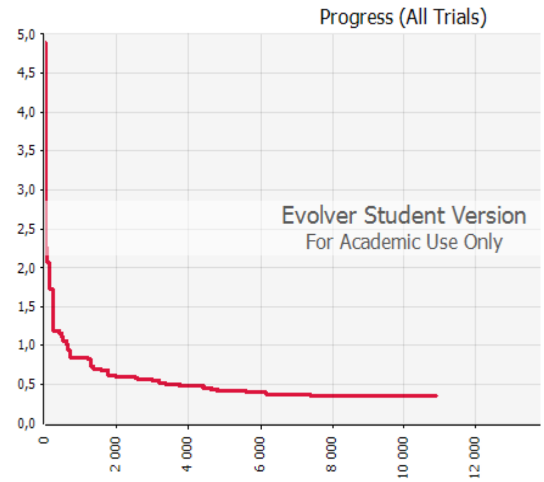
The variance of the portfolio ranged between 0.234 and 0.353 showing a range of 0.119. This equates to a variance of 0,00239476 (approximately zero) over the 5 simulations. The average total number of runs completed was equal to 11308.4. Therefore a range in the total number of runs completed was between 411.4 and just 17.4 runs away from the average. The best solution was found between 8,645 runs (75.47% of the simulation) and 10 693 runs (with 92.63% of the simulation being completed). The time to find the best solution ranged between 7:37 minutes and 9:15 minutes.

The graphs of each of the 5 simulations are shown in the figure below:

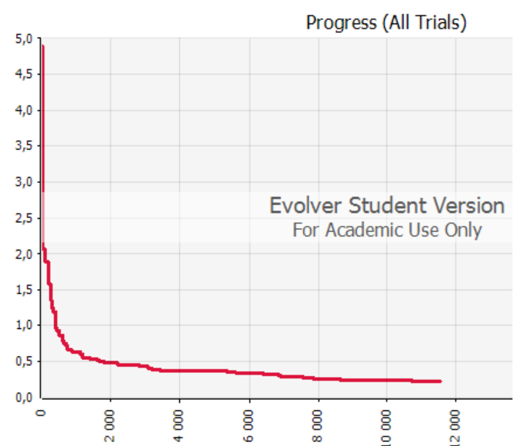
Simulation 1:



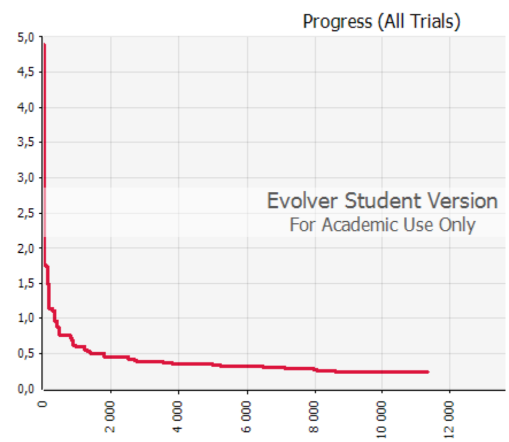
Simulation 2:



Simulation 3:



Simulation 4:



Simulation 5:

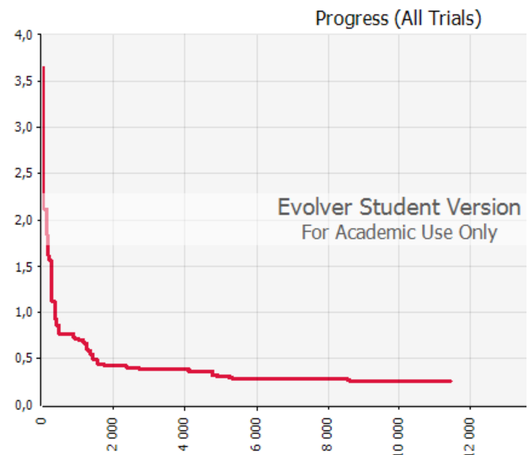


Figure 23: Graphs illustrating the variance minimisation progress achieved over the above simulations

4.1.2b Unequally Weighted Portfolio

The methodology described above was then applied under the constraints where the weights on each share were allowed to vary, however the weights of all the shares to be included in the portfolio is required to sum to 1. Applying the genetic programming approach yielded table of results shown in Appendix 2. These results are displayed graphically below. As in the simulation approach with unequal weightings, the results are far more erratic compared to the equally weighted counterpart due to the increasing possibilities of portfolio construction, only a fraction of which can be explored in the ten minute time restriction. However similar findings hold, the minimum variance portfolio can be constructed by using between 10 and 15 shares.

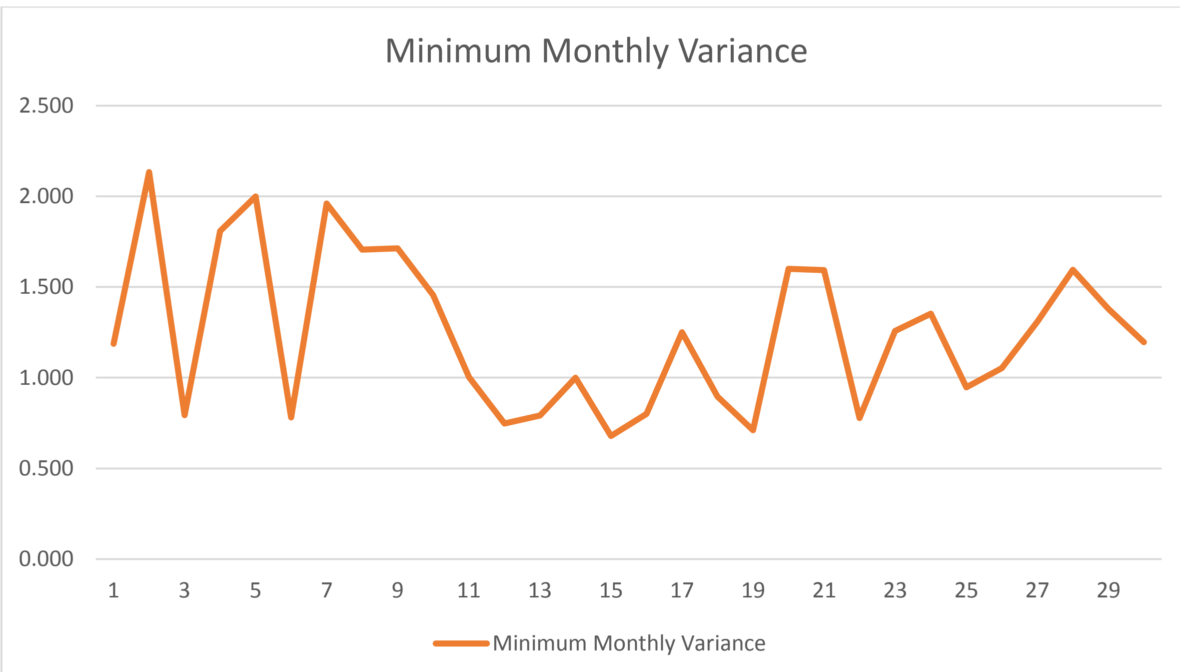


Figure 24: Graphical representation of the results achieved utilising a genetic programming approach while allowing the weightings on each share in the portfolio to vary

The annualised series of the above is graphically displayed below:

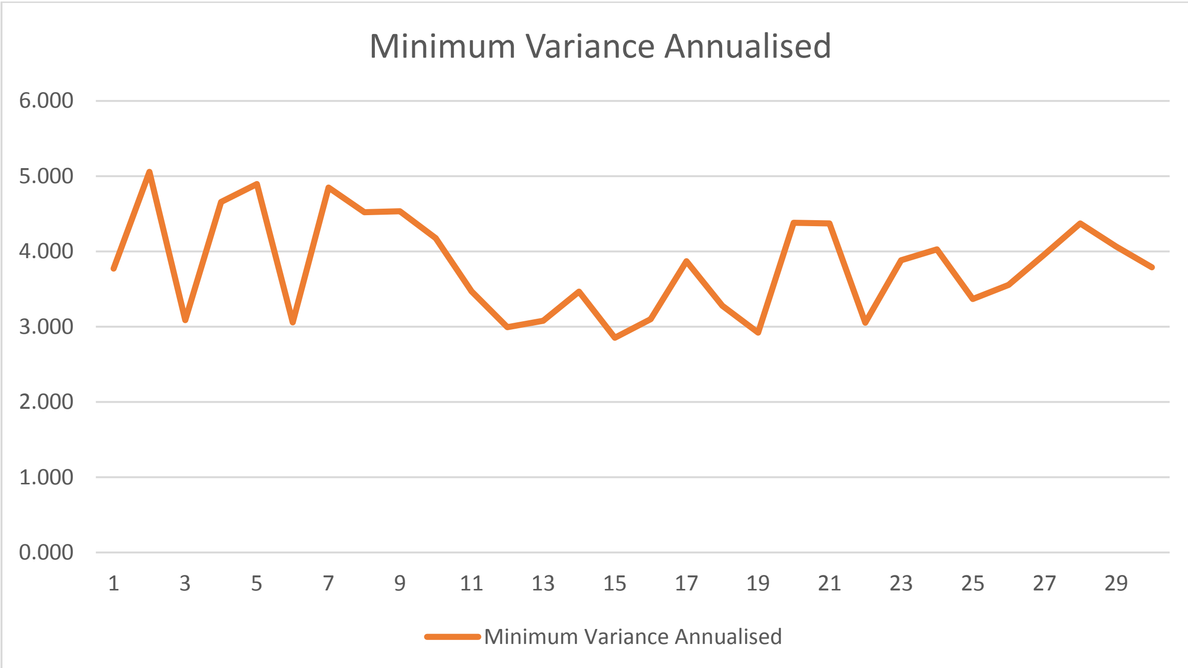


Figure 25: Adjusts the data in Figure 24 to be annualised

Once again it is demonstrated that despite using a relatively more advanced approach, the resulting series remains erratic. Obtaining an unstable series across both the traditional simulation approach as well as the genetic programming approach provides a practical demonstration illustrating the extent to which the number of potential portfolios is increased when the weightings on each share are permitted to vary. Figure 25 however still displays that optimal results are achieved with a portfolio of between 10 to 15 shares. As previously shown, each individual simulation can be analysed, here the 14th simulation will be further analysed (this equates to the portfolio that consists of 15 shares).

In this simulation a total of 9 931 runs were completed. The first portfolio that was generated that satisfied all constraints was discovered on the 502nd run at a time of 55 seconds. This gave the minimum variance portfolio to be equal to 1.889. The overall minimal portfolio variance was found at the run time of 9 minutes 47 seconds, on trial number 9 657.

The progress graph of how the portfolio variance was progressively minimised in this simulation can be seen below:

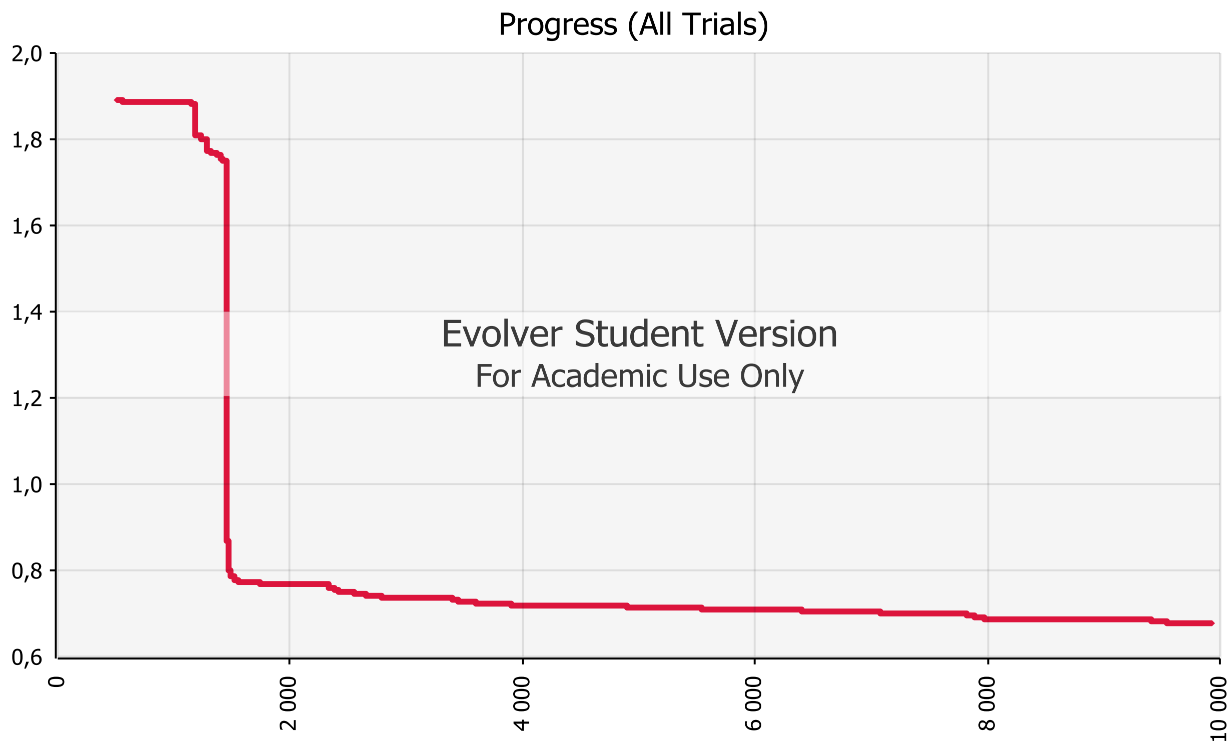


Figure 26: Progression of simulation when the portfolio is restricted to consist of 15 shares

Figure 26 shows that the first successful run achieved a portfolio variance of 1.889. The algorithm took 1:46 minutes to achieve a portfolio variance of below 1 – at the 1 438nd run a portfolio variance of 0.926 was achieved. This shows an improvement of approximately 204%. The final best simulation, which was the 9 657th simulation achieved at a time of 9 min 47 seconds, solved for a minimum portfolio variance of 0.679. This is an improvement on the initial value of 1.21 or an improvement of 278%.

The summary table is shown below:

Table 14: *Summary table of the simulation displayed graphically in Figure 26*

Results	
Valid Trials	2146
Total Trials	9931
Best Value Found	0,679013707
+ soft constraint penalties	0,00
= result	0,679013707
Best Trial Number	9657
Time to Find Best Value	0:09:47
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Table 14 provides a summary of the full simulation. Here it is shown that 9 931 trials were run, 2 146 of which were valid (21.1%). The best value (minimal portfolio variance) found was 0.679 as demonstrated above. Furthermore the total optimisation time allowed was for 10 minutes.

Chapter 4 opened in section 4.1.1 by expounding upon the results achieved by the pre-specified equally weighted simulation approach. The chapter continued in section 4.1.2 with a similar inspection into the results achieved by the pre-specified equally weighted genetic programming approach. This technique obtained a less unstable series than that of the simulation approach. When equal weightings were employed the optimally diversified portfolio consisted of 14 shares and a variance of 0.266 was observed. When the weightings on each share were allowed to vary the approach solved for an optimally diversified portfolio containing 15 shares, exhibiting a variance of 0.307. These results demonstrate increased strength of the approach compared to the traditional simulation approach and to past literature. The comparable improvements are discussed further in section 4.4. The chapter now continues by examining the results achieved across mathematical approaches when an unconstrained approach was employed.

4.2 Unconstrained Analysis (Absent of Pre-Selected Number of Shares)

The next section aimed to solve the research question regarding how many shares the minimum variance portfolio would consist of, if the restriction that the portfolio should consist of a pre-specified number of shares was removed. The minimum variance in that optimal portfolio will be thereafter calculated. Firstly, the traditional simulation approach was applied to the data.

This approach used slightly different constraints that what were previously applied in the simulations above. The only constraint on the number of shares in the portfolio was that the portfolio needed to consist of at least 1 share. If this were not a constraint, the simulations would consistently solve for a portfolio consisting of 0 shares. Secondly, once the simulation had solved for the initial portfolio variance, the next best portfolio variance needed to be at least 0.01% better in order to replace the current best portfolio. In addition, if the portfolio variance is not bettered by 0.01% over 100 runs the simulation will be stopped and the current lowest portfolio variance utilised. This is in an attempt to restrict the simulation from continuously adding shares to the portfolio, each in turn with decreasing impact on lowering the variance of the portfolio. Initially a simulation approach was used on an equally weighted portfolio of shares and thereafter on an unequally weighted portfolio of shares. Secondly, the same test was run but employed a genetic programming approach.

4.2.1 Simulation Results

4.2.1a Equally Weighted Portfolio

When the simulation approach was made use of in the test above on an equally weighted portfolio, the minimum variance portfolio consisted of 37 shares and achieved a portfolio variance of 0.39542. Notably, in the simulation analysis previously, this portfolio variance was achieved using only 12 shares. Presumably the increased potential search space prevented the simulation analysis from solving for this portfolio construction. This is discussed in more detail in section 4.4.

In this simulation a total of 3 583 simulations were run. The minimum variance portfolio was solved for at a time of 9:57 minutes (simulation number 3 559) and achieved a monthly portfolio variance of 0.395 which equates to an annualised variance of 2.178. This portfolio solved for the minimum portfolio variance using 37 shares. The graph of the progression of the simulation is displayed below:

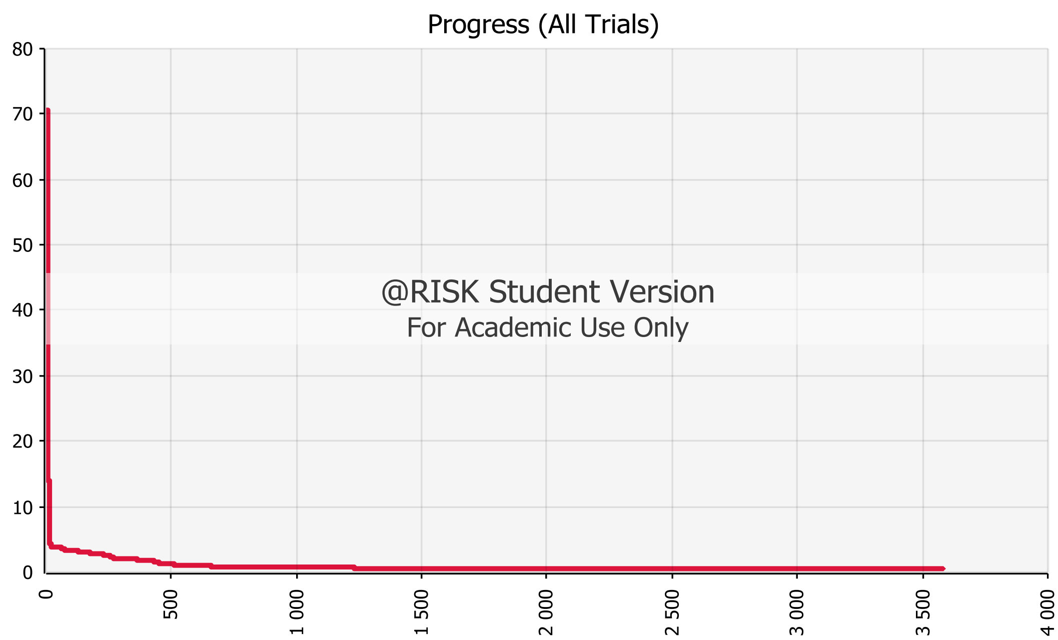


Figure 27: Shows the progression of the simulation when the number of shares to be included in the portfolio is not pre-specified, weightings on each share are equal

In Figure 27 one can visually determine that the initial portfolio achieved a monthly portfolio variance of just over 70 – the portfolio variance was 70.654. Within 500 simulations the minimum monthly portfolio variance was equal to 1.329. This is a significant improvement of 5 316%. A table of the overall performance of the simulation is displayed below:

Table 15: *Summary table of the simulation displayed graphically in Figure 27*

Results	
Valid Trials	3582
Total Trials	3583
Best Value Found	0,395420345
+ soft constraint penalties	0,00
= result	0,395420345
Best Trial Number	3559
Time to Find Best Value	0:09:57
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

This summary shows that 99.97% of the simulations run were valid. The minimum variance portfolio was equal to 0.3954 and was achieved at a time of 9:57 minutes.

Computation of the Benchmark Portfolio

The primary objective of this dissertation is to discover the optimally diversified portfolio through the use of the three mathematical approaches (simulation approach, genetic programming approach and particle swarm optimisation approach). Solving for the optimally diversified portfolio requires modelling the relationship between the number of shares in the portfolio and the risk (standard deviation) inherent in the portfolio. Thereafter portfolios are modelled to include various adjustments to their constraints, such as allowing the weights on each share in the portfolio to vary as opposed to be equally weighted. Throughout the analysis of this dissertation the same assumptions were applied as utilised throughout past literature, as described in the literature review section. This focus of study is significant as it identifies the fundamental relationship between the number of shares and the risk in a portfolio; this was the initial step which thereafter enables one to concentrate on the relationship between the risk and reward (return) of a portfolio.

Although not the primary focus of the dissertation, it is interesting to examine the return profiles of the portfolio's identified as optimally diversified in the various stages of analysis. The returns achieved by the optimally diversified portfolios should then be compared to the returns an investor would earn if he simply bought the market portfolio, referring to the portfolio containing every available share on the JSE Main Board, post the respective adjustments described in Chapter 3. In this way, the market portfolio acts as a benchmark. The analysis below is completed with the aim to state whether holding an optimally diversified portfolio allows the investor to earn excess, benchmark beating returns.

In line with the above, firstly the mean returns of the equally weighted market portfolio are calculated. This is completed by computing the average return of all the respective listed shares in each month over the observed time period. Secondly, the median of the market returns is computed for each month. The median represents the midpoint of the frequency distribution of observed values and as such represents the middle most value in a range of data such that there is an equal probability of falling above or below it. The median value is immune to outliers which would otherwise effect the mean.

