Hybrid Optimisation and Formation of Index Tracking Portfolio in TSE

Mohammad Ali Assadi

December 2021

Supervisor

Dr Rasol Eskandari

University of Salford MANCHESTER

Abstract

Asset allocation and portfolio optimisation are some of the most important steps in an investors decision making process. In order to manage uncertainty and maximise returns, it is assumed that active investment is a zero-sum game. It is possible however, that market inefficiencies could provide the necessary opportunities for investors to beat the market. In this study we examined a core-satellite approach to gain higher returns than that of the market. The core component of the portfolio consists of an index-tracking portfolio which has been formulated using a meta-heuristic genetic algorithm, allowing for the efficient search of the solution space for an optimal (or near-optimal) solution. The satellite component is made up of publicly traded active managed funds and the weights of each component are optimised using mathematical modelling (quadratics) to maximise the returns of the resultant portfolio.

In order to address uncertainty within the model variables, robustness is introduced into the objective function of the model in the form of risk tolerance (Degree of uncertainty). The introduction of robustness as a variable allows us to assess the resultant model in worst-case circumstances and determine suitable levels of risk tolerance. Further attempts at implementing additional robustness within the model using an artificial neural network in an LSTM configuration were inconclusive, suggesting that LSTM networks were unable to make informative predictions on the future returns of the index because market efficiencies render historical data irrelevant and market movement is akin to a random walk. A framework is offered for the formation and optimisation of a hybrid multi-stage core-satellite portfolio which manages risk through the implementation of robustness and passive investment, whilst attempting to beat the market in terms of returns. Using daily returns data from the Tehran Stock Exchange for a four-year period, it is shown that the resultant core-satellite portfolio is able to beat the market considerably after training.

Results indicate that the tracking ability of the portfolio is affected by the number of its constituents, that there is a specific time frame of 70 days after which the resultant portfolio needs to be re-assessed and readjusted and that the implementation of robustness as a degree of uncertainty variable within the objective function increases the correlation coefficient and reduces tracking error.

Keywords: Index Funds, Index Tracking, Passive Portfolio Management, Robust Optimisation, Core Satellite Investment, Quadratic Optimisation, Genetic Algorithms, LSTM, Heuristic Neural Networks, Efficient Market Hypothesis, Modern Portfolio Theory, Portfolio optimisation

Table of Contents

Chapter 1: Introduction

As highlighted by a study conducted by the CBI (Central Bank of Iran) in 2014, the local money market (comprised wholly of local banks and credit institutions) accounts for approximately ninety percent of all finance and lending operations inside of the country (NadAli et al, 2017). Alarmed by the inherent risks caused by the over reliance on the money markets and the banking sector's over-exposure to short-term debt, law makers have recently been highlighting the need for the diversification of investment instruments which are able to cater to the tastes of different market participants whilst shifting the economy towards a more market-based approach (Feghhi Kashani et al, 2013).

After being introduced in 2007 the number of investment funds operating within the Iranian capital market rapidly grew from 7 to 339 within a span of 14 years with an overall AUM of over 631,000 million IRR equating to 22.97 billion USD at the time of writing (FIPiran.ir). The massive growth of the sector and its inherent adoption can be attributed to the underlying need of local investors for more customisable tailormade financial products which can create an equilibrium between returns attained and the risks involved (systematic and unsystematic). The study of the efficient allocation of assets can act as a powerful catalyst when aiming to attain the trust of investors whilst striving to maximise returns and mitigate risks and costs.

Furthermore, the need for innovative asset allocation strategies is highlighted by the current uncertain climate of the financial markets (themselves a result of the ever-increasing economic sanctions imposed by the UN and the United States in particular) which not only act as a barrier of entry for foreign funds but also further complicate matters for local managers who have to compete with ever-increasing rates of inflation when trying to attract new capital whilst juggling the various needs of market participants.

The approach formulated in this research attempts to optimise classical models of asset allocation whilst utilising robust optimisation techniques with the objective of addressing the aforementioned uncertainty and leveraging the benefits of both passive and active investment strategies with an emphasis on index tracking products.

1-2- Current Economic Climate of Iran

In order to fully understand the current economic climate of Iran, we first need to delve into the underlying issues which have arisen as a result of strained diplomatic relations between the country and the rest of the world.

In response to Iran's reported nuclear program the United States congress signed into law the Iran Sanctions act of 1996, which targeted both US and non-US businesses/persons who proceeded to make investments which contributed to the enhancement of Iran's ability to develop petroleum resources (Iran Libya Sanctions Act, 1996). This was further escalated in the CISADA act of 2010 which expanded upon the original sanction (aimed specifically at the country's ability to maintain and expand its domestic production of refined petroleum) and also established further provisions which allow for the prohibition of specified foreign exchange, banking and property transactions (Comprehensive Iran Sanctions, Accountability and Divestment Act, 2010).

By targeting all institutions and companies who contributed to Iranian oil production and/or refinery operations the US sanctions effectively crippled foreign investment in Iran whilst making all Iranian financial institutions pariahs on the international stage because of their involvement in various oil exploration, extraction and subsidiary projects. Coupled with the fact that at the time nearly eighty percent of the governmental revenue depended on oil exports these sanctions had an immense impact on the economic stability of Iran (Katzman, 2011).

As can be seen in the figure below the GDP in Iran was worth nearly 599 billion dollars in 2011, a number which has subsequently been reduced by sixty one percent in 2020 with a value of approximately 231.55 billion USD. (World Bank, 2020).