The returns in the above analysis are calculated absent of transaction costs. The majority of brokers enabling an investor to transact (buy and sell shares) on the JSE main board issue either a fixed cost or a sliding fee dependant on the value of each completed transaction (Corbbett, 2013). Although there are other costs involved (such as a monthly account fee) the transaction costs will constitute the majority of the costs faced by an investor. If the fees faced by an investor are purely dependant on the value of the portfolio to be initiated, both the optimally diversified portfolio and the benchmark portfolio would experience equivalent transaction costs. Alternatively, if the fees faced by an investor are dependant on the number of transactions to be completed, representative of the number of shares in the portfolio, fees on the optimally diversified portfolio provide a lower bound to the amount of fees an investor could incur while the benchmark portfolio represents the upper bound (encompassing all available shares). Transaction costs are not incurred post initiation of the respective portfolios due to their buy-to-hold nature; shares are purchased on the first observed day and held until termination of the observed period. As the analysis below is comparative in nature it is not considered imperative to include an exact figure in the return calculation representative of the transaction fees. The aim of the analysis below is to indicate whether the optimal portfolios achieve benchmark beating returns as opposed to quantifying the exact post-cost outperformance.

The figure below compares the mean market returns against the median market returns as a ratio over time. This ratio should remain fairly constant; if it does not, it indicates there may be outliers in the population. This can be assumed as an outlier would drag the mean in the respective direction but would not affect the median. The graph depicting this ratio is displayed below:

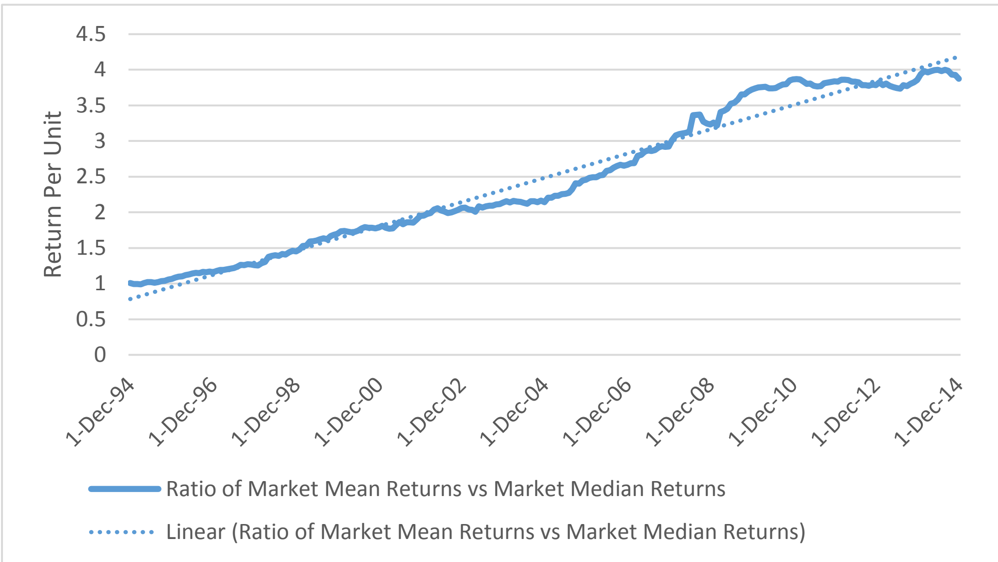


Figure 28: Ratio of Returns using a Mean Market Portfolio against the returns of the Median Market Portfolio

In Figure 28 the solid line displays the ratio over time of the mean market returns against the median market returns while the dotted line reflects the linear trend line thereof. As stated above, with no notable outliers the ratio of the mean return series to the median return series should be approximately constant. Figure 28 illustrates that this is not the case; the mean return over the time period observed consistently outperforms the median. A ratio of 2 in the above figure can be interpreted as the mean return of the market portfolio being double that of the median portfolio at that point in time. Over the observation period, the mean portfolio earns a return of approximately 3.9 times the median portfolio. This gives evidence that there may be notable outliers in the population. Additionally it can be observed that the aforementioned outliers occurred across the time series in the population - if there was a single outlier occurring in a single month, the ratio would adjust at that point in time but would thereafter remain at the new ratio constantly. As this is not the case it, as displayed by Figure 28, it can be stated that there is at least one share in the population in the majority of months observed whose return can be considered a relative outlier.

Notably, this is not indicative that it is the same share each month. Analysis of the underlying return distributions demonstrates that this is indeed the case – the share with an outlying return in each month over the time series differs from month to month. Due to this observation, it is prudent to make adjustments to the mean market portfolio, as to remove the outliers prevalent in the period or to reduce their weighting. This is completed through the use of two adjustments to the mean return portfolio.

The two adjustments made to the mean return series were completed in order to obtain a return series that is a better representation of the true obtainable market portfolio. This would be observable if the return series of the portfolio, post adjustments, better approximated the returns achieved by the median return portfolio. The adjustments made entailed removing the returns of the single best and single worst performing shares from the market portfolio each month over the time period observed. In a similar fashion to the above representation, the ratio of the adjusted mean market return portfolio to the median market return portfolio is shown in Figure 29. Here the solid line reflects the ratio of the mean market returns in the adjusted sample against the median market returns while the dotted line represents a linear trend line thereof.

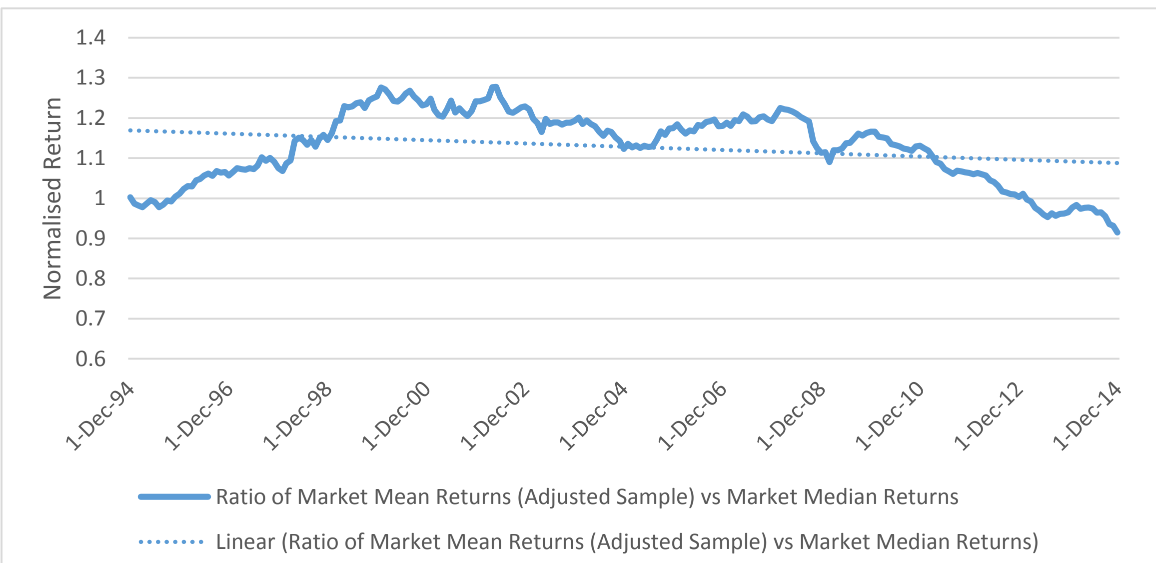


Figure 29: Ratio of Returns of the Adjusted Mean Market Portfolio against the returns of the Median Market Portfolio

Figure 29 demonstrates that the adjusted mean market portfolio is a better representative of the market return series, as observed by the ratio to the median portfolio. The ratio is more stable linear than the unadjusted return series. In addition, this is further demonstrated by the smaller range of values within the series. The range of the ratio in Figure 28 was between 0.99 and 3.99.

Figure 29 however demonstrated a narrower range, between 0.91 and 1.27, the average of which is centred very near to 1 (1.095). This reflects that the return series from the adjusted market portfolio is a relatively better approximate of the population as it is displays an increased amount of similarity to the median returns. As such, the resultant portfolio post the above two adjustments is selected to represent the benchmark portfolio. The return series generated from both the average market portfolio and the adjusted mean market portfolio (benchmark portfolio) is shown below:

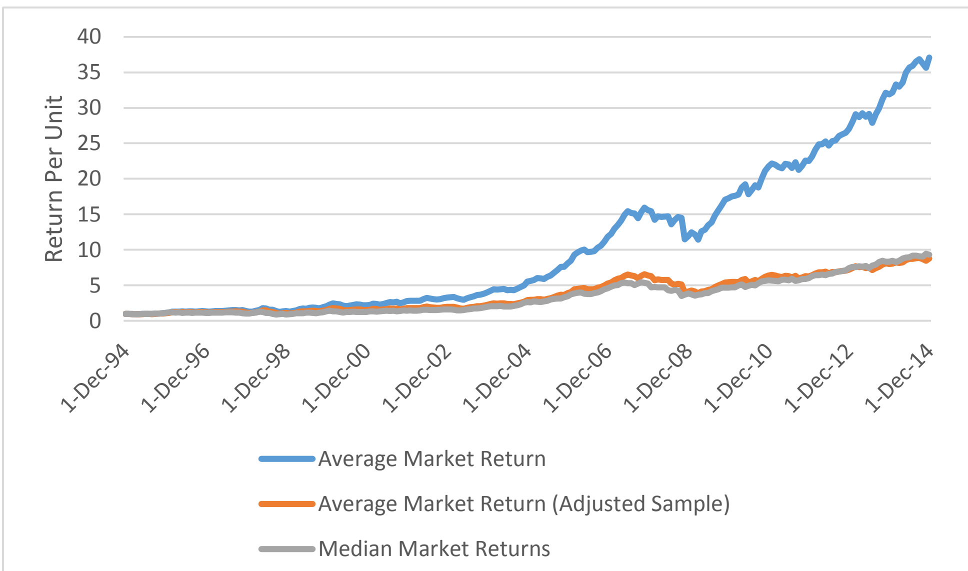


Figure 30: The unitised return series comparing the difference between the returns of the mean market portfolio, adjusted mean market portfolio and the median market portfolio

The vertical axis in Figure 30 represents a unitised return, for example, if the portfolio reaches a value of 5 it can be interpreted that if an investor had invested one unit on the first trading day of the time series, they would currently have 5 units – the investor would have earned a 500% return, absent of costs as previously discussed. Figure 30 illustrates the marked difference between the return series of the original mean market portfolio and the adjusted mean market (benchmark) portfolio. The median return series is also plotted above and is shown to be much better approximated by the benchmark portfolio as opposed to the unadjusted mean market return series.

Analysis of Returns

As mentioned above, this section includes the analysis of the returns achieved by both the unconstrained equally weighted portfolios. This examination compares the monthly returns achieved by both the simulated and genetic programming approaches to the benchmark portfolio described previously. The monthly returns are calculated as an equally weighted average of the returns achieved by each listed share included in each respective portfolio. As previously stated, shares that were not listed over the entire observation period were still included in the population with the aim to reduce survivorship bias. The inclusion of these shares can, however, skew the portfolio returns downward. This can occur, for example, where a share is assigned a weighting and as such is included into the portfolio. If this share is not listed over a period, thus presenting a return of 0 for the respective unlisted month, it will drag the return of the portfolio downward.

This limitation can be overcome in the case of an equally weighted portfolio. In the event a share is not listed over a specific time period, the weightings of each of the remaining shares included in the portfolio can be subsequently reweighted on an equal basis. These weightings are derived from the number of shares actively employed by the portfolio in a specific month. This construct is employed for the unconstrained equally weighted portfolios generated by the simulated approach and the genetic programming approach as well as the when calculating the returns achieved by the benchmark portfolio expounded upon above. Notably, this construct would not be necessary if shares included in the population were restricted to only those that were listed over the entire period. However, employing this restriction would lead to survivorship bias as previously explained.

The above observation insinuates that the returns achieved by the unequally weighted portfolios cannot be as transparently evaluated compared to the equally weighted portfolios. This is primarily due to the reasoning that the portfolio cannot be rebalanced with equally distributed weightings, as this would not reflect the varying weightings assigned to each share. As such, it would be necessary to include the return series of all selected shares, even over periods when they were not listed and thus reflected a return of 0. This would lead to a false reflection of the returns earned by the portfolio as including the return series of unlisted shares leads to the portfolio return being drawn lower, as described previously. The effect of this would be exacerbated if the weightings on the unlisted shares are relatively large. Due to this, the return analysis is performed across approaches but only on the portfolios where equal weightings were employed.

Through the use of the method mentioned above, the return series achieved by the unconstrained equally weighted simulation approach and by the equally weighted genetic programming approach can be analysed and compared to the benchmark portfolio. These comparisons are completed in the sections to follow. The return performance of the optimally diversified portfolio obtained when the equally weighted simulation approach is applied is shown in the graph below:

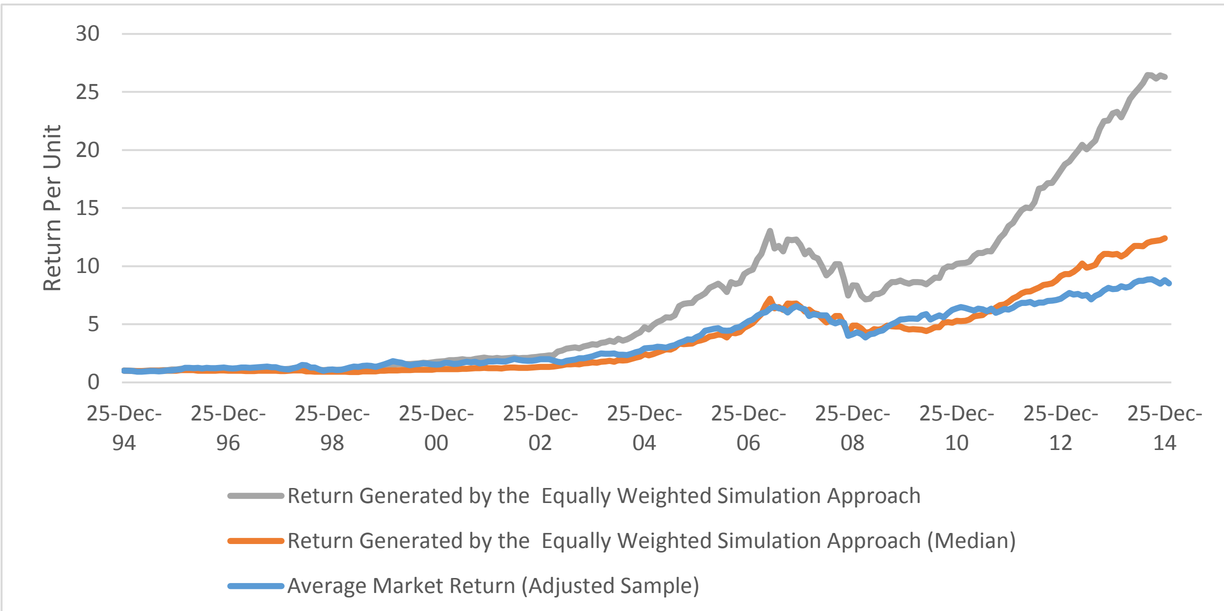


Figure 31: Unitised return comparison of the equally weighted simulation approach against the equally weighted median simulation approach and the benchmark portfolio

Figure 31 demonstrates the unitised return series (as described above) generated by the optimally diversified portfolio when the equally weighted simulation approach was applied, absent of transaction costs. This is compared to the median points achieved by the equally weighted simulation approach as well as against the benchmark portfolio. It is observed that although the return of the portfolio was not directly optimised, the return of the equally weighted portfolio generated by the simulation approach did in fact outperform the benchmark portfolio. This was true of both the portfolio computed as well as when the median return series was employed. The return of the actual portfolio appears to be unrealistically high and markedly different to the median return earned by the portfolio, indicating that there may be one or more shares with an outlying return distribution. This is more likely in the simulated portfolio given the relatively larger number of shares (37) in the portfolio, as discussed previously.

The figure below demonstrates the performance of the equally weighted simulated portfolio against the benchmark portfolio as a ratio:

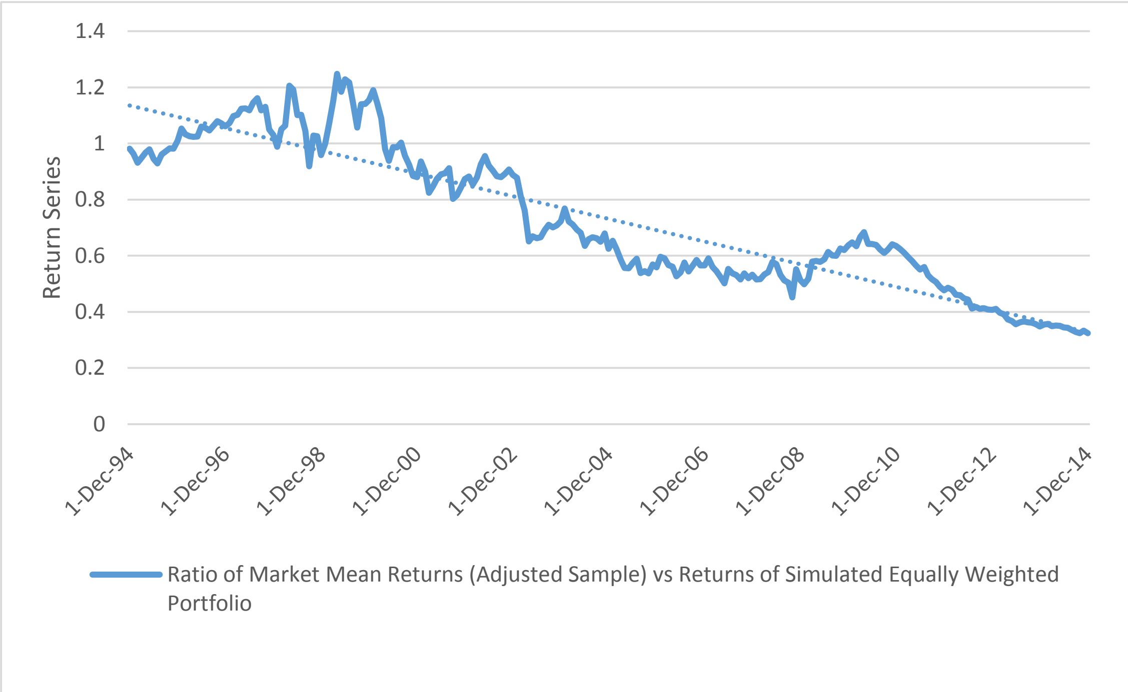


Figure 32: Ratio of market mean portfolio returns (benchmark) against the returns achieved using a simulated equally weighted portfolio

Figure 32 displays the ratio of the return series generated by the benchmark portfolio and the series generated by the optimally diversified portfolio solved for by using an equally weighted simulated approach. The downward trend illustrates that over the time period observed, the simulated portfolio continued to outperform the adjusted benchmark portfolio.

Further insights can be gleaned by examining the industry split achieved from the above simulated portfolio and thereafter completing a comparison to the industry spilt displayed by the entire population. These insights can be collected through the analysis of the figure below. Figure 33 is constructed through analysis of the number of companies present in each sector of the simulated portfolio, as opposed to being weighted by the market capitalisation of the respective companies. The figure is created in this way as the simulation assumed equal weightings on shares and as such is the correct method to depict the diversification across industries of the resulting portfolio. The figure displays that 8 of the 11 potential industries of the population are represented in the above portfolio. Moreover the financial sector is shown to be notably over-weighted while the basic materials sector is slightly underweighted compared to the portfolio consisting of the full population. The industries that were not represented in the above portfolio constituted 5.3% of the population portfolio cumulatively. Due to the various industries that the above portfolio has included, it can be said that the above portfolio has been diversified across sectors.

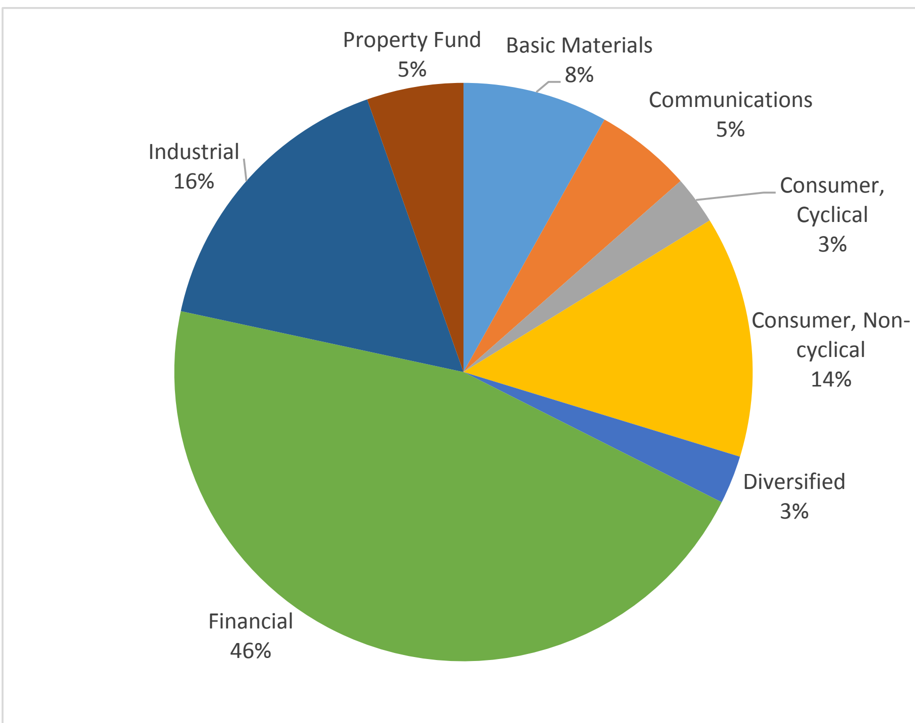


Figure 33: Industry split achieved by the optimal portfolio generated by the equally weighted simulation approach

4.2.1b Unequally Weighted Portfolio

The simulation approach was then applied to a portfolio of shares where the weighting on each share is not restricted to be equal. The minimum portfolio variance solved for in this approach consisted of a portfolio of 13 shares, giving a monthly portfolio variance of 0.592. This was solved for at a time of 9:50 minutes and at simulation number 3229. Although this variance is larger than the equally weighted portfolio, this is due to the exponentially larger possible number of possibilities the portfolio could utilise. The graph of the simulation progress is shown below:

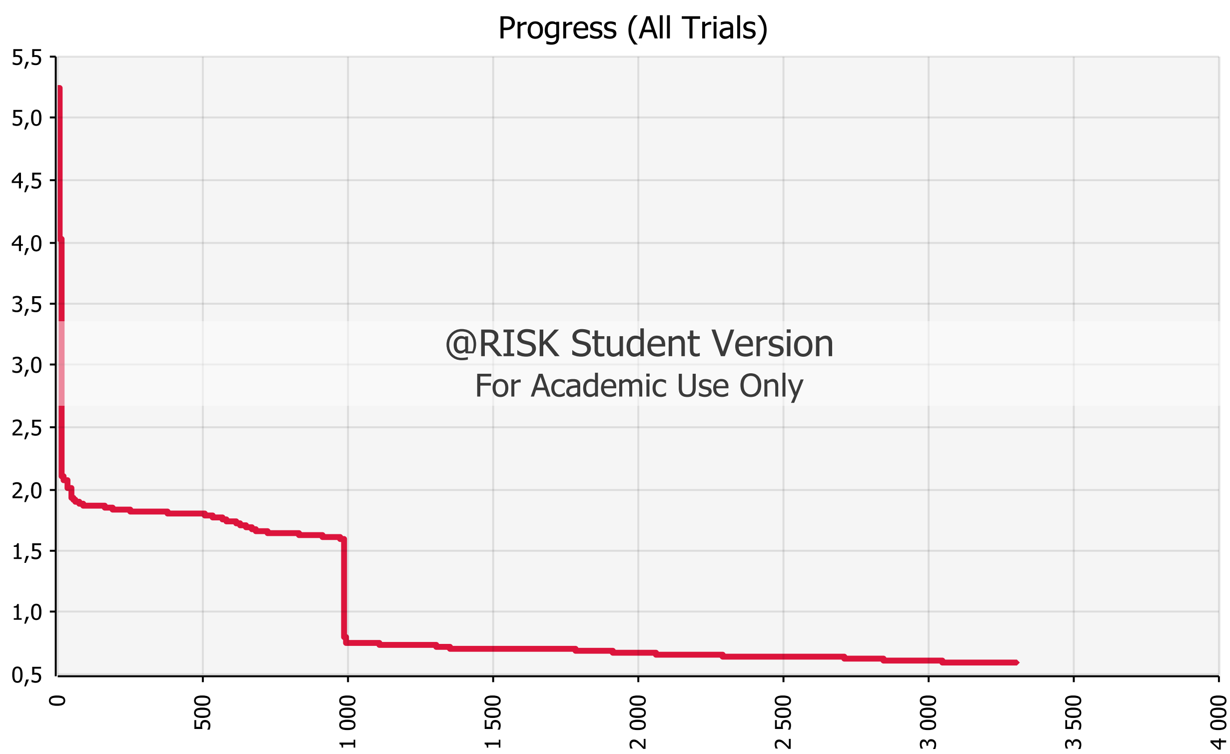


Figure 34: Progression of the simulation when the number of shares to be included in the portfolio is not pre-specified and the weightings on each share are not required to be equal

The initial portfolio variance was solved to be 5.245. Within 500 iterations this was reduced to 1.8. The table of the overall performance of the simulation is shown below:

Table 16: *Summary table of the simulation displayed graphically in Figure 34*

Results	
Valid Trials	3302
Total Trials	3303
Best Value Found	0,592347141
+ soft constraint penalties	0,00
= result	0,592347141
Best Trial Number	3229
Time to Find Best Value	0:09:50
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Table 16 shows that a total of 3 303 simulations were run and of that 3 302 were valid (99.7%). The minimum variance portfolio found was at a time of 9:50 minutes and was equivalent to 0.592.

The industry segmentation achieved from the above portfolio is then compared to the industry spilt of the population portfolio. The industry split of the above simulation is shown in Figure 35. The figure demonstrates the industry split of the resulting portfolio based on the number of shares representing each industry. This is equivalent to the resultant portfolio being equally weighted. The industry split displayed of the above portfolio, when each share is weighted according to the optimal weighting scheme illustrated above is then discussed further in the following section.

Comparing the diversification across industries achieved by the above approach against the industry split of the population portfolio shows that the approach covered 5 of the available 11 industries. The industries that were not included constituted 16.57% of the population portfolio. Although below 50% of the industries were utilised in the above approach, the industries that constituted the majority of the population were represented. As previously stated, the industries that were not utilised in the above approach amounted to 16.57%, which is notably higher compared to the 5.3% not used from the equal weightings simulation. This points against the initial hypothesis that allowing the weighting of the shares to vary will provide a better optimally diversified portfolio with a minimum number of shares compared to the equally weighted simulation. It should however be noted that the variance of the first simulation, 0.395, is only slightly lower than the variance of this simulation of 0.592. Conversely this variance is achieved with less than half of the number of shares – 37 shares were utilised in the first simulation while only 13 were necessary in this simulation.

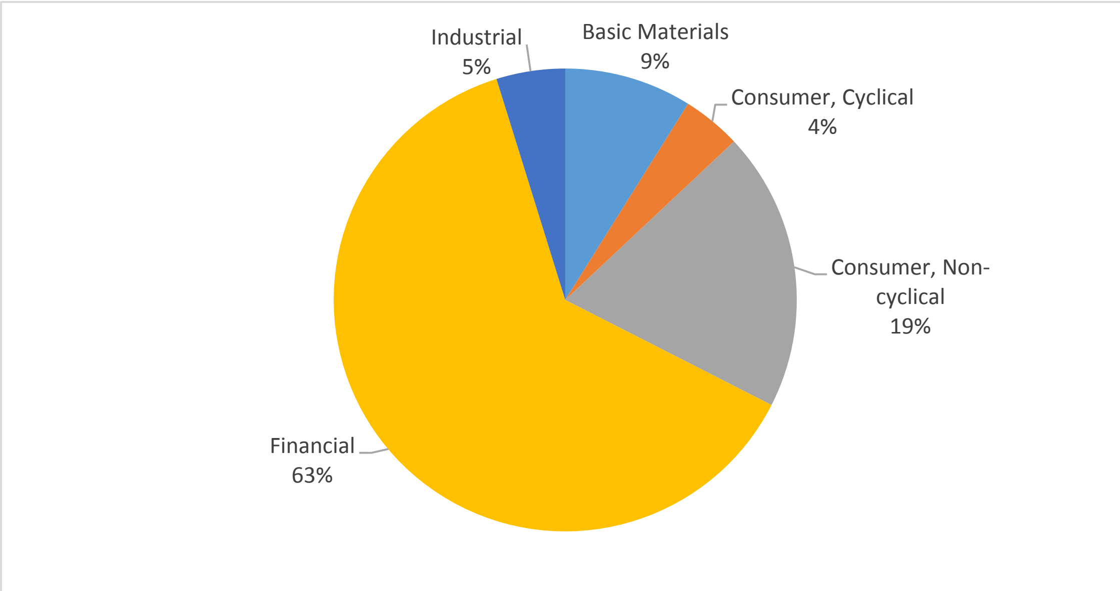


Figure 35: Industry spilt (number of shares representing each industry) achieved by the optimal portfolio solved for using a simulation approach with unequal weightings.

The figure below illustrates the resulting industry split when the shares constituting the optimally diversified portfolio are weighted according to the optimal weights calculated using the simulated approach:

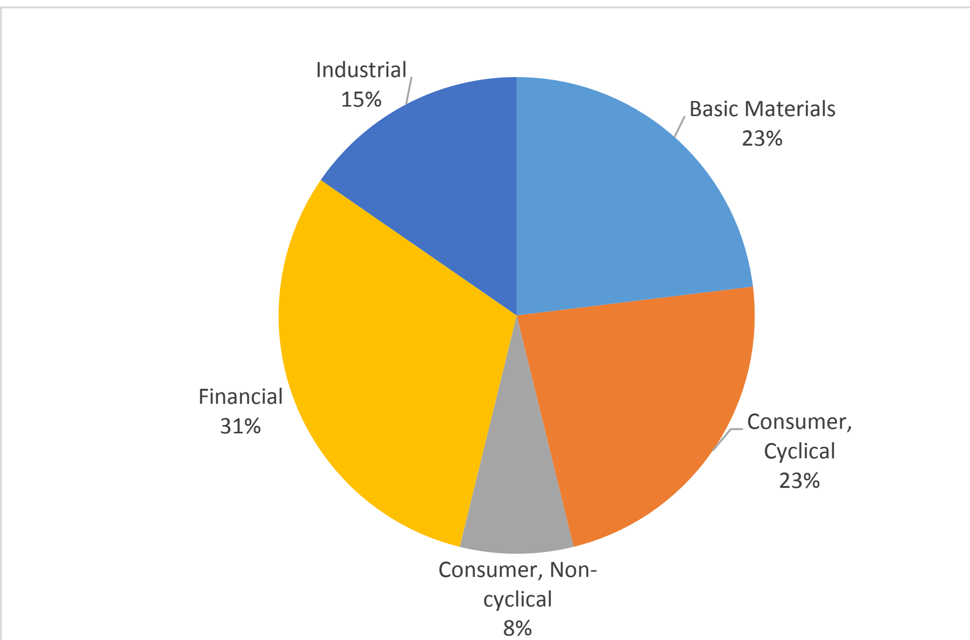


Figure 36: Industry spilt (weighted result) achieved by the optimal portfolio solved for using a simulation approach with unequal weightings.

When the weightings solved for by the unequally weighted simulation approach are applied to the portfolio Figure 36 is obtained. This is the true representation of the resulting diversification benefits obtained by an investor. This demonstrates that a much larger weighting is placed on the financial sector when compared to the industry split demonstrated by the market portfolio. It is interesting to note however, that of the financial firms included in the above portfolio, none were banks. The financial shares included were non-bank institutions such as insurance providers and financial companies with large exposures to property and real estate. The companies with large exposure to property and real estate dominated the representation of financial institutions – of the 63% of the portfolio that was represented by financial institutions, 98% of these were companies with exposure to local property and real estate as opposed to the remaining 2%, represented by insurance firms. It is interesting to observe that of the financial firms selected by this approach, the majority were not listed over the entire observed period. One firm was listed from 1991-1996 while the remaining were all listed post the financial crisis that occurred in 2008. Therefore it is understandable why these companies featured with the relatively large weighting as demonstrated in the figure above – firms with large exposure to property experienced a relatively stable return series (with the exception of during the period of the 2008 financial crisis). The lower volatility in the property sector in addition to its relatively non-cyclical nature explains why the approach would have heavily weighted these shares.

Chapter 4 continued in this section (4.2) which aimed to solve for the minimum number of shares necessary to obtain a minimum variance portfolio if the restriction that the portfolio should consist of a pre-specified number of shares was removed. This analysis commenced with section 4.2.1a, analysing the results obtained through the application of the unconstrained equally weighted simulation approach. It was demonstrated that the optimally diversified portfolio should employ 37 shares, exhibiting a minimum variance of 0.3954. This was notably the only simulation across all tests that employed greater than 30 shares as part of the optimally diversified portfolio. The computation of the benchmark portfolio returns was then completed in order to contribute to the analysis wherein the returns achieved by the unconstrained equally weighting simulation approach is compared to the returns achieved by the benchmark portfolio. The unconstrained equally weighting simulation approach outperformed the benchmark portfolio. Analysis of the diversification across industries was thereafter examined and the portfolio deemed optimally diversified was shown to be considered representative of the population portfolio.

In section 4.2.1b the results from the unconstrained unequally weighted simulation portfolio were examined. The optimally diversified portfolio employed 13 shares and obtained a minimum variance of 0.592. Additionally, the optimally diversified portfolio could once more be considered representative of the population portfolio, although, an overweighting of the financial sector was found. These results are further compared and expounded upon in section 4.4. This chapter continues with section 4.2.2 which inspects the results obtained by the unconstrained portfolios when a genetic programming approach is applied.

4.2.2 Genetic Programming Results

4.2.2a Equally Weighted Portfolio

In the same method as above, genetic algorithms were then applied to the data under the same restrictions. When the portfolio was restricted so that each share included had an equal weighting, the following results were obtained. The minimum variance portfolio consisted of 24 shares and yielded a portfolio variance of 0.254. This was found in the simulation 10 092. The portfolio variance calculated is notably 64.24% smaller than the 0.39542 calculated using the simulation approach. Nevertheless one should also note that over twice the number of simulations were able to be run in the same amount of time (genetic algorithms completed 11 361 simulations while the simulation approach completed only 3 583 simulations). The graph of the simulation performance is displayed below:

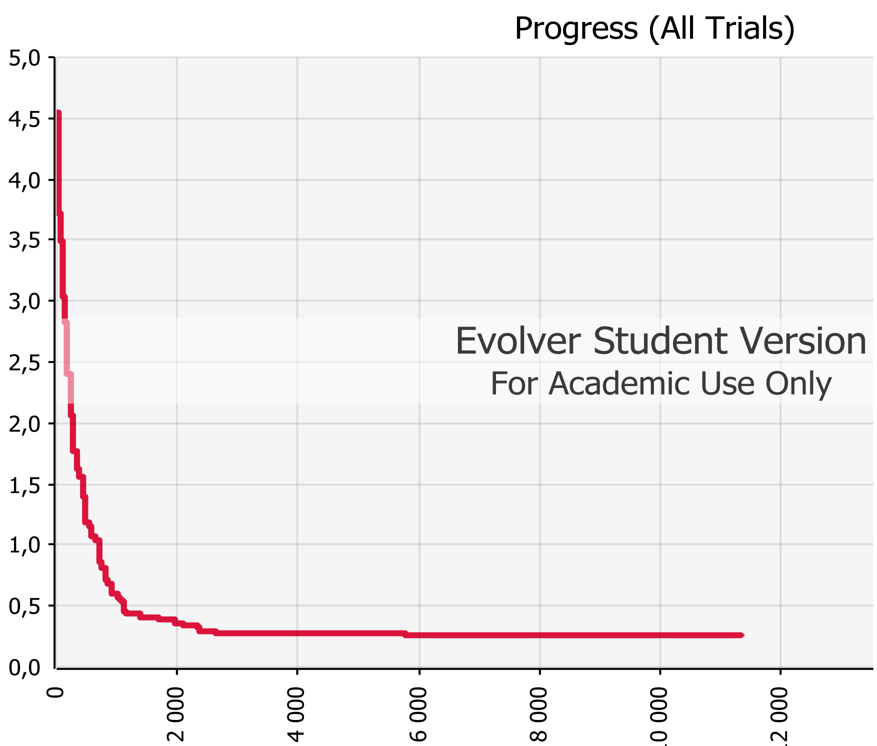


Figure 37: Progression of the simulation when the number of shares to be included in the portfolio is not pre-specified (equally weighted)

The initial valid simulation using the genetic algorithm approach yielded a minimum portfolio variance of 4.5. This is notably better than the simulation approach, producing an initial minimum portfolio variance of 70.654. The summary of the performance of this simulation can be seen in the table below:

Table 17: *Summary table of the simulation displayed graphically in Figure 37*

Results	
Valid Trials	11360
Total Trials	11361
Best Value Found	0,253713261
+ soft constraint penalties	0,00
= result	0,253713261
Best Trial Number	10092
Time to Find Best Value	0:08:58
Reason Optimization Stopped	Elapsed time
Total Optimization Time	0:10:00

Table 17 shows that in total, 11 361 trials were run of which 11 360 were valid. The minimum variance portfolio was found to consist of 24 shares and had a variance of 0.254. This was solved for on simulation number 10 092 at a time of 8:58 minutes. In this simulation genetic algorithms demonstrate a better performance than the traditional simulation method. Interestingly, this result is correspondingly better than all portfolios solved for in the previous analysis (where the number of shares to be included in the portfolio was entered as a constraint) up until the portfolio consisted of 30 shares.

Analysis of Returns

Although not the primary aim of this dissertation it is interesting to analyse the performance (with respect to the returns earned) of the portfolio classified as optimally diversified. Figure 38 exhibits the performance of the optimally diversified equally weighted portfolio when the genetic programming approach was applied. The figure demonstrates the return series achieved by the equally weighted portfolio when a genetic programming approach was utilised. As in the instance where the optimally diversified portfolio was obtained using the equally weighted simulation approach, the equally weighted genetic programming approach solves for an optimally diversified portfolio that also outperforms the benchmark portfolio. This holds true both when the actual equally weighted (mean) portfolio was utilised as well as when the median portfolio at each time period was used. The similarity between the actual performance of the portfolio and the performance of the median portfolio indicates that there is a relatively lower likelihood of the portfolio containing shares with an outlying return distribution.

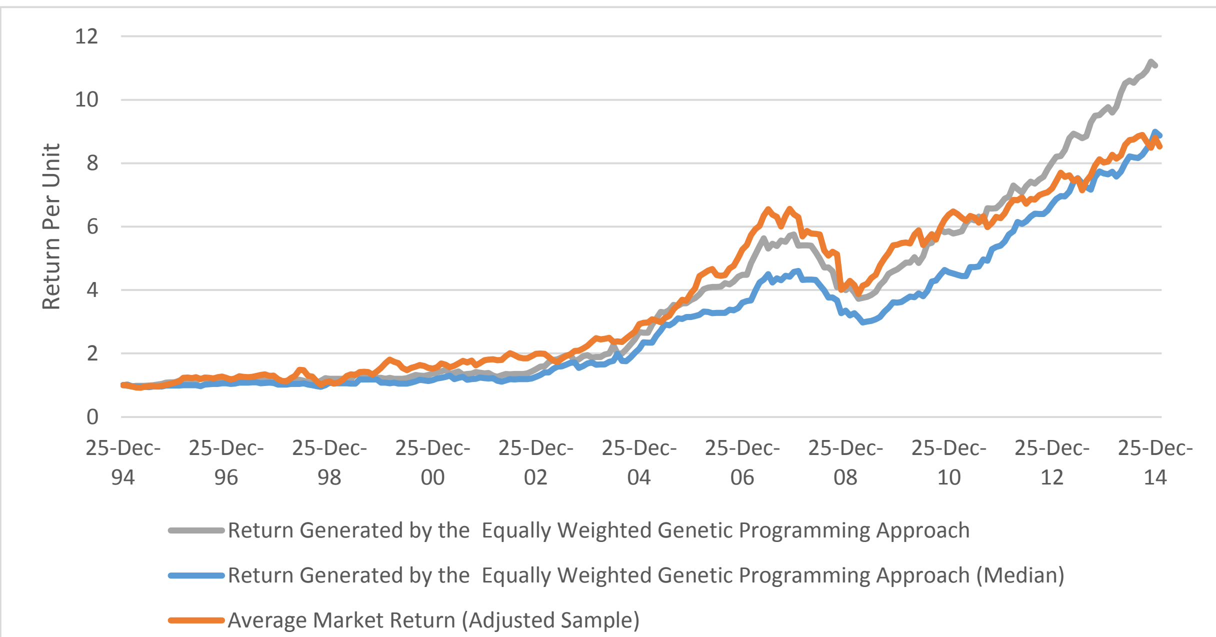


Figure 38: Return series of the optimally diversified portfolio when an equally weighted genetic programming approach was utilised benchmarked against the median return series achieved by the equally weighted genetic programming approach as well as the returns series of the benchmark portfolio.

The outperformance of the equally weighted genetic programming approach is displayed as a ratio to the benchmark portfolio in the figure below. It is observed that the ratio tends downwards across time which demonstrates the constant outperformance by the genetic programming portfolio. This is an important observation as this establishes that the outperformance of the equally weighted genetic programming approach against the benchmark portfolio is not due to one outlier in the series. Rather, genetic programming shows a constant trend of outperformance over the observed period.

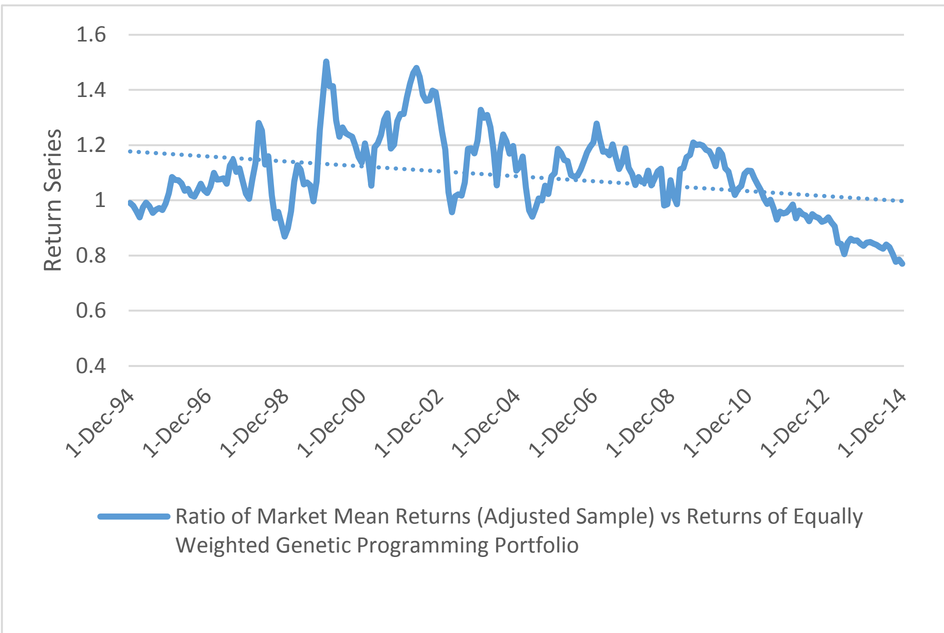


Figure 39: Ratio of returns earned by the benchmark portfolio against the returns earned by the equally weighted portfolio when a genetic programming approach was utilised

The analyses of industry split achieved from the above portfolio compared to the split observed in the population portfolio was then completed. Figure 40 illustrates these findings. It is displayed, in Figure 40, that 7 of the 11 potential industries of the population are represented in the above portfolio. Moreover, the financial sector is shown to be notably over-weighted while the basic materials sector is slightly underweighted compared to the portfolio consisting of the full population. The industries that were not represented in the above portfolio constituted 18.54% of the population portfolio cumulatively. A relatively fewer number of shares were employed in this portfolio (24 as opposed to 37 in the unconstrained equally weighted simulation approach) which led to a larger proportion of the population portfolio not being represented (18.54% opposed to 5.3% in the unconstrained equally weighted simulation approach). Despite this, the portfolio achieved a lower minimum variance than that obtained by the unconstrained equally weighted simulation approach. A further elaboration of this is carried out in section 4.4. Due to the various industries that the above portfolio has included, it can be said that the above portfolio has been diversified across sectors.

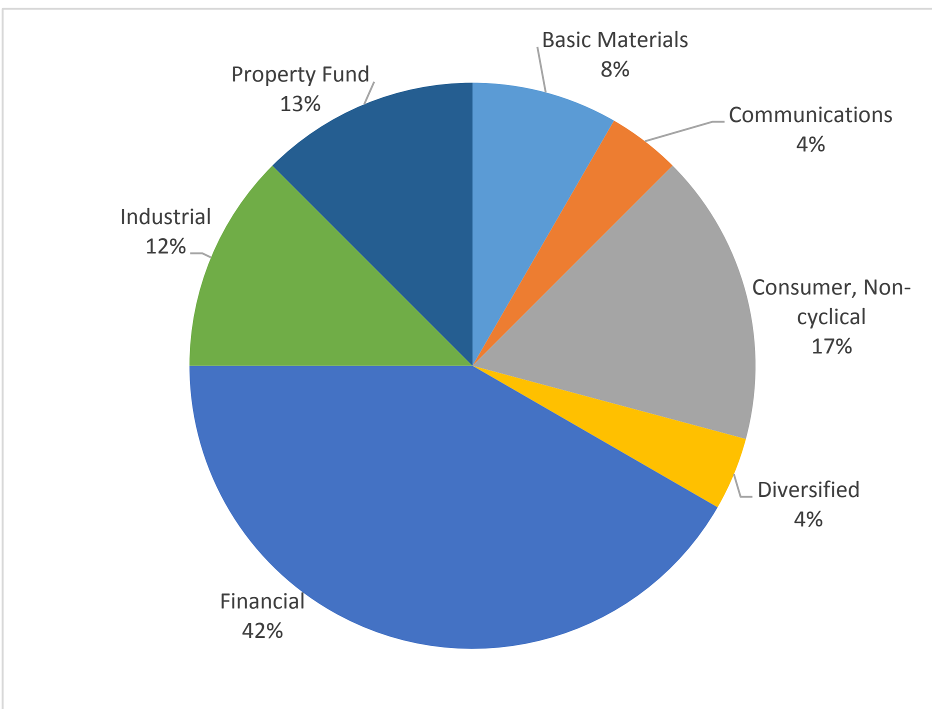


Figure 40: Industry split achieved by the optimal portfolio when implementing an equally weighted Genetic Programming Approach

4.2.2b Unequally Weighted Portfolio

When the weighting on each share in the portfolio were not restricted to be equal, the following set of results was found. The minimum variance portfolio was found at a time of 5:23 minutes, simulation number 4 553, and was found to be equal to 0.497. The initial valid minimum variance portfolio was found to be 3.519. This demonstrates an improvement to the initial minimum variance portfolio found using the simulation approach of 5.245. Furthermore, more simulations were able to be run in the same amount of time using genetic programming, 4 640 as compared to 3 303 in the traditional simulation approach. The minimum variance portfolio of 0.497 using 11 shares is better than the simulation approach of computing for a minimum portfolio variance of 0.592 whereby 13 shares were utilised. Once more the genetic programming approach outperforms the traditional simulation approach.

The graph of the performance of this simulation can be seen in Figure 41. The initial minimum variance portfolio is shown to be equivalent to 3.518. Within 500 simulations the minimum variance portfolio was 3.13. However, as the simulation continued the minimum portfolio variance tended towards the final figure found of 0.497 (achieved at a time of 5:23 minutes and at simulation number 4 553).

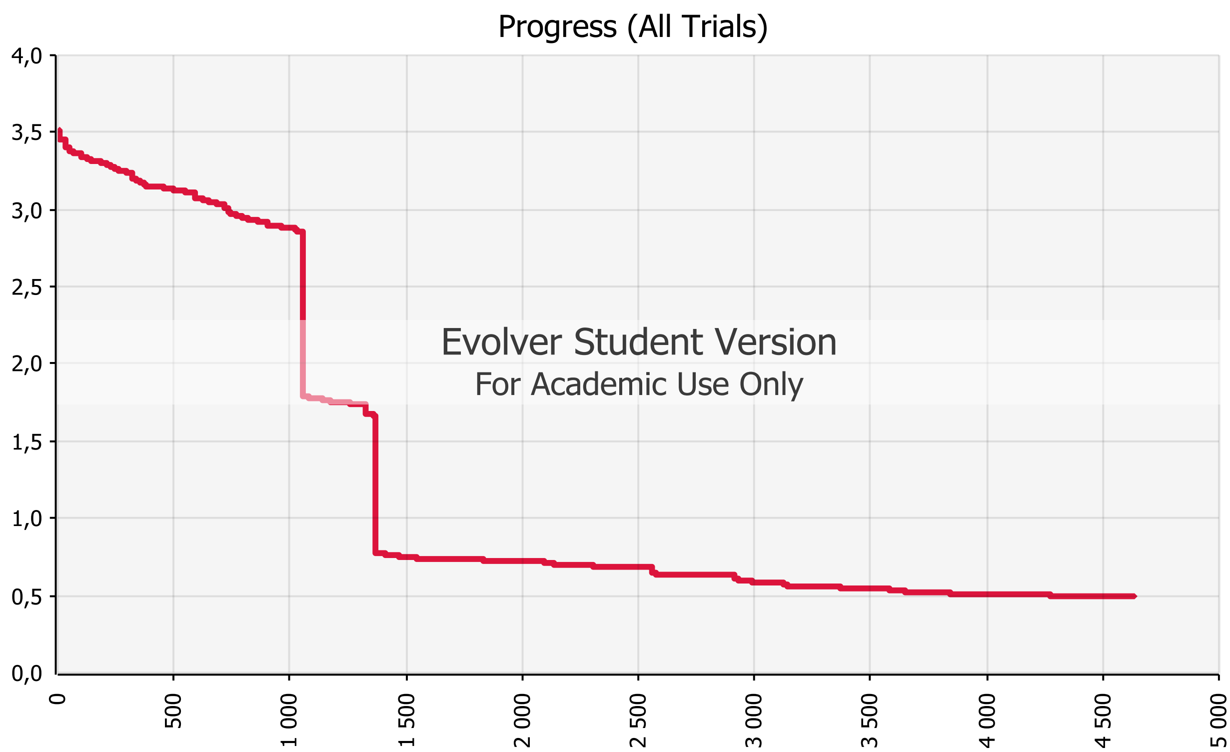


Figure 41: Progression of the simulation when the number of shares to be included in the portfolio is not pre-specified, weightings on each share are not required to be equal

The summary table of this simulation is shown below:

Table 18: *Summary table of the simulation displayed graphically in Figure 41*

Results	
Valid Trials	4639
Total Trials	4640
Best Value Found	0,497380714
+ soft constraint penalties	0,00
= result	0,497380714
Best Trial Number	4553
Time to Find Best Value	0:05:23
Reason Optimization Stopped	Progress condition
Total Optimization Time	0:05:27

Table 18 demonstrates that 4 640 trials were run, of which 4 639 were valid and a minimum portfolio variance of 0.497 was solved for at a time of 5:23 minutes.

An analysis of the industry diversification of the above portfolio yields the results shown graphically in Figure 42. The industry split displayed represents the number of companies included in the portfolio from each sector. It is shown that of the 11 potential industries that could be utilised, 6 were included. The industries that were not included contributed a combined total of 12.08% of the population portfolio. Although this portfolio obtained a higher variance than that of the equally weighted portfolio when applying genetic programming (0.498 against 0.254) it achieved this variance with less than half of the number of shares, using 11 against the 24 used by the previous simulation.

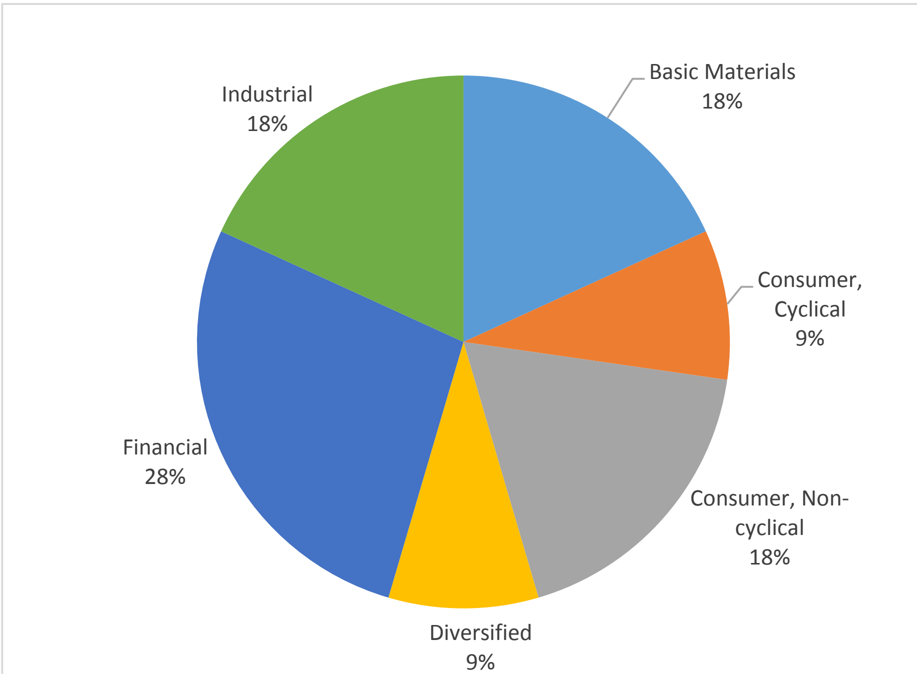


Figure 42: Industry split (number of shares representing each industry) achieved by the optimal portfolio when applying the Genetic Programming Approach with unequal weightings.

Once the respective weightings were applied to the shares included in the portfolio, the weighted industry split is obtained and displayed in Figure 43. As in the results obtained from the unequally weighted simulation approach, the financial sector is heavily over weighted. The companies representing the financial sector are the same companies as in the results obtained using the unequally weighted simulation approach described previously with one exception. There are no insurance companies included in the portfolio generated from using an unequally weighted genetic programming approach. The overweighting of the property and real estate industry is likely to have occurred for the same reasons as elaborated upon previously.

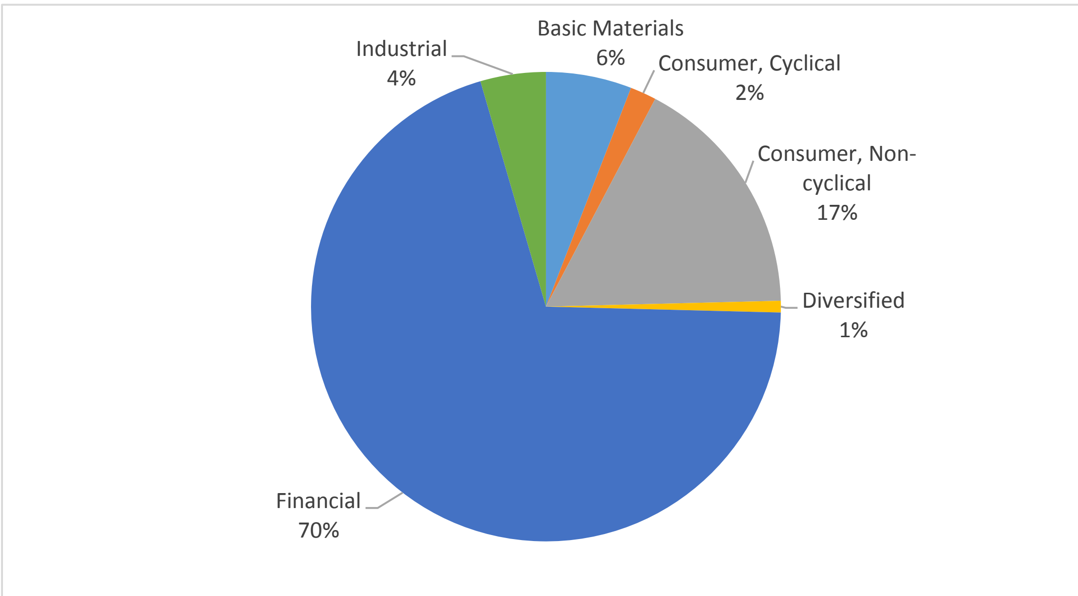


Figure 43: Industry spilt (weighted result) achieved by the optimal portfolio solved for using a genetic programming approach with unequal weightings.

Section 4.2.2 inspected the results obtained by the unconstrained portfolios when a genetic programming approach was applied. This was initiated with section 4.2.2a, examining the results achieved by the unconstrained equally weighted genetic programming approach. It was revealed that the optimally diversified portfolio consisted of 24 shares and exhibited a portfolio variance of 0.254. Subsequently, through the use of comparable analysis, it was discovered that the unconstrained equally weighted genetic programming approach delivered benchmark beating returns. Section 4.2.2b continued the chapter with the analysis of the unconstrained unequally weighted genetic programming approach. This approach made use of 11 shares to obtain a minimum portfolio variance of 0.497. In both sub-sections of 4.2.2 the optimally diversified portfolios could be observed as being representative of the population across industries, albeit that the financial sector experienced an overweighting. This is further discussed in section 4.4. The chapter continues with the analysis of the results achieved by the particle swarm optimisation approach.

4.3 Particle Swarm Optimisation Results

Notably absent in the above results are the results obtained from the Particle Swarm Optimisation simulations. This is due to the algorithm displaying extremely poor performance when attempting to minimise the variance of the portfolio. On numerous occasions hard constraints, such as the weight allowed on each share and the number of shares to be included, were not adhered to which led to numerous runs within each simulation being invalid and unable to be used. Furthermore outputs were extremely unstable, achieving a wide range of results, albeit none close to those achieved using a traditional simulation approach or the genetic programming approach.

In the first test, which attempted to solve for the minimum variance portfolio consisting of a pre-specified number of shares while maintaining equal weightings on each, the algorithm frequently achieved only 1 valid run where the constraints were met. Furthermore many of the simulations achieved no valid output throughout the ten minute run. When solving for the optimum portfolio when the number of shares to be included in the portfolio was not specified , but equal weightings on each share was maintained (test 2) the following set of results were obtained:

Table 19: *The Optimal number of Shares to be included in the Portfolio, requiring equal weightings on each share while required the sum to add to 1 using a Particle Swarm approach*

	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5
Optimal Number of Shares:	176	171	175	170	177
Monthly Portfolio Variance:	4.077	3.496	3.693	4.120	4.045
Annual Portfolio Variance:	6.994	6.477	6.657	7.032	6.967

The optimal number of shares to be included in the portfolio was significantly greater using this approach. It also did not achieve a lower portfolio variance than in the traditional simulation and genetic programming approaches which used a lower number of shares. When the weightings on the shares were allowed to vary, but the number of shares to be included in the portfolio was pre-specified (test 3), similar problems were encountered as in the approach that maintained equal weightings. Furthermore, the final test demonstrated that when the number of shares to be included in the portfolio was not pre-specified and the weightings were allowed to vary, no simulations gave valid results within a ten minute run.

The results above demonstrate the poor performance achieved by the Particle Swarm Optimisation approach. The performance is indiscriminately poor across each of the four tests. This is hypothesised to be due to the fact that when initialised, many particles did not satisfy the appropriate constraints. This severely limited the number of valid *pbest* (p_{id}) positions experienced by the particles. In addition, if no particles had valid positions when initiated, the *gbest* (p_{gd}) variable would not be assigned a value. This significantly reduces the influence that variables such as velocity (V_{id}) and acceleration (c_1 and c_2) ascribed to each particle can have in solving for the optimal portfolio. The combination of the above factors led to a minimal amount of interaction between particles; as explained previously, a particle by itself has almost no explanatory power. As such this limited the effectiveness of the approach. Finally, the restrictions placed on the above simulations led to a violation of the principles put forward by Millonas (1994), as discussed previously. Specifically, the third principle was violated as the required restrictions forced the model to act along an excessively narrow channel. The above results demonstrate an unstable series and illustrate the unsuitability of Particle Swarm Optimisation to be applied to combinatorial type problems with strict hard constraints. This establishes that the employment a more advanced approach does not ensure increased performance. These results agree with the findings of Kendall and Su (2005) where it was noted that in an instance where the population consists of more than 15 available shares, the PSO approach experiences extreme difficulties in performance. As mentioned previously, the current data set consists of 356 shares.

4.4 Direct Comparisons

The section below will directly address each hypothesis of this study. The first research question was: "what is the optimal number of shares a portfolio should consist of in order to achieve a minimum variance?" The hypothesis was that the optimal portfolio would consist of between 10 to 30 shares. In addition a sub-research question to this aimed to discover whether allowing the weights on each share in the portfolio to vary would unleash further diversification benefits. It was hypothesised that the more flexibility allowed through varying weights would unlock further benefits of diversification. The second research question aimed to examine whether the more complex mathematical models such as genetic programming and particle swarm optimisation would display increased strength in solving for the optimal number of shares to include into a minimum variance portfolio. The hypothesis was that genetic programming and particle swarm optimisation would indeed demonstrate increased strength when compared to the results found by the traditional simulation approach.

This section will test the above hypotheses through an analysis using the following structure. Firstly, the results found using a simulation approach with equal weightings will be compared to the results established using a simulation approach that allowed for varying weightings on each share. This analysis will take place by observing the trend of each test as well as comparative analysis of the optimal points achieved. This same analysis will then be completed through the comparison of the results achieved through genetic programming, where the results from portfolios consisting of equal weightings are compared to those achieved when the weightings were allowed to vary. Thereafter a comparison of the results where the number of shares to be included in the portfolio was not pre-specified will take place across both the traditional simulation and genetic programming approaches. Finally, an analysis of the diversification across industries of the above approaches will be completed.

In order to investigate the second research question and test the applicable hypothesis, similar comparisons and tests as mentioned above are carried out however, in this instance the simulation approach results are compared directly against the results achieved by the genetic programming methodology. The equally weighted portfolio achieved through the traditional simulation approach will be compared to the equally weighted portfolio of the genetic programming approach and so forth.

The table comparing the results of the optimal portfolio when utilising each approach is presented on the following page. This table will be referenced throughout the below analysis in order to answer the two research questions, thereby offering evidence to either reject or fail to reject the initial hypotheses.

Table 20: *Comparative table summarising the results of the Optimal Portfolios solved for across each Mathematical Approach*

	Simulation Approach, Equal Weightings	Simulation Approach, Unequal Weightings	Genetic Programming Approach, Equal Weightings	Genetic Programming Approach, Unequal Weightings	Simulation Approach, Single Share, Equal Weightings	Simulation Approach, Single Share, Unequal Weightings	Genetic Programming Approach, Single Share, Equal Weightings	Genetic Programming Approach, Single Share, Unequal Weightings
Optimal Number of Shares	10	15	14	15	37	13	24	11
Minimum Variance Achieved	0.307	0.744	0.266	0.679	0.3954	0.592	0.254	0.497
Optimal Solution Simulation Number	3652	3983	9234	9657	3559	3229	10092	4553
Optimal Solution Simulation Time	7 min 49 sec	8 min 55 sec	8 min 5 sec	9 min 47 sec	9 min 57 sec	9 min 50 sec	8 min 58 sec	5 min 23 sec
Minimum Variance achieved on First valid Trial	9.282	0.832	7.152	1.889	70.65	21.81	21.81	21.81
First Valid Trial Number	2	553	2	502	2	2	2	2
First Valid Trial Time	46 sec	1 min 42 sec	46 sec	55 sec	1 min 21 sec	43 sec	33 sec	33 sec
Total Trials; Per cent of Valid trials	4762; 99.98%	4484; 19.29%	11969; 99.9%	9931; 21.1%	3583; 99.97%	3303; 99.7%	11361; 99.99%	4640; 99.98%

4.4.1 Simulation Approach (Equal Weightings) compared to Simulation Approach (Unequal Weightings)

Reviewing the above results it is shown that both of the above approaches improved upon the results found by previous literature. In chapter 2 it was revealed that numerous studies inferred that an optimally diversified portfolio should consist of between 10 and 30 shares (Elton & Gruber, 1977; Neu-Ner & Firer, 1997). Both the simulation approach with equal weightings and the simulation approach with unequal weightings indicated that the optimally diversified portfolio should consist of 15 shares or less. This reinforces the first respective hypothesis – the optimally diversified portfolio is shown to consist of between 10 and 30 shares. The second hypothesis however, was not supported from the above analysis. This hypothesis stated that by allowing the weights on each share to vary, the optimal portfolio should be enhanced and further diversification benefits unlocked. This proved not to be the case, the first approach solved for a portfolio consisting of 10 shares that yielded a variance of 0.307. This is notably superior to the portfolio solved for by the second approach; this portfolio consisted of 15 shares and yielded a variance of 0.744. The first approach also solved for this superior portfolio using a lesser number of simulations and in a quicker amount of time. However, it should be noted however that a result achieved on a portfolio where all of the shares are required to be weighted equally, should act as the lower bound of the potential best portfolio that could be established if the weights were allowed to vary. This is due to the fact that adding further flexibility to the methodology should only improve the results; if an equally weighted portfolio is indeed the optimal portfolio, using an approach where the weights on each share were allowed to vary would merely solve for the optimal portfolio which in this case would be an equally weighted portfolio.

The results found are not in agreement with the above argument. This could be due to the limited computing power available to the simulations. As previously discussed, allowing the weights on each share to vary exponentially increases the number of potential portfolios that can be constructed. It can be assumed that if sufficient computing power was available or if the simulation was allowed to run for a vastly increased amount of time, an improved result may have been established. This theory explains why varying the weights on each share as opposed to maintaining an equally weighted portfolio may have led to inferior results. This is supported by the results – the first valid trial number when the weights were equal was solved for in 46 seconds and found on the second simulation. When the weights were allowed to vary the first valid trial number was found after 1 minute and 42 seconds – this was simulation number 533. Furthermore, 99.98% of the trials run in the equally weighted simulation were valid but merely 19.29% of the trials were valid when the weightings on each share were allowed to vary. In summary, the above comparison reinforces the first hypothesis that an optimal portfolio does indeed consist of between 10 to 30 shares. The second hypothesis is not however proven – allowing the weights to vary does not in fact lead to an enhanced performance in solving for the optimal portfolio.

A graphic illustration of the results achieved from the traditional simulation approach (equally weighted and unequally weighted) when the portfolio was restricted to incrementally consist of a number of shares, ranging from 1 to 30 is shown in Figure 44. It can be seen from the figure that the expected trend (the minimum variance portfolio initially begins higher for a lower number of shares and dramatically decreases up to a point where including additional shares to the portfolio adds only a slight diversification benefit) is observed more noticeably in the simulation where the weights have been restricted to be equal. The trend of the equally weighted approach above is unstable (in some cases adding a share in fact increases monthly variance). This instability is exacerbated when the weightings allowed on each share are variable. It is expected that both trends should display a stronger trend if the simulations were allowed to run for longer (albeit potentially excessive) amounts of time, as mentioned previously. The reasoning for the weaker trend in the instance with unequally weighted constituents is discussed previously, and can be said to be due to the increased quantum of potential portfolios. Furthermore, the figure reiterates the findings mentioned earlier - allowing the weights to vary does not provide additional benefits to diversification. It is shown however, that the minimum variance portfolio contains less than 30 shares, demonstrating alignment to the findings of Neu-Ner and Firer (1997) and those of Statman (1987), with a slight improvement as fewer shares were deemed necessary to be employed (between 10 and 15), amongst other previous authors as discussed in Chapter 2. The analysis now continues to further answer the first research question through a similar comparative analysis, now aimed at the results achieved from the genetic programming approach.

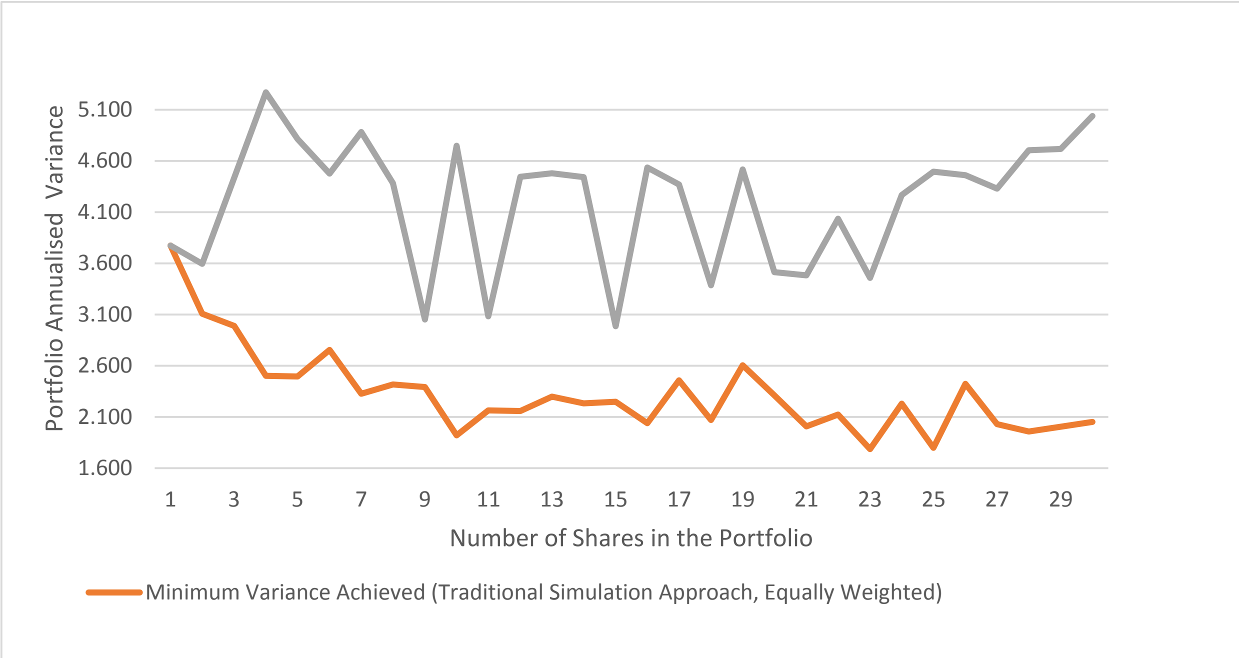


Figure 44: Traditional Simulation Approach: Comparison of equally weighted and unequally weighted results

4.4.2 Genetic Programming Approach (Equal Weightings) compared to Genetic Programming Approach (Unequal Weightings)

The results achieved when using the genetic programming approach reinforce those found by the traditional simulation approach. Both of the equally weighted and unequally weighted portfolios were shown to consist of between 10 to 30 shares – the equally weighted approach consisted of 14 shares while the unequally weighted approach consisted of 15 shares. This is again in line with previous literature, such as Neu-Ner and Firer (1997) and that of Statman (1987) as discussed in Chapter 2, as well as with the hypothesis of this dissertation. Similarly to the traditional simulation approach, the genetic programming approach does not find additional benefits to diversification when allowing the weightings of individual shares to vary. When the weightings on each share were equal, the optimal portfolio consisted of 14 shares which yielded a portfolio variance of 0.266. Utilising an unequally weighted portfolio increases the number of shares necessary to be held in the portfolio to 15 while also achieving a higher portfolio variance (less efficient) of 0.679. Furthermore, the equally weighted approach utilised less simulations to solve for the optimal portfolio and did so in a shorter time period. The earlier argument again applies to the underlying reasoning as to why this may be the case and is given further impotence due to the results illustrating that 99.9% of the trials were valid when equal weights were used while only 21.1% of trials were valid when the weights were allowed to vary. The first valid trail number was also the second trial on the equally weighted approach but was only the 502nd trial on the variably weighted portfolio.

A graphical illustration of the results achieved by the genetic programming approach is shown in Figure 45. Once again it is shown that the simulation with the clearer trend is the equally weighted simulation. The graph displayed in Figure 45 again demonstrates that the optimal portfolio consists of a portfolio of 15 shares or less, reinforcing our initial hypothesis. The results do however again exhibit that allowing for more flexibility in the simulations does not in fact enhance the diversification benefit available to the investor. This could be due to the arguments noted previously in this section.

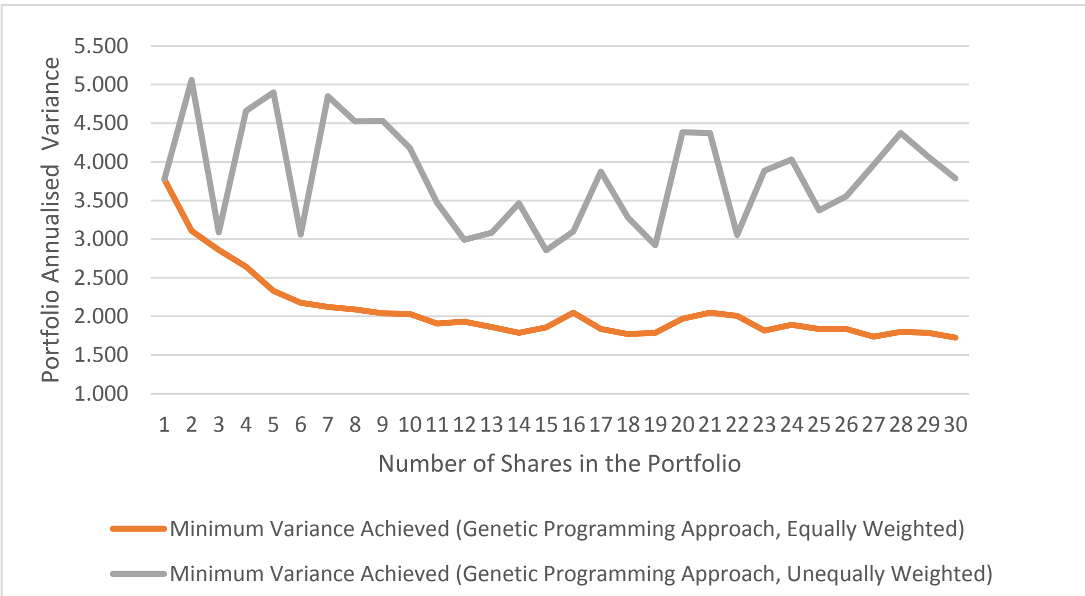


Figure 45: Genetic Programming Approach: Comparison of equally weighted and unequally weighted results

Through the use of the above comparative analysis the first research question and associated sub question are answered. The optimal portfolio can be seen to consist of between 10 to 15 shares, in agreement with the initial hypothesis. Secondly, contrary to the hypothesis, allowing for more flexibility in the simulations by way of allowing the weights of each share in the portfolio to vary, does not in fact enhance the diversification benefits available to an investor.

This section now continues with further comparative analysis with the aim to answer the second research question: "Of the three selected mathematical approaches, which demonstrates the most strength when solving for the optimally diversified portfolio in a South African environment?" It is hypothesised that the more computationally intensive and complex model of genetic programing will display increased strength in solving for the optimally diversified portfolio as opposed to the traditional simulation method.

4.4.3 Simulation Approach (Equal Weightings) compared to Genetic Programming Approach (Equal Weightings)

Both the equally weighted simulation approach and the equally weighted genetic programming approach achieve an optimal portfolio containing between 10 and 15 shares. The equally weighted simulation approach utilises a portfolio containing 10 shares yielding a variance of 0.307. The equally weighted genetic programming approach draws upon 4 additional shares to obtain a variance of 0.266 which is a 13% reduction in portfolio variance compared to the equally weighted simulation approach. In addition, although the genetic programming approach used 2.5 times more trials to find the optimal portfolio it also completed 2.5 times more trials in total compared to the equally weighted simulation approach, with similar levels of these trials being valid (satisfying all constraints). Both approaches found the first valid simulation on the second trial although it is acknowledged that this trial yielded a lower variance when using genetic programming. The results of these tests, when the number of shares to be included in the portfolio increases from 1 to 30 incrementally is depicted graphically in Figure 46. The figure shows that an equally weighted genetic programming approach achieves a stronger, more stable trend when compared to the results from an equally weighted traditional simulation approach. Furthermore, it can be seen that an equally weighted genetic programming approach solves for a minimum variance that is smaller than that solved for by the traditional simulation approach for 22 out of the 29 observations.

This reinforces the second hypothesis: the more complex mathematical approach of genetic programming does indeed yield better diversification benefits to an investor.

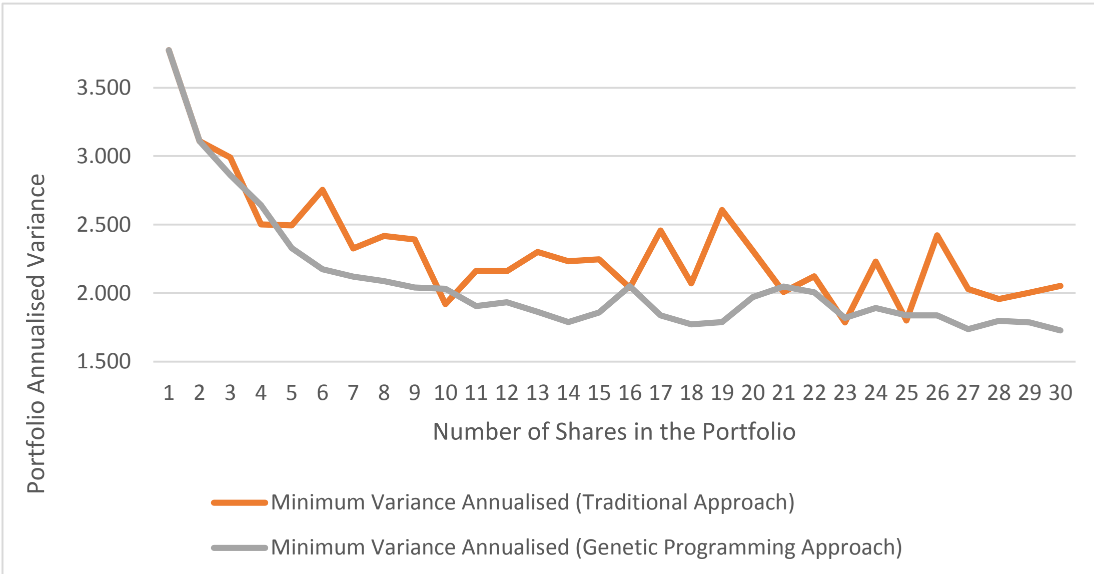


Figure 46: Traditional Simulation Approach and Genetic Programming Approach: An equally weighted comparison when the number of shares to be included in the portfolio was pre-specified

4.4.4 Simulation Approach (Unequal Weightings) compared to Genetic Programming Approach (Unequal Weightings)

The above finding is reinforced when comparing the results achieved by each approach when the weights of the shares are allowed to vary. Both approaches solve for an optimal portfolio which consisted of 15 shares, however, the portfolio variance achieved by genetic programming is 8.74% lower than that achieved by the traditional simulation approach. Yet the genetic programming approach used 2.4 times more trials to solve for the optimal portfolio but completed 2.2 times more trials in total. Both tests experienced a similar level of trials being valid, 19.29% using a traditional simulation approach and 21.1% using a genetic programming approach. The genetic programming approach did however take longer to solve for the optimal portfolio.

An illustration of the results achieved over both approaches is presented in Figure 47. This figure illustrates that both approaches yield a less clear trend. Although, it is calculated that the genetic programming approach remains superior as on 21 of the 29 observations, the genetic programming approach yields a lower variance while using the same number of shares in the portfolio. Consequently, this comparison highlights that genetic programming demonstrated more power in solving for the minimum portfolio variance compared to the traditional simulation approach.

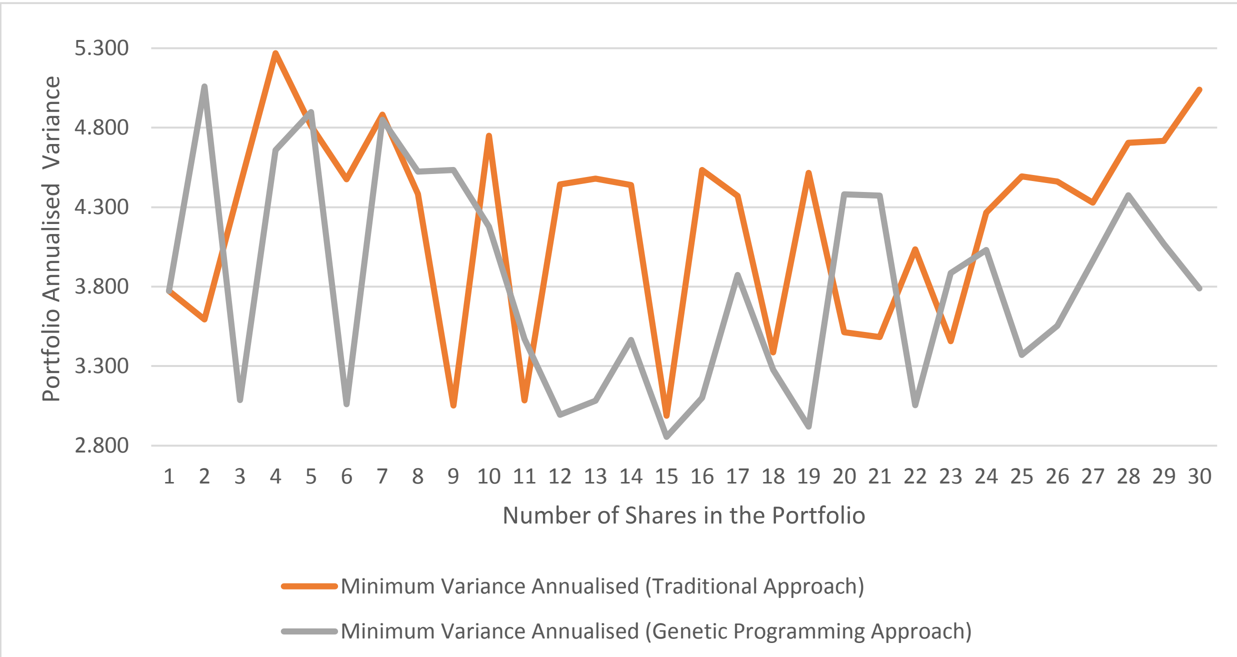


Figure 47: Traditional Simulation Approach and Genetic Programming Approach: An unequally weighted comparison when the number of shares to be included in the portfolio was pre-specified

4.4.5 Unconstrained Portfolio Analysis

The analysis now continues through a similar comparative procedure, now aimed at the tests when the number of shares to be included in the portfolio were not pre-specified. When this approach was used under the constraint of equal weightings, the following results were obtained:

The traditional simulation approach solved for a portfolio consisting of 37 shares which yielded a portfolio variance of 0.3954. This was notably the only result that was not in agreement with the first hypothesis that an optimal portfolio should consist of between 10 and 30 shares. When the simulation approach was applied without the constraint of equal weightings the optimal portfolio consisted of 13 shares and yielded a portfolio variance of 0.592. The results show that although less shares were utilised when the traditional simulation approach with unequal weightings was used, the additional flexibility allowed to the approach did not lead to further benefits of diversification being uncovered for the investor.

A similar result was uncovered when the genetic programming methodology was applied. Under this methodology, the equally weighted portfolio solved for the optimal portfolio consisting of 24 shares which yielded a portfolio variance of 0.254. When the weights were allowed to vary the optimal portfolio was shown to consist of 11 shares which yielded a portfolio variance of 0.497. Once more it is observed that although less shares were utilised, the more flexible approach did not uncover further additional benefits of diversification available to the investor. There was however, a notable drop in both the total number of trials completed as well as the trial number when the optimal solution was found. These results again reinforce the findings made previously, the optimal portfolio can be said to consist of between 10 to 30 shares and additionally, the more flexible approach does not in fact yield additional diversification benefits to an investor.

In order to further answer the second research question, regarding which approach displays the most strength in solving for the optimal portfolio, the comparisons are now made across the mathematical approaches when the number of shares to be included in the portfolio were not pre-specified. When the weightings on each share are required to be equal the simulation approach uses a portfolio consisting of 37 shares yielding a variance of 0.3954. The genetic programming approach uses both fewer shares and achieves a lower portfolio variance, the minimum variance portfolio consisted of 24 shares and yielded a portfolio variance of 0.254. The portfolio variance calculated is notably 35.76% smaller than the 0.39542 calculated using the simulation approach while also utilising fewer shares. In addition it should also be noted that although the genetic programming approach used 2.8 times more trials in order to solve for the optimal portfolio, the approach completed 3.2 times more trials in total.

When the weightings on each share in the portfolio were allowed to vary the traditional simulation approach solved for an optimal portfolio consisting of 13 shares which yielded a variance of 0.592. Once again the genetic programming approach utilised a fewer number of shares to obtain a smaller portfolio variance – 11 shares were utilised to yield a portfolio variance of 0.497. These results again reiterate previous findings. The optimal portfolio is demonstrated to include between 10 and 30 shares, with one exception. Secondly, allowing increased flexibility of the methodology by allowing the weightings on each share to vary in fact detracts from the benefits achievable by an investor. Finally, it is established that the genetic programming approach displays increased strength over the traditional simulation approach in solving for the optimal number of shares to include into a minimum variance portfolio.

4.4.6 Diversification Benefits

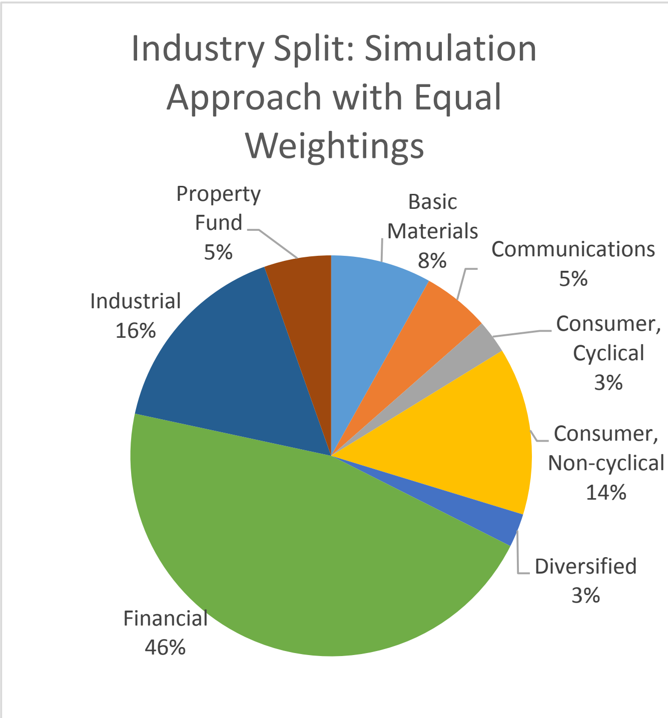
A final comparative analysis is carried out in this section. Below the diversification achieved across industries obtained by both mathematical approaches where the number of shares to be included in the portfolio was not pre-specified is examined. A table demonstrating the industry diversification obtaining by each approach is shown in Table 21.

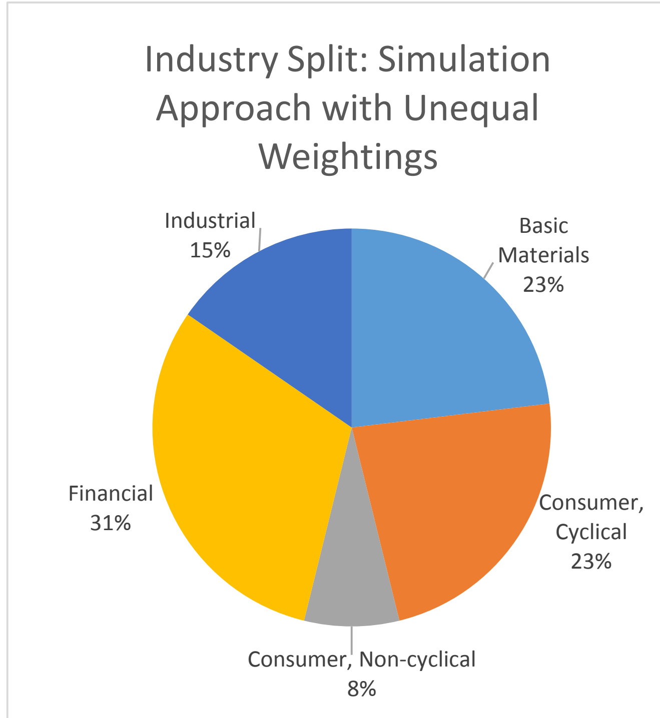
When comparing the approaches where traditional simulation was applied it is observed that a higher number of industries are included in the equally weighted share portfolio, albeit with a higher number of shares. In addition the percent of the population industries that are not included is less in the equally weighted example. As such the equally weighted traditional simulation approach can be said to be better diversified and more representative of the population than the unequally weighted traditional simulation approach (which is demonstrated through a lower portfolio variance). When comparing the equally weighted genetic programming approach to the unequally weighted genetic programming approach, it is observed that the equally weighted portfolio is diversified over more industries, however of the industries that were neglected the unequally weighted portfolio included the industries that carried a greater weighting of the population. The equally weighted genetic programming approach can however be said to be better diversified within industries and as such achieves a lower portfolio variance.

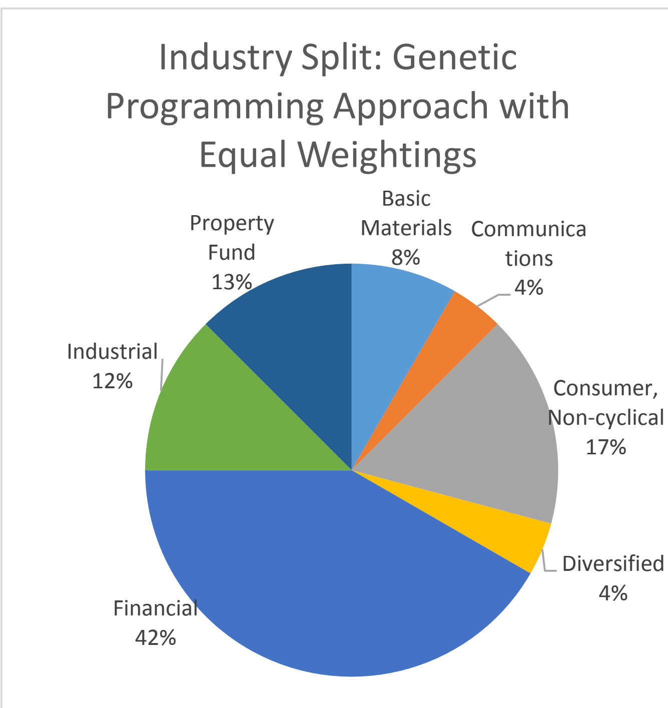
When weights remain equal, comparing the traditional simulation approach to the genetic programming approach shows that although the traditional simulation approach included a higher number of industries and was more representative of the population, the genetic programming approach, as previously argued, is better diversified within industries and thus achieved a lower portfolio variance. When the weightings on the share were allowed to vary, the genetic programming approach outperformed the traditional simulation approach. It utilised shares across more industries that the traditional simulation approach, was more representative of the population and also achieved a lower portfolio variance. A graphical summary of the diversification across industries achieved by each approach is displayed in Figure 48.

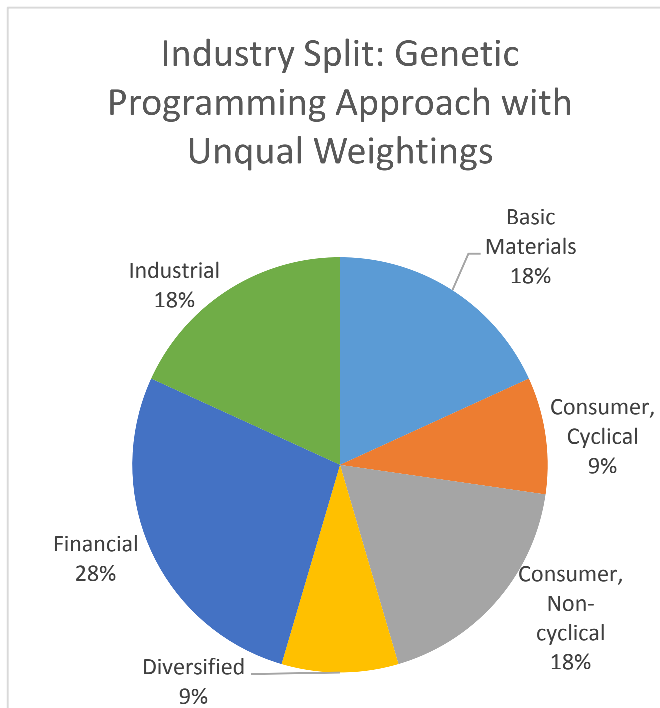
Table 21: *Diversification across Industries achieved by each Approach*

	Simulation Approach, Absent of Pre-Selected Number of Shares, Equal Weightings	Simulation Approach, Absent of Pre-Selected Number of Shares, Unequal Weightings	Genetic Programming Approach, Absent of Pre-Selected Number of Shares, Equal Weightings	Genetic Programming Approach, Absent of Pre-Selected Number of Shares, Unequal Weightings
Number of Industries Included	8	5	7	6
Per Cent of Population Industries not included	5.3%	16.57%	18.54%	12.08%
Number of Shares in respective Simulation	37	13	24	11
Minimum Variance Achieved in respective Simulation	0.3954	0.592	0.254	0.497


Industry Split: Simulation Approach with Equal Weightings
Property Fund 5%
Basic Materials 8%
Communications 5%
Consumer, Cyclical 3%
Consumer, Non-cyclical 14%
Diversified 3%
Financial 46%
Industrial 16%


Industry Split: Simulation Approach with Unequal Weightings
Industrial 15%
Basic Materials 23%
Consumer, Cyclical 23%
Consumer, Non-cyclical 8%
Financial 31%


Industry Split: Genetic Programming Approach with Equal Weightings
Property Fund 13%
Basic Materials 8%
Communica tions 4%
Consumer, Non-cyclical 17%
Diversified 4%
Financial 42%
Industrial 12%


Industry Split: Genetic Programming Approach with Unqual Weightings
Basic Materials 18%
Industrial 18%
Consumer, Cyclical 9%
Consumer, Non-cyclical 18%
Diversified 9%
Financial 28%

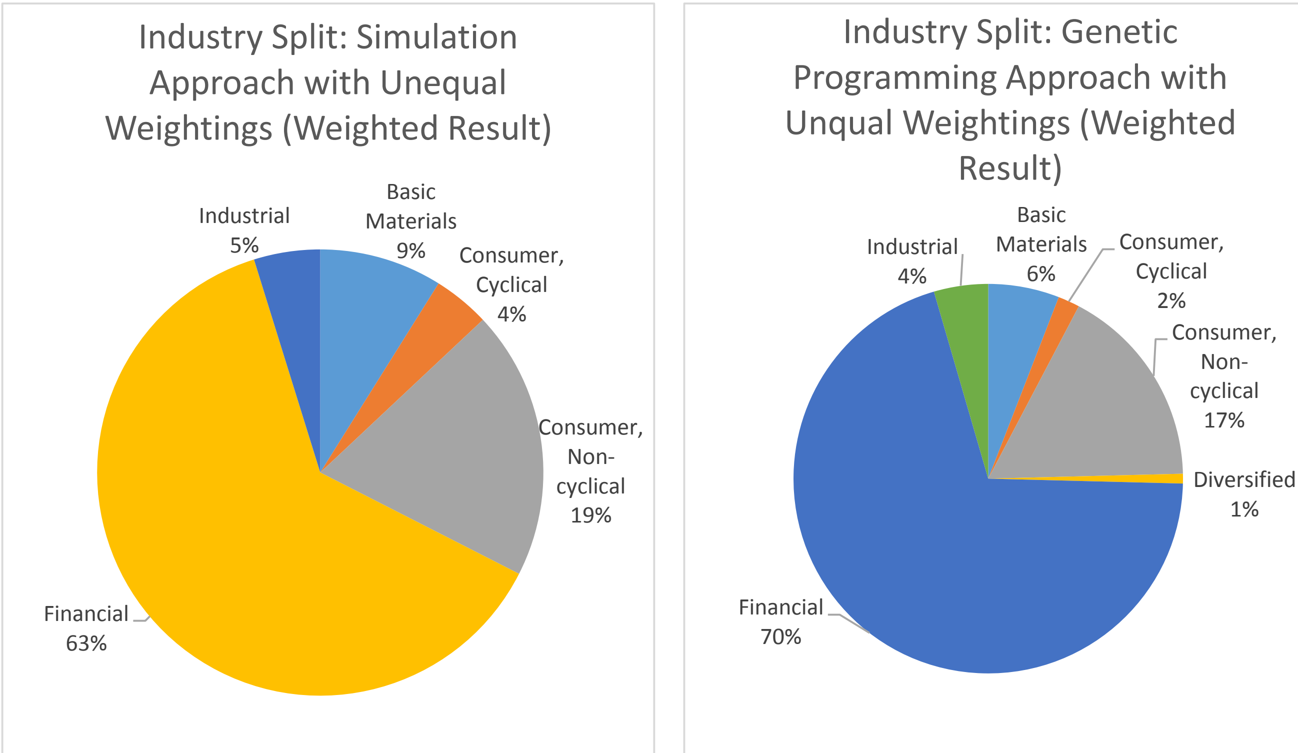


Figure 48: Graphical representation of the industry diversification achieved by each mathematical approach

4.4.7 Summary

This chapter aimed to provide a thorough analysis of the results achieved by this study as well as discussion as to the respective implications. This was accomplished with the intention of answering the two research questions. With the exception of one instance, all results were in agreement as to the first hypothesis – an optimally diversified portfolio should consist of between 10 and 30 shares (the results ranged between making use of between 10 and 24 shares, apart from the singular instance employing 37 shares which occurred in the unconstrained equally weighted simulation approach). Secondly, contrary to initial expectations, on no occasion did allowing the weights on each share to vary provide an investor with additional diversification benefits. This is believed to be primarily due to the increased quantum of potential possible portfolios leading to a smaller probability of the globally optimally diversified portfolio being discovered.

In response to the second research question, the results showed that genetic programming was the strongest mathematical approach with regards to solving for the optimally diversified portfolio. In all instances the genetic programming approach solved for a smaller minmum variance of the optimally diversified portfolio than that discovered by the corresponding traditional simulation approach, albeit by occasionally employing a larger number of shares. Particle swarm optimisation was the notable underperformer. This was hypothesised to occur due to the fact that when initialised, many particles did not satisfy the appropriate constraints. Finally, in the unconstrained portfolios, in all instances the portfolios described as optimally diversified outperformed the benchmark portfolio. Additionally, through an examination of the diversification across industries, all optimally diversified portfolios were found to display an adequate level of diversification and thus could be said to be representative of the population portfolio. The equally weighted simulation approach achieved a greater level of diversification across industries when compared to the unequally weighted simulation approach. This was validated once more when the results obtained from the equally weighted genetic programming approach were compared to the unequally weighted genetic programming approach. Furthermore, in both respective instances the genetic programming approach achieved a greater level of diversification across industries than the corresponding traditional simulation approaches. The study continues in Chapter 5 provides a conclusion to the dissertation accompanied by a section detailing the considerations and limitations of the current study, a summary of the results as well as an examination of the potential areas where future research could aim to contribute and improve upon current literature.

Chapter 5: Conclusion

Diversification continues to be a central tenant of investment theory due to its far reaching impacts, both theoretical and practical in nature. The act of diversification refers to spreading one's assets so as to minimise the effect of a singular event affecting all assets held, negatively. As such, diversification is regarded as a risk mitigation technique. It is imperative to solve for the optimal amount of diversification investors should obtain due to the reality that a marginal increase in the amount of diversification is accompanied by a marginal increase in associated costs.

With this in mind this dissertation aimed to answer two research questions. The first research question was: "what is the minimum number of shares to include into an equity portfolio, in order to achieve an optimal level of diversification in a South African environment?" The hypothesis was that, in line with previous literature, an optimally diversified portfolio should consist of between 10 to 30 shares. Additionally, the sub-question to the above was: "will allowing the weights on each share to vary uncover increased diversification benefits to an investor?" It was hypothesised that this would be the case due to the increased flexibility allowed to the simulation.

The second research question was: "of the three selected mathematical approaches, which demonstrates the most strength when solving for the optimally diversified portfolio in a South African environment?" It was hypothesised that the more computationally intensive and complex models of genetic programing and particle swarm optimisation would display increased strength in solving for the optimally diversified portfolio as opposed to the traditional simulation method.

The research questions above were answered through the application of the traditional simulation approach as well as the methodology of genetic programming and particle swarm optimisation. The study employed monthly share price data from the JSE with the aim being to solve for the optimal number of shares that should be included to a portfolio in order to achieve a minimum variance portfolio. The results achieved by each approach were thereafter compared to one another. The analysis was completed with various noted considerations and limitations to the methodology. These are discussed in section 5.1 below, which is followed by section 5.2 containing a summary of the results. The dissertation is thereafter concluded with an examination of possible future extensions to the current research.

5.1 Considerations and Limitations

When carrying out the methodology and analysis in the dissertation a number of considerations were made. These will be discussed below as well as the limitations facing the current study. The first consideration made was that transaction costs were taken into account. Elton and Gruber (1977) noted that if transaction costs for a security were strictly proportional to the size of the transaction, then the total amount of transaction costs would be independent of the number of securities in the portfolio. If, however, transaction costs increased as N, the number of shares in a portfolio, increased then an investor would prefer a portfolio consisting of a lower number of shares. In previous literature transaction costs were taken into account in three ways. The first was to include a fixed cost (monetary or secondly, statistical) to each additional share added to the portfolio while the third was to include a cost in proportion to the size of the amount placed on each share (with higher values incurring higher costs). The approach of this dissertation however, was rather than to include transaction costs directly, a termination criterion was required to be met in order for the processes to terminate (Magill & Constantinides, 1976; Fang, Lai & Wang, 2006). This termination criterion was discussed previously in the methodology section and with its inclusion, a statistical transaction cost is accounted for.

The second consideration highlighted was that the optimal number of shares to include in a diversified portfolio through testing on a data set may not necessarily have predictive capability for other data sets. This does not pose a problem to the research if the assumption is made that a South African investor is restricted to South African Equity. The third consideration made was that the liquidity of shares is taken into account, as previously discussed. This was completed by removing the shares in the population whose share price fell below 100 cents at any time over the observed period. This was completed as many shares that trade below 100 cents are viewed as ‘penny-stocks’ and as such their liquidity is exceptionally low and they trade with an excessively, irrational nature which results in extremely large daily moves followed by periods of extremely muted volatility. This was further expounded upon in the methodology section.

The fourth consideration is behavioural in nature - if two portfolios (or indexes) are perfectly diversified, an investor could still not be indifferent to the portfolios due to various behavioural factors towards the leadership of the various companies; the outlook that one portfolio is more likely to experience a negative black swan event (referring to an unforeseen event with an extremely low probability of occurring) as well as due to a plethora of others. Thus, it may be believed by the investor to be necessary to further diversify a 'completely' diversified portfolio. There is potential to find various behavioural factors to add into several models in order to analytically take them into account – although this is not the aim of this study.

The final consideration was that although optimal portfolio weightings may be found using the methodology, these weightings may be unachievable in reality as one cannot buy a portion of a share which may be necessary for the exact weightings to be achieved. This problem is not however considered extremely limiting as contracts for difference can be used to overcome this difficulty (Norman, 2010). This would allow one to achieve the exact weightings specified on each share to be included into the portfolio. This approach does not however take into account the potentially increased transaction costs of using contracts for difference to implement the same strategy, in place of using equity.

Furthermore, there are three limitations of this study, firstly taxes are not taken into account and secondly the costs with regards to constant rebalancing of the portfolio through time are not explicitly taken into account. The exclusion of the above two factors does not have a material impact on the primary aims of this dissertation. This is for the reason that the relationship between the number of shares and the risk in the portfolio is not substantially exposed to the costs of implementation. The aforementioned costs would affect the returns achieved by the portfolios which would be exhibited through inferior returns realised by the optimal portfolios. The inclusion of taxes could alter the shares selected by the simulations if returns are considered in addition to the relationship between the number of shares in the portfolio and the risk of the portfolio. For example, if dividends are taxed at a favourable rate compared to capital gains, shares with higher dividend pay-out ratios would be preferred over those with lower pay-out ratios. Similarly, if the costs of rebalancing the portfolio on a recurrent basis (daily, monthly, annually and the like) is considered (in order to maintain the respective weightings on each share), it will play an integral role in determining the returns achieved by the optimal portfolio. In this dissertation it is assumed that the weightings are set in the beginning of the period and held constant throughout the remainder of the period. There is potential for future studies to include the two above factors however, this does not pose an excessive threat to the value of the current dissertation as previous literature, in general as discussed in the literature review section, did not account for these limitations. As such this dissertation maintained a similar methodology to previous studies in order to ensure results attained are comparable.

The final limitation is that the portfolio is restricted to be comprised of only equity – holdings in cash, bonds and other alternative investments are not in the investors' potential universe. This is an intentional limitation included in this dissertation as it is necessary to restrict the portfolio to comprise of only equity in order to make the findings directly comparable to previous literature, the majority of which included the same limitation as previously elaborated. This chapter continues with a summary of the results of this dissertation and thereafter concludes with a discussion incorporating the aims that future research could focus on in order to continue to advance and make significant contributions to this division of financial literature.

5.2 Summary of Results

This dissertation aimed to answer the two research questions highlighted previously through a comparison of the results achieved by each of the three different approaches. This was accomplished through the use of four tests. The first test pre-specified the number of shares to be included into the portfolio and maintained an equal weighting on all shares utilised. This was completed using a number of shares ranging from 1 to 30. Thereafter the optimisation approaches were applied to a test where a number of shares were not specified beforehand, but equal weightings on all shares were to be maintained. The final two tests were similar to the first two tests but with a key difference – in these tests the weightings on each share in the portfolio were not required to be equal.

The completion of the above analysis allowed one to answer both research questions. The first research question targeted discovering the minimal number of shares to be included to obtain an optimally diversified portfolio in a South African environment. It was hypothesised that an optimal portfolio should consist of between 10 to 30 shares, as described by previous literature discussed in Chapter 2. The findings of this dissertation are in agreement with this and demonstrated that in all but one simulation, the optimal portfolio consisted of between 10 to 30 shares; the majority of results illustrated that a portfolio can be optimally diversified by utilising up to only 15 shares. It is demonstrated that one can however continue to add shares to the portfolio which is shown to further reduce the variance of a portfolio, however this becomes less and less significant as more shares are added to the portfolio - the positive impact of adding an additional share to a portfolio will further decrease if transaction costs are taken into account. These results provide an improvement upon those found by previous literature.

The objective of the sub-question was to evaluate whether allowing the weights on each share to vary would uncover increased diversification benefits. It was hypothesised that this would be the case; an equally weighted portfolio should construe a lower bound regarding the potential available diversification benefits available to an investor. However, evidence pointed towards the contrary; the portfolios that allowed for variable weights underperformed the equally weighted portfolios. This is likely due to the increased computing power and time necessary to thoroughly compute the enlarged search space. The search space is increased exponentially due to the vastly increased number of potential portfolios that can be constructed under the new constraints, as previously discussed.

The second research question centred on ascertaining which of the three mathematical approaches was the strongest in solving for the optimally diversified portfolio. It was hypothesised that the performance of the more computationally complex models would dominate the results achieved by the traditional simulation method. This was found to be true for the genetic programming approach - the genetic programming approach was recognised to be stronger than the traditional simulation approach. The genetic programming approach solved for portfolios yielding lower portfolio variances than the traditional simulation approach in all of the tests where the number of shares to be included in the portfolio was incrementally increased. Furthermore, where the number of shares included to the portfolio was not pre-specified the genetic programming approach solved for portfolios yielding a lower variance while also utilising fewer shares than the portfolios computed using the traditional simulation methodology. Conversely, the notable underperformer was the particle swarm optimisation approach, clearly showing its weaknesses. This approach performed significantly worse than both the genetic programming and the traditional simulation approaches. Reasons for the underperformance of the particle swarm optimisation approach were outlined in the results section.

Finally, although not the primary aim of the dissertation, the return profiles of the portfolio's identified as optimally diversified using the unconstrained equally weighted simulation approach and the unconstrained equally weighted genetic programing approach were analysed and compared. This established that in both cases the portfolios deemed as optimally diversified outperformed the benchmark portfolio. Additionally, through an examination of the diversification achieved across industries, all optimally diversified portfolios were found to display a sufficient level of diversification and thus could were deemed to be representative of the population portfolio. The equally weighted simulation approach achieved a greater level of diversification across industries when compared to the unequally weighted simulation approach. This was validated once more when the results obtained from the equally weighted genetic programming approach were compared to the results achieved by the unequally weighted genetic programming approach. Furthermore, in both respective instances the genetic programming approach achieved a greater level of diversification across industries than the corresponding traditional simulation approaches.

The dissertation now concludes with a discussion summarising the potential aims for future research.

5.3 Aims for Future Research

Previous literature thus far, with regards to solving for the optimally diversified portfolio, has primarily focused on the utilisation of a single asset class – equity. This holds true in the analysis of this dissertation too. There is thus the potential for future research to include various additional asset classes into the analysis with the aim of being able to further diversify the wealth of an investors' portfolio. There are numerous asset classes that could be included such as fixed income, currencies, property and the like. In addition to the inclusion of a larger number of asset classes, there is potential to include assets that range across a number of geographic locations. The above inclusion could further enhance the diversification benefits available to an investor. This dissertation focused solely on South African equity intentionally in order to ensure the results achieved were directly comparable to those achieved by previously discussed literature. In doing so, one gains a like for like comparison enabling a fair discernment in answering the proposed two research questions.

Future research could also aim to include the effect that taxes may have on an investors' portfolio. For example, if dividends are taxed at a lower rate than capital gains, with respect to equities, the diversification analysis should aim to optimally diversify an investors' portfolio while favouring high dividend paying shares over shares that obtain returns primarily through capital appreciation, as previously described. Similarly, the costs of constantly rebalancing a portfolio in order to obtain the targeted weightings could also be brought into account. The above two future research aims assume that the study will be focused on both the relationship between the number of shares in a portfolio and the risk inherent in that portfolio as well as the returns achieved by the specified optimal portfolio. As discussed previously, this is not the primary concentration of this dissertation albeit that an analysis of the returns achieved by the unconstrained equally weighted portfolios was completed and compared to a benchmark portfolio.

Finally, future research focusing on applying artificial intelligence to minimise the variance of a portfolio could include incorporating various other artificial intelligence methodologies such as simulated annealing, fuzzy logic, reinforcement learning as well as involving other paradigms of thinking, such as cybernetics. The simulated annealing algorithms were initially inspired by the annealing process in metal work. In this process metal is heated and then cooled with the aim to alter its physical properties. The algorithm works by ‘heating’ the data which allows for various solutions to be found in different areas of the search space. The algorithm is then ‘cooled’ which involves the process of gradually decreasing the probability of a less optimal solution being accepted in place of the best solution found at any time. This approach is relatively simple and aims to solve for an appropriate solution, although it may not be the globally optimal solution.

Fuzzy logic is a mathematical approach that stands in contrast to Boolean logic. Boolean logic refers to the fact that variables may only have values of 0 or 1 whereas fuzzy logic allows values of variables to be any value between 0 and 1. This could be applied by allowing values, such as the weightings on shares, to represent a range of continuous values rather than a simple discreet weighting. Variables are thus based on a ‘degree of truth’ as opposed to the ‘true’ or ‘false’ variables used by Boolean simulations. Fuzzy logic was initially founded with the aim of assisting computers to understand natural language and is said to be closer to the way in which the human brain works – data is aggregated from a number of partial truths to reach new, higher levels of truth.

Reinforcement learning refers to a type of machine learning that is able to automatically determine the ideal behaviour within a specific context, aiming to maximise performance. The behavioural technique revolves around the simple framework of reward based learning. The algorithm is rewarded for behaving in a certain manner. This leads the algorithm to learn this behaviour as a positive one and as such when a similar situation is encountered again the algorithm should behave in a similar way to the one for which it was previously rewarded. Cybernetics grew from a desire to understand and build goal orientated systems. Cybernetics is applicable when a system incorporates a closed signalling loop. This is understood as a system wherein an action generates a change in the environment which is then reflected in the environment. The change in the environment is thereafter learned from and another change is then introduced to the environment with the aim of achieving an optimal state. There is potential for each of the above approaches to be applied to the problem of achieving an optimally diversified portfolio. Finally, future literature can aim to include a predictive component in the analysis which would allow an investor to forecast which shares should be included in order to generate the minimum variance portfolio in the future.

In conclusion, it was found that optimal portfolios should consist of between 10 to 30 shares, in line with previous literature. This outcome was emphasised by the fact that the results of both the traditional simulation and genetic programming approaches were in agreement with this hypothesis. Secondly, contrary to expectations, it was demonstrated that allowing the weights on each share to vary decreased the performance of each approach, limiting the diversification benefits available to an investor. Finally, the genetic programming approach was revealed as the strongest approach in solving for the optimally diversified portfolio, followed by the traditional simulation approach and thereafter by the significantly underperforming particle swarm optimisation approach. Future research can aim to incorporate multiple asset classes across numerous geographies. Additional costs such as taxes and rebalancing costs could be incorporated into the analysis when an optimally diversified portfolio is considered in conjunction with the objective of maximising returns. Finally, applying additional advanced mathematical approaches escalates the possibility of discovering further diversification benefits available to an investor. Through a focus on the above mentioned aims, future research would be enabled to contribute to the current literature available on the topic of diversification while making significant strides forward in potentially uncovering previously unknown superior diversification benefits available to investors.

References

Abraham, A., & Nath, B. (2000). Hybrid Intelligent Systems Design: A Review of a Decade of Research. *IEEE Transactions in Systems, Man & Cybernetics (Part C) August, 3*(1), 1-37.

Akgiray, V. (1989). Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts. *Journal of Business*, 55-80.

Allen, F., & Karjalainen, R. (1993). Using Genetic Algorithms to Find Technical Trading Rules. *Journal of Financial Economics, 51*(2), 20-93.

Álvarez-Díaz, M., & Álvarez, A. (2005). Genetic Multi-Model Composite Forecast for Non-Linear Prediction of Exchange Rates. *Empirical Economics, 30*(3), 643-663.

Ansoff, H. I. (1957). Strategies for Diversification. *Harvard Business Review, 35*(5), 113-124.

Arrow, K. J. (1963). Uncertaintity and the Welfare Economics of Medical Care. *The American Economic Review*, 941-973.

Baker, K. H., & Nofsinger, J. R. (2010). *Behavioral Finance: Investors, Corporations and Markets* (Illustrated ed.). New York: John Wiley & Sons.

Battiti, R., & Brunato, M. (2014). *The LION Way: Machine Learning plus Intelligent Optimization* (Vol. 1). Italy: University of Trento.

Bauer, H. (1996). *Probability Theory*. Berlin: Walter de Gruyter.

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica, 22*(1), 23-36.

Blumer, A., Ehrenfeucht, A., Haussler, D., & Warmuth, M. K. (1987). Ocam's Razor. *Informational Processing Letters, 24*(1), 377-380.

Bradfield, D. (1993). The Effect of Diversification on Risk in Portfolio Construction. *Department of Statistical Sciences, University of Cape Town*, Unpublished Working Paper.

Burrell, P. R., & Forlarin, B. O. (1997). The Impact of Neural Networks in Finance. *Neural Computing and Applications*, 193-200.

Chen, Y., Yang, B., & & Abraham, A. (2007). Flexible Neural Trees Ensemble for Stock Index Modeling. *Neurocomputing, 70*(4-6), 697-703.

Chen, Y., Yang, B., & Dong, J. (2004). Nonlinear System Modelling Via Optimal Design of Neural Trees. *International Journal of Neural Systems, 14*(2), 1-13.

Chen, Y., Yang, B., Dong, J., & Abraham, A. (2005). Time-Series Forecasting using a Flexible Neural Tree Model. *Information Sciences, 124*(3-4), 219-235.

Clerc, M. (2006). *Particle Swarm Optimization.* London: ISTE.

Comment, R., & Jarrell, G. A. (1995). Corporate Focus and Stock Returns. *Journal of Financial Economics, 37*(1), 67-87.

Corbbett, J. (2013, October 30). *Moneyweb: Online Stock Broker Fees Compared.* Retrieved from Moneyweb: http://www.moneyweb.co.za/archive/online-stock-broker-fees-compared-2/

Daniel, K., & Titman, S. (1999). Market Efficiency in an Irrational World. *Financial Analysts Journal, 55*(6), 28-40.

Darwin, C. (1859). *On The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life.* London: W. Clowes and Sons.

Domingos, P. (1999). The Role of Occam's Razor in Knowledge Discovery. *Department of Computer Science and Engineering*, 1-19.

Eberhart, R. C., & Kennedy, J. (1995). A New Optimizer Using Particle Swarm Theory. *In Proceedings of the Sixth International Symposium on Micro Machine and Human Science. 1*, pp. 39-43. Piscataway: IEEE.

Eberhart, R. C., & Shi, Y. (2001). Particle Swarm Optimization: Developments, Applications and Resources. *Evolutionary Computation Proceedings. 1*, pp. 81-86. Piscataway: IEEE.

Elton, E. J., & Gruber, M. J. (1977). Risk Reduction and Portfolio Size: An Analytical Solution. *The Journal of Business, 50*(4), 414-437.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzman, W. N. (2009). *Modern Portfolio Theory and Investment Analysis.* New Jersey: John Wiley & Sons.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2009). *Modern Portfolio Theory and Investment Analysis* (8th ed.). New York: John Wiley & Sons.

Evans, J. L., & Archer, S. H. (1968). Diversification and the Reduction of Dispersion: An Empirical Analysis. *The Journal of Finance, 23*(5), 761-767.

Fang, Y., Lai, K. K., & Wang, S. Y. (2006). Portfolio Rebalancing Model with Transaction Costs based on Fuzzy Decision Theory. *European Journal of Operational Research, 175*(2), 879-893.

Fernández, A., & Gómez, S. (2007). Portfolio Selection using Neural Networks. *Computers and Operations Research, 34*(4), 1177-1191.

Fischer, D. E., & Jordan, R. J. (1995). *Sercurity Analysis and Portfolio Management* (6th ed.). New Jersey: Prentice Hall International.

Francis, J. C. (1986). *Investments: Analysis and Management* (4th ed.). New York: McGraw-Hill.

Freitas, F. D., Souza, A. F., & de Almeida, A. R. (2009). Prediction-Based Portfolio Optimization Model using Neural Networks. *Neurocomputing, 72*(10), 2155-2170.

French, K. R., & Poterba, J. M. (1993). Investor Diversification and International Equity Markets. In R. H. Thaler, *Advances in Behavioural Finance* (pp. 383-384). New York: Russell Sage Foundation.

Gelman, A., & Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

German, S., Bienenstock, E., & Doursat, R. (1992). Neural Networks and the Bias/Variance Dilema. *Neural Computation, 4*(1), 1-58.

Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition.* Cambridge: Cambridge University Press.

Gibson, J. P., & Gibson, T. R. (2009). *Natural Selection.* New York: Infobase Publishing.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* New York: Springer.

Goonatilake, S., & Treleaven, P. C. (1995). *Intelligent Systems for Finance and Business.* New York: John Wiley & Sons, Inc.

Gregoriou, G. N. (2008). *Encyclopedia of Alternative Investments.* New York: CRC Press.

Gup, B. E. (1983). *The Basics of Investing* (2 ed.). New York: John Wiley & Sons.

Hagan, M., Demuth, H., Beale, M., & De Jesus, O. (1996). *Neural Network Design.* Boston: PWS Publishing Company.

Hall, R. E., & Lieberman, M. (2008). *Microeconomics: Principles and Applications* (Vol. 4E). Ohio: Thomson Learning.

Holland, J. H. (1962). Outline for a logical theory of adaptive systems. *Journal of the ACM, 9*(3), 297-314.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* Oxford, England: U Michigan Press.

Huang, W., Lai, K. K., Nakamori, Y., Wang, S., & Lean, Y. (2007). Neural Networks in FInance and Economics Forecasting. *International Journal of Information Technology & Decision Making, 6*(1), 113-140.

Jacob, N. L. (1974). A Limited-Diversification Portfolio Selection Model for the Small Investor. *The Journal of Finance, 29*(3), 837-875.

Jin, N., Tsang, E., & Li, J. (2009). A Constraint-Guided Method with Evolutionary Algorithms for Economic Problems. *Applied Soft Computing, 9*(3), 924-935.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation Bias in Sequential Information Search after Preliminary Decisions: An expansion of Dissonance Theoretical Research on Selective Exposure to Information. *Journal of Personality and Social Psychology, 80*(4), 557-571.

JSE Equities Rules. (2016, January 18). Retrieved from Johannesburg Stock Exchange: https://www.jse.co.za/services/market-regulation

Kaastra, I., & Boyd, M. (1996). Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing, 10*, 215-236.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kahneman, D., & Tversky, A. (1984). Choices, Values, and Frames. *American Psychologist, 39*(4), 341-350.

Kendall, G., & Su, Y. (2005). A Particle Swarm Optimisation Approach in the Construction of Optimal Risky Portfolios. *Artificial Intelligence and Applications, 453*, 140-145.

Kennedy, J., & Eberhart, R. C. (1995). Particle Swarm Optimization. *In Proceedings of the IEEE Congress on Evolutionary Computation* (pp. 1942-1948). Piscataway: IEEE.

King, M. A., & Leape, J. I. (1998). Wealth and Portfolio Composition: Theory and Evidence. *Journal of Public Economics, 69*(2), 155-193.

Knight, F. H. (2012). *Risk, Uncertaintity and Profit.* New York: Dover Publications.

Koza, J. R. (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* Cambridge: The MIT Press.

Krishnaswamy, C. R., Gilbert, E. W., & Pashley, M. M. (2000). Neural Network Applications in Finance: A Practical Introduction. *Financial Practice and Education*, 75-84.

Lang, L. H., & Stultz, R. M. (1994). Tobin's q, Corporate Diversification, and FIrm Performance. *Journal of Political Economy, 102*(6), 1248-1280.

Liang, J. J., & Suganthan, P. N. (2005). Dynamic Multiswarm Particle Swarm Optimizer. *IEEE Swarm Intelligemce Symposium (SIS)* (pp. 124-129). Piscataway: IEEE.

Liu, F., Ng, G. S., & Quek, C. (2007). RLDDE: A Novel Reinforcement Learning-Based Dimension and Delay Estimator for Neural Networks in Time Series Prediction. *Neurocomputing, 70*(7), 1331-1341.

Magill, M. J., & Constantinides, G. M. (1976). Portfolio Selection with Transaction Costs. *Journal of Economic Theory, 13*(2), 245-263.

Malkiel, B. G. (2015). *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing* (11 ed.). New York: W. W. Norton.

Markowitz, H. (1952a). The Utility of Wealth. *The Journal of Political Economy, 60*(2), 151-158.

Markowitz, H. (1952b). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91.

Markowitz, H. (1991). Foundations of Portfolio Theory. *The Journal of Finance, 46*(2), 469-477.

Markowitz, H. (1999). The Early History of Portfolio Theory: 1600-1960. *Financial Analysts Journal*, 5-16.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments.* New York: John Wiley & Sons.

Mayshar, J. (1979). Transaction Cost in a Model of Capital Market Equilibrium. *Journal of Political Economy, 87*, 673-700.

McCarthy, T. J., & McCluskey, S. (2009). A Particle Swarm Optimisation Approach to Reinforced Concrete Beam Design Acording to AS3600. *Proceedings of the First International Conference* *on Soft Computing Technology in Civil, Structural and Environmental Engineering* (pp. 1-14). Madeira, Portugal: Civil-Comp Press.

McLeish, D. L. (2011). *Monte Carlo Simulation and Finance.* Hoboken: John Wiley & Sons.

Miles, S., & Smith, B. (2010). Testing the Adaptive Efficiency of U.S. Stock Markets: A Genetic Programming Approach. *Journal of Business & Economics Research, 8*(11), 87-112.

Miller, F. P., Vandome, A. F., & McBrewster, J. (2009). *Confirmation Bias.* Saarbrücken: VDM Publishing.

Miller, J. H. (1986). *A Genetic Model of Adaptive Economic Behaviour.* University of Michigan: Mimeo.

Millonas, M. (1994). Swarms, Phase Transitions, and Collective Intellignece. In C. G. Langton, *Artificial Life III* (pp. 417–445). Massachusetts: Addison Wesley.

Moody, J., & Saffell, M. (2001). Learning to Trade via Direct Reinforcement. *IEEE Transactions on Neural Networks, 12*(4), 875-889.

Morck, R., Shleifer, A., & Vishny, R. W. (1990). Do Managerial Objectives Drive Bad Acquisitions? *Journal of Finance, 45*(1), 31-48.

Neu-Ner, M. A., & Firer, C. (1997). The Benefits of Diversification on the JSE. *Investment Analysts Journal, 46*, 45-59.

Nevid, J. S. (2013). *Psychology: Concepts and Applications* (4 ed.). California: Wadsworth Cengage Learning.

Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology, 2*(2), 175-220.

Norman, D. J. (2010). *CFDs - The Definitive Guide to Contracts for Difference.* Hampshire: Harriman House Limited.

Pai, P. F., & Lin, C. S. (2005). A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting. *Omega, 33*(6), 497-505.

Pearson, K. (1925). "James Bernoulli's Theorem". *Biometrika*, 201-210.

Poli, R. (2008). Analysis of the Publications on the Applications of Particle Swarm Optimisation. *Journal of Artificial Evolution and Applications, 3*, 1-10.

Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle Swarm Optimization: An Overview. *Swarm Intelligence, 1*, 33-57.

Reilly, F. K. (1985). *Investment Analysis and Portfolio Management* (2 ed.). San Francisco: Dryden Press.

Roehner, B. M. (1995). *Theory of Markets: Trade and Space-time Patterns of Price Fluctuations, A Study in Analytical Economics.* New York: Springer-Verlag.

Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica, 20*, 431-449.

Rubinstein , M. (2002). Markowitz's "Portfolio Selection': A Fifty-Year Retrospective. *The Journal of Finance, 57*(3), 1041-1045.

Rumelt, R. P. (1982). Diversification Strategy and Profitability. *Strategic Management Journal, 3*(4), 359-369.

Shachmurove, Y. (2002). *Applying Artificial Neural Networks to Business, Economics and Finance.* University of Pennsylvania: Center for Analytic Research in Economics and Social Sciences.

Shakespeare, W. (1905). *The Merchant of Venice.* Hayes Barton Press.

Shiller, R. J. (1990). Speculative Prices and Popular Models. *The Journal of Economic Perspectives*, 55-65.

Singh, S. (2003). *Advanced Sampling Theory With Applications* (Vol. 1). Boston: Kluwer Academic Publishers.

Solnik, B. H. (1974). Why not Diversify Internationally rather Domestically? *Financial Analyst Journal*, 48-54.

Statman, M. (1987). How Many Stocks Make a Diversified Portfolio? *The Journal of Financial and Quantative Analysis, 22*(3), 353-363.

Stevenson, A. (2010). *Oxford Dictionary of English.* Oxford: Oxford University Press.

Stevenson, R. A., & Jennings, E. H. (1984). *Fundamentals of Investments* (3 ed.). San Francisco: West Publishing Co.

Suganthan, P. N. (1999). Particle Swarm Optimiser with Neighbourhood Operator. *IEEE International Congress on Evolutionary Computation (CEC)* (pp. 1958-1962). Piscataway: IEEE.

Swensen, D. F. (2009). *Pioneering Portfolio Management: An Unconventional Approach to Institutional Investment.* New York: Free Press.

Thawornwong, S., & Enke, D. (2004). The Adaptive Selection of Financial and Economic Variables for use with Artificial Neural Networks. *Neurocomputing, 56*, 205-232.

Thomopoulos, N. T. (2012). *Essentials of Monte Carlo Simulation: Statistical Methods for Building Simulation Models.* Chicago: Springer.

Tornay, S. C. (1938). *Ockham: Studies and Selections.* La Salle: Open Court.

Tversky, A., & Fox, C. R. (1995). Weighing Risk and Uncertainty. *Psychological Review, 102*(2), 269-283.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for Judging Frequency and Probability. *Cognitive Psychology, 5*(2), 207-232.

Tversky, A., & Kahneman, D. (1990). Judgment Under Uncertainty: Heuristics and Biases. In P. K. Moser, *Rationality in Action: Contemporary Approaches* (pp. 171-188). Cambridge: Cambridge University Press.

Wagner, N., & Brauer, J. (2007). Using Dynamic Forecasting Genetic Programming (DFGP) to Forecast United States Gross Domestic Product (US GDP) with Military Expenditure as an Explanatory Variable. *Defence and Peace Economics, 18*(5), 451-466.

Wilcox, C. (2010). *Groupthink: An Impediment to Success.* Bloomington: Xlibris Corporation.

Wilson, E. O. (1975). *Sociobiology: The New Synthesis.* Cambridge, M.A: Belknap Peress.

Zhu, H., Wang, Y., Wang, K., & Chen, Y. (2011). Particle Swarm Optimization (PSO) for the Constrained Portfolio Optimization Problem. *Expert Systems with Applications, 38*(8), 10161-10169.

Zhu, Y., & Zhou, G. (2009). Technical Analysis: An Asset Allocation Perspective on the Use of Moving Averages. *Journal of Financial Economics, 92*(3), 519-544.

Appendix 1

Procedures in Excel

Below is a step by step guide revealing how the data was retrieved and utilised in the current dissertation:

As described previously the monthly closing prices of the respective shares were imported from the Bloomberg data base. Thereafter the adjustments previously mentioned were performed on the data and the return series as well as the standard deviation for each share was calculated respectively, an example of the output is illustrated in Table 22 below:

Table 22: *The 6 Month Return Series with Respective Standard Deviations for 4 Hypothetical Shares*

Returns:				
	Share A	Share B	Share C	Share D
Month 1	2.00%	4.00%	0.50%	-2.00%
Month 2	1.50%	3.00%	1.00%	-1.00%
Month 3	-0.50%	2.00%	-0.75%	2.00%
Month 4	0.25%	1.00%	-0.25%	0.45%
Month 5	1.00%	2.00%	0.00%	-0.10%
Month 6	0.50%	1.00%	1.00%	1.00%

Standard Deviations	0.009	0.012	0.007	0.014

Note: The table above shows the return matrix of a hypothetical portfolio consisting of 4 shares over the period of 6 months followed by the standard deviation of each shares returns respectively

In order to calculate the portfolio variance the correlation matrix was required to be evaluated. This was accomplished by utilising the default excel functioning as follows: under the ‘data’ tab one should click on ‘data analysis’. Thereafter one can select the series of returns as the input range and output the correlation matrix into a new sheet. This will give one half of the correlation matrix. The output of the calculated correlation matrix is illustrated in the table below:

Table 23: *Correlation matrix calculated from the return series data in Table 22*

	Column 1	*Column 2*	*Column 3*	*Column 4*
Column 1	1			
Column 2	0.752241	1		
Column 3	0.667651	0.241943	1	
Column 4	-0.97481	-0.77987	-0.53489	1

As the excel based portfolio variance calculation requires a full correlation matrix, the table above needed to be filled completely. This can be achieved by copying the data from Table 23, selecting ‘paste special’, and the option ‘transpose’ and then pasting this output data to a separate section in the sheet. Thereafter the newly pasted data can be selected, copied and then pasted into the original correlation table (Table 23) using the options ‘paste special’ and ‘skip blanks’ when selecting the first data point in the correlation matrix. This will yield the result of a full correlation matrix, as seen below:

Table 24: *The full correlation matrix of the share returns shown in Table 22*

	Column 1	*Column 2*	*Column 3*	*Column 4*
Column 1	1	0.752241	0.667651	-0.97481
Column 2	0.752241	1	0.241943	-0.77987
Column 3	0.667651	0.241943	1	-0.53489
Column 4	-0.97481	-0.77987	-0.53489	1

Thereafter, the section in Table 22 showing the series of standard deviations for each share return respectively should be transposed to be vertical. One can then create the table below:

Table 25: *The table necessary to compute the portfolio variance*

Used	**Standard Deviations**	**Weight**	**Used*S.D.*Weight**
1	0.009	0.5	0.005
1	0.012	0.5	0.006
0	0.007	0.5	0
0	0.014	0.5	0

In the table above the standard deviations are the transposed matrix of the standard deviations for the shares previously reported in Table 22. The column headed 'Used' contains a binary field (that will contain either 1 or 0, indicating 1 for when a share is included in the portfolio and 0 for when a share is not included into the portfolio). The column entitled 'weight' demonstrates the weight to be allocated to each share in the portfolio. The derivation of this figure will be explained below. The final column holds the horizontal product, being whether or not a share is included into a portfolio (binary field) multiplied by the standard deviation of the respective share multiplied by the weight assigned to the specific share. The portfolio variance is then calculated by the following formula:

=MMULT(MMULT(TRANSPOSE(T7:T10),K4:N7),T7:T10) *(15)*

Cells T7:T10 represent the product cell (the final column in Table 25) while cells K4:N7 represent the full correlation matrix shown in Table 24. It is important to note that as this a formula involving matrix multiplication the keys ‘Ctrl’, ‘Shift’ and ‘Enter’ must be pressed simultaneously in order to calculate the portfolio variance. The final answer (using the data above) should be equivalent to the table below:

Table 26: *Displaying the final portfolio variance for the hypothetical portfolio exhibited previously*

Portfolio Variance	0.000106134

The optimisation process for the equally weighted test is run under the following constraints:

Table 27: *The constraints utilised on an Equally Weighted Portfolio*

Constraints:	Number of shares used	2
	Weights	0.5
	Number of shares to be used	2

The ‘number of shares used’ contains a ‘countif’ formula that calculates how many shares are used (how many shares possess a ‘1’ value in the ‘Used’ column). The weight is then calculated as 1 divided by the number of shares included in the portfolio. The final constraint assists the optimisation process by restricting results to only be successful trials when the number of shares used equals the number of shares to be used (which is a measure of the number of shares the user desires the optimal portfolio to consist of). The number of shares to be used is increased from 1 to 30 as described previously and should always be equal to the number of cells actually used. Test one then optimises the portfolio variance by adjusting the binary cells in the ‘Used’ column until an optimum solution is found that meets all constraints or until a termination criterion is met as described previously.

The optimum overall portfolio is then found (test 2) by removing the first restriction (that the number of shares included in a portfolio equal the number of shares requested by the investor) and allowing the optimisation to run with the constraint that a specific number of shares is deemed necessary by the optimisation process. The process is however still subject to the other termination criteria while retaining equal weightings on all shares included in the portfolio.

Test 3 is then computed under the following constraints:

Table 28: *Showing the constraints needed to compute the optimal variance for an unequally weighted portfolio*

Constraints:	Number of shares used	2
	Sum weights :	1
	Number of shares to be used	2
	Number of shares with a 0 weighting by used	0

The ‘number of shares used’ as well as the ‘number of shares to be used’ cells serve the same purpose as those in test 1 and 2 respectively. The ‘sum weights’ cell is a ‘sumifs’ function that is equal to the sum of all the weights under the ‘weights’ column where the binary ‘used’ cell is equal to 1. This constraint should always be equal to 1 (meaning that the full portion of an investors funds are always invested into the portfolio, without short selling or leverage). The sum of the number of shares with a 0 weighting but where the binary used cell is equal to 1 is the final constraint. This constraint should always be equal to 0 as to not affect the count of how many shares were included into the portfolio as a share included into a portfolio with a weight of 0 should not be counted as being included into the portfolio.

Test 4 is then computed by removing the same constraints as were removed from test 1 while keeping the remaining constraints applied to test 3. The final test will then optimise the variance of the portfolio until the termination criterion is met, as described previously. Furthermore, over all the tests the constraint that the weights on each of the shares should always range between 0 and 1 and add up to 1 should always remain true.

Appendix 2

The table below shows the results obtained in the first simulation, when the minimum variance portfolio for each pre-specified number of shares (ranging from 1 to 30) was calculated using the traditional simulation approach. As discussed previously, the portfolio is restricted to be equally weighted and for the sum of the weights on each share to always equal 1.

Table 29: *Simulation approach solving for minimum variance portfolio for each number of shares (1-30) using an equally weighted portfolio, restricting the sum of the weights to be equal to 1*

Number of Shares in Portfolio	Minimum Monthly Variance (%)	Minimum Variance Annualised (%)
1	1.186	3.773
2	0.806	3.109
3	0.745	2.990
4	0.522	2.502
5	0.519	2.495
6	0.632	2.754
7	0.451	2.327
8	0.487	2.418
9	0.477	2.393
10	0.307	1.920
11	0.390	2.164
12	0.389	2.160
13	0.441	2.300
14	0.416	2.233
15	0.421	2.248
16	0.346	2.039
17	0.503	2.458
18	0.358	2.071
19	0.566	2.606
20	0.445	2.310
21	0.336	2.009
22	0.376	2.123
23	0.266	1.786
24	0.415	2.231
25	0.27	1.801
26	0.489	2.423
27	0.343	2.030
28	0.320	1.958
29	0.335	2.005
30	0.351	2.053

The table below demonstrates the results obtained in the second simulation. Here the minimum variance portfolio for each pre-specified number of shares (ranging from 1 to 30) was calculated using the traditional simulation approach. This differs from the above table as in this simulation the weights on each share were allowed to vary however the sum of the weights was still required to be equal to 1.

Table 30: *Simulation approach solving for the minimum variance portfolio for each number of shares (1-30) using a portfolio that is not required to be equally weighted, while still restricting the sum of the weights to be equal to 1*

Number of Shares in Portfolio	Minimum Monthly Variance (%)	Minimum Variance Annualised (%)
1	1.186	3.773
2	1.076	3.594
3	1.640	4.436
4	2.314	5.269
5	1.929	4.812
6	1.670	4.476
7	1.987	4.883
8	1.600	4.382
9	0.777	3.053
10	1.879	4.749
11	0.792	3.084
12	1.646	4.444
13	1.672	4.480
14	1.643	4.440
15	0.744	2.988
16	1.713	4.534
17	1.592	4.370
18	0.955	3.385
19	1.700	4.516
20	1.028	3.513
21	1.011	3.484
22	1.356	4.034
23	0.996	3.457
24	1.517	4.267
25	1.682	4.493
26	1.658	4.461
27	1.561	4.328
28	1.844	4.705
29	1.855	4.718
30	2.116	5.039

The genetic programming approach was then applied to the data. The minimum variance portfolio was once again calculated for each pre-specified number of shares utilising an equally weighted portfolio. The table below displays the results obtained from the utilisation of this approach.

Table 31: *Genetic Programming approach solving for the minimum variance portfolio for each number of shares (1-30) using an equally weighted portfolio, restricting the sum of the weights to be equal to 1*

Number of Shares in Portfolio	Minimum Monthly Variance (%)	Minimum Variance Annualised (%)
1	1.186	3.773
2	0.806	3.109
3	0.682	2.861
4	0.582	2.643
5	0.452	2.328
6	0.394	2.175
7	0.375	2.121
8	0.364	2.089
9	0.347	2.042
10	0.344	2.032
11	0.303	1.906
12	0.312	1.934
13	0.290	1.864
14	0.266	1.788
15	0.288	1.859
16	0.350	2.050
17	0.282	1.839
18	0.262	1.772
19	0.267	1.790
20	0.324	1.971
21	0.350	2.049
22	0.335	2.006
23	0.276	1.819
24	0.298	1.891
25	0.282	1.839
26	0.282	1.839
27	0.251	1.737
28	0.270	1.799
29	0.266	1.787
30	0.249	1.727

The results below are obtained from applying the genetic programming approach to the data, however in this instance the weights on each respective share were allowed to vary. The weights on the shares were however still restricted to sum to 1.

Table 32: *Genetic Programming approach solving for the minimum variance portfolio for each number of shares (1-30) without requiring equal weightings on each share while still restricting the sum of the weights to be equal to 1*

Number of Shares in Portfolio	Minimum Monthly Variance (%)	Minimum Variance Annualised (%)
1	1.186	3.773
2	2.33	5.060
3	0.794	3.087
4	1.090	4.659
5	1.990	4.898
6	0.780	3.060
7	1.610	4.851
8	1.060	4.524
9	1.120	4.533
10	1.550	4.178
11	1.040	3.471
12	0.747	2.993
13	0.791	3.081
14	1.010	3.465
15	0.679	2.854
16	0.801	3.100
17	1.500	3.873
18	0.896	3.278
19	0.711	2.920
20	1.990	4.381
21	1.930	4.372
22	0.777	3.054
23	1.580	3.886
24	1.530	4.030
25	0.947	3.371
26	1.530	3.554
27	1.080	3.962
28	1.940	4.374
29	1.790	4.068
30	1.960	3.788

Appendix 3

The table below lists the share codes and corresponding company names of all shares that were incorporated in the analysis of this dissertation.

Table 33: *Share Codes and Respective Company Names*

Share Code	*Company Name*		*Share Code*	*Company Name*
AAC	Anglo American Corp S Africa		IPF	Investec Property Fund Ltd
AAL	Alpha Limited		IPL	Imperial Holdings Ltd
ABI	Amalgamated Beverage Inds		IPR	Iprop Holdings Ltd
ABT	Ambit Properties Ltd		IRV	Irvin & Johnson Limited
ACP	Acucap Properties Ltd		ITL	Interleisure Limited
ADC	Adcock Ingram Ltd		ITU	Intu Properties Plc
ADN	Adcock Ingram Ltd-"N" Shs		IVG	Invego Investment Limited
ADV	Advent Properties Limited		JCG	Jci Gold Ltd
AEG	Aveng Ltd		JDG	Jd Group Ltd
AEL	Allied Electronics Cor-A Shr		JNC	Johnnic Holdings Ltd
AEN	Allied Electronics Co-N Shrs		JOE	Hj Joel Gold Mining Co Ltd
AFE	Aeci Ltd		JSE	Jse Ltd
AFT	Afrimat Ltd		KEH	Keaton Energy Holdings Ltd
AFX	African Oxygen Ltd		KIO	Kumba Iron Ore Ltd
AGL	Anglo American Plc		KLO	Kloof Gold Mining Co Ltd
AGR	A M Moolla Group Limited		KOH	Kohler Limited
AHH	Afrox Healthcare Ltd		KTL	Kunene Technology Ltd
AHV	African Harvest Ltd		LBH	Liberty Holdings Ltd
AIA	Ascension Properties Ltd-A		LDM	Lindum Reefs Gold Mining Co
AIB	Ascension Properties Ltd-B		LES	Leslie Gold Mines Limited
AIP	Adcock Ingram Holdings Ltd		LEW	Lewis Group Ltd
AKJ	Arthur Kaplan Jewellery Hldg		LGB	Langeberg Holdings Limited

ALT	Allied Technologies Ltd		LGL	Liberty Group Ltd
AMB	Amb Holdings Ltd		LHC	Life Healthcare Group Holdin
AMC	Anglo American Coal Corp Ltd		LNC	Lenco Holdings Limited
AMG	Anglo Amer Gold Investment		LON	Lonmin Plc
AMI	Anglo Amer Industrial Corp		LOR	Loraine Gold Mines Limited
AMS	Anglo American Platinum Ltd		LTA	Lta Limited
AND	Andulela Investment Holdings		LTH	Lithotech Limited
ANG	Anglogold Ashanti Ltd		MAF	Mutual & Federal Insurance
ANN	Anglovaal Ltd-N Shares		MAS	Masonite Africa Ltd
APA	Apexhi Properties-Unit Cl A		MDG	Mdm Growth Investments Ltd
APB	Apexhi Properties-Unit Cl B		MDI	Master Drilling Group Ltd
APG	Autopage Holdings Limited		MDN	Madison Property Fund Manage
APP	Amb Private Equity Partn-Uts		MDS	Midas Pty Ltd
AQP	Aquarius Platinum Ltd		MEG	Millennium Entmt Grp Africa
ARD	Ardor Sa		MES	Messina Ltd
ARI	African Rainbow Minerals Ltd		MHH	Mih Holdings Pty Ltd
ARL	Astral Foods Ltd		MKL	Makalani Holdings Ltd
ARO	Anglo American Properties		MLB	Malbak Ltd
ASC	Ascendis Health Ltd		MMI	Mmi Holdings Ltd
ASG	Assmang Limited		MND	Mondi Ltd
ASR	Assore Ltd		MNO	Mainstreet Property Fund
ATT	Attacq Ltd		MNP	Mondi Plc
AUK	Aukland Health Limited		MNR	Minorco Sa
AVG	Avgold Ltd		MNS	Supersport International Hol
AVM	Avmin Ltd		MPC	Mr Price Group Ltd
AVS	Avis Southern Africa Ltd		MPR	Metboard Property Fund
AVU	Avusa Pty Ltd		MPT	Mpact Ltd

AWA	Arrowhead Properties Ltd		MSM	Massmart Holdings Ltd
AWB	Arrowhead-B		MTE	Marshall Monteagle Holding S
AWR	Allwear Limited		MTK	Metkor Group Pty Ltd
AXC	Apexhi Properties-Unit Cl C		MTN	Mtn Group Ltd
BAT	Brait Se		MTX	Metorex Ltd
BAW	Barloworld Ltd		MUR	Murray & Roberts Holdings
BCX	Business Connexion Group		MZG	Metje & Ziegler Ltd
BET	Beatrix Mines Limited		NCW	New Central Witwatersrand
BGA	Barclays Africa Group Ltd		NED	Nedbank Group Ltd
BIL	Bhp Billiton Plc		NEP	New Europe Property Invest
BJM	Barnard Jacobs Mellet Hldgs		NIB	Nedcor Investment Bank Hldgs
BLU	Blue Label Telecoms Ltd		NIN	Ninian & Lester Holdings Ltd
BLY	Blyvooruitzicht Gold Mining		NIV	Niveus Investments Ltd
BPRD	Barprop Limited-11% Min Ln		NPK	Nampak Ltd
BTG	Bytes Technology Group Ltd		NPN	Naspers Ltd-N Shs
BTI	British American Tobacco Plc		NPT	Newport Property Fund
BTS	British American Tobacco Hld		NRB	Nrb Holdings Ltd
BUF	Buffelsfontein Gold Mines		NT1	Net 1 Ueps Technologies Inc
BUR	Burlington Industries Ltd		NWH	Norwich Holdings Sa Ltd
BVT	Bidvest Group Ltd		NWL	Nu-World Holdings Ltd
BZK	Berzack Brothers (Holdings)		ODM	Ocean Diamond Mining Holding
CAN	Canadian Overseas Packaging		ODV	Oceana Investment Corp Plc
CAS	Cadbury Schweppes S Africa		OML	Old Mutual Plc
CAT	Caxton And Ctp Publishers An		OMN	Omnia Holdings Ltd
CBH	Country Bird Holdings Ltd		ORS	Orion Selections Limited
CCO	Capital & Counties Propertie		PAM	Palabora Mining Co Ltd
CFC	Commercial Finance Co Ltd		PAS	Protea Assurance Co Limited

CFR	Financiere Richemont-Dep Rec		PEI	Pep Limited
CGS	C.G. Smith Limited		PEN	Penrose Holdings Ltd
CGU	Cgu Holdings Limited		PEP	Pepkor Limited
CGW	Consol Limited/Old		PFG	Pioneer Foods Group Ltd
CHE	Chemical Services Ltd		PGL	Pallinghurst Resources Ltd
CHR	Charter Plc		PIK	Pick N Pay Stores Ltd
CIL	Consolidated Infrastructure		PIN	Polifin Limited
CKS	Crookes Brothers Ltd		PKN	Pick'N Pay Stores Ltd-N Shs
CLE	Clientele Life Assurance Co		PMA	Primedia Ltd/South Africa
CLH	City Lodge Hotels Ltd		PMG	Primegro Properties-Link Unt
CLI	Clientele Ltd		PMN	Primedia Ltd-'N' Shrs
CLR	Clover Industries Ltd		PON	Profurn Limited
CLS	Clicks Group Ltd		POT	Brian Porter Holdings Ltd
CMI	Cons Metallurgical Inds Ltd		POW	Power Technologies Ltd
CML	Coronation Fund Managers Ltd		PRA	Paramount Property Fund Ltd
CMP	Cipla Medpro South Africa Lt		PRI	Primedia Ltd-Units
CNC	Concor Limited		PRP	Premier Pharmaceutical Co Pt
CNF	Congella Federation Limited		PSL	Psg Financial Services Ltd
CNX	Conafex Hldgs Sa		PSY	Plessey Corporation Limited
COH	Curro Holdings Ltd		RBP	Royal Bafokeng Platinum Ltd
COT	Coates Bros. (South Africa)		RBX	Raubex Group Ltd
CPB	Micawber 274 Ltd		REB	Rebosis Property Fund Ltd
CRM	Ceramic Industries Ltd		REI	Reinet Investments Sa-Dr
CRW	Corwil Investments Ltd		REM	Remgro Ltd
CVH	Capevin Holdings Ltd		RES	Resilient Reit Ltd
CXT	Caxton Publishers & Printers		RFN	Randfontein Estates Limited
DAG	Da Gama Textile Co Ltd		RHE	Rhoex Ltd

DDT	Dimension Data Holdings Plc		RLO	Reunert Ltd
DIV	Diversified Property Fund		RMI	Rand Merchant Investment Hol
DLF	Del Monte Royal Foods Ltd		RMR	Rms Property Holdings Ltd
DLK	Deelkraal Gold Mining Co Ltd		ROM	Romatex Limited
DLT	Delta Property Fund Ltd		RPL	Redefine International Plc
DRD	Drdgold Ltd		RTN	Rex Trueform Clothing-N Shs
DST	Distell Group Ltd		RTO	Rex Trueform Clothing Co Ltd
DSY	Discovery Ltd		SAB	Sabmiller Plc
DTC	Datatec Ltd		SAP	Sappi Limited
DUK	Duiker Mining Limited		SBK	Standard Bank Group Ltd
ECO	Edgars Consolidated Stores		SBO	Saambou Holdings Ltd
EGN	Engen Limited		SDG	South African Druggists Ltd
EHS	Evraz Highveld Steel And Van		SFW	Stellenbosch Farmers' Winery
ELA	Elandsrand Gold Mining Co		SGG	Sage Group Ltd
ELH	Ellerine Holdings Ltd		SGL	Sibanye Gold Ltd
EMI	Emira Property Fund Ltd		SHF	Steinhoff Intl Holdings Ltd
ENR	Energy Africa Ltd		SHP	Shoprite Holdings Ltd
EQS	Eqstra Holdings Ltd		SHV	Sea Harvest Corporation Ltd
ERG	East Rand Gold & Uranium Co		SLM	Sanlam Ltd
ESV	Eastvaal Gold Holdings Ltd		SLU	Investment Solutions Hldgs
EUR	Eureka Industrial Ltd		SMA	Samancor Chrome Ltd
EVR	Evander Gold Mines Ltd		SNT	Santam Ltd
EVT	Everite Group Limited		SOL	Sasol Ltd
EXX	Exxaro Resources Ltd		SON	Southern Life Association
FAM	Frame Group Limited		SPE	Spearhead Property Hold -Uts
FDC	Foodcorp Limited		SPG	Super Group Ltd
FFA	Fortress Income Fund Ltd-A		SPN	Specialty Stores Ltd-N Shs

FIN	Fintech Ltd		SPP	Spar Group Limited/The
FIT	First International Trust		SPU	Spur Steak Ranches Limited
FRE	Free State Development & Inv		SRL	Sa Retail Properties Ltd
FRG	Free State Cons Gold Mines		SRT	Smart Group Holdings Ltd
FSB	First Natl Bank Holdings Ltd		SRY	Sentry Group Limited
FSC	Lion Match Co Pty Ltd/The		SSK	Stefanutti Stocks Holdings
FSP	Freestone Property Hldgs		STH	St Helena Gold Mines Limited
FWD	Kansai Plascon Africa Ltd		SUI	Sun International Ltd
GAR	Guardian National Insurance		SUN	Suncrush Limited
GBL	Genbel South Africa Limited		SVL	Southvaal Holdings Limited
GDA	Glodina Holdings Ltd		SYA	Siyathenga Property Fund
GDO	Gold One International Ltd		TAM	Tamboti Property Fund Ltd
GFC	Gold Fields Coal Limited		TBS	Tiger Brands Ltd
GFI	Gold Fields Ltd		TCP	Transaction Capital
GFN	Griffin Shipping Hldgs Ltd		TDH	Tradehold Ltd
GIJ	Gijima Group Ltd		TFG	The Foschini Group Ltd
GOC	General Optical Company Ltd		TIW	Tiger Wheels Ltd
GPL	Grand Parade Investments Ltd		TKG	Telkom Sa Soc Ltd
GPN	Group Five Ltd-"N" Shs		TLF	Tile Afrika Holdings Limited
GPT	Global Capital Private Eq		TLJ	Teljoy Holdings Limited
GRA	Gray Security Services Ltd		TME	Johnnic Publishing Ltd
GRC	Aveng Grinaker-Lta Ltd		TMG	Times Media Group Ltd/South
GSC	Genbel Securities Limited		TMX	Telemetrix Plc
GTA	Gentyre Industries Ltd-A		TON	Tongaat Hulett Ltd
GTB	Gentyre Industries Ltd-B		TOY	Toyota South Africa Pty Ltd
GUB	Gubb & Inggs Ltd		TRE	Trencor Ltd
HAG	Haggie Limited		TRU	Truworths International Ltd

HAL	Halogen Holdings		TSX	Trans Hex Group Ltd
HAR	Harmony Gold Mining Co Ltd		TUN	T & N Holdings Limited
HBN	Hartebeestfontein Gold Mng		UMN	Umdoni Property Fund
HCT	Hoechst South Africa Limited		USV	United Service Technologies
HDC	Hudaco Industries Ltd		UTR	Unitrans Ltd
HGT	Higate Property Fund		UUU	Uranium One Inc
HLH	Hunt Leuchars & Hepburn Hldg		VIL	Village Main Reef Ltd
HLM	Hulamin Ltd		VKE	Vukile Property Fund Ltd
HPA	Hospitality Property Fund-A		VLY	Velocity Holdings Limited
HPB	Hospitality Property Fund-B		VNF	Venfin Pty Ltd
HSP	Holdsport Ltd		VOD	Vodacom Group Ltd
IBM	Ibm South Africa Group Ltd		WAL	Waltons Stationery Co Ltd
ICH	Indus & Commercial Hldgs Grp		WAR	Gold Fields Operations Ltd
ICS	Ics Holdings Limited		WDL	Western Deep Levels Limited
IDW	Independent Newspapers Hldgs		WET	Wetherlys Investment Hldgs
IFR	Ifour Properties Ltd		WGR	Witwatersrand Consolidated G
IGE	Ingwe Coal Corporation Ltd		WHL	Woolworths Holdings Ltd
ILV	Illovo Sugar Ltd		WPH	Women Investment Portfolio-B
IMP	Impala Platinum Holdings Ltd		WRC	West Rand Consolidated Mines
INL	Investec Ltd		ZED	Zeder Investments Ltd
INP	Investec Plc		ZSA	Zurich Insurance Co South Af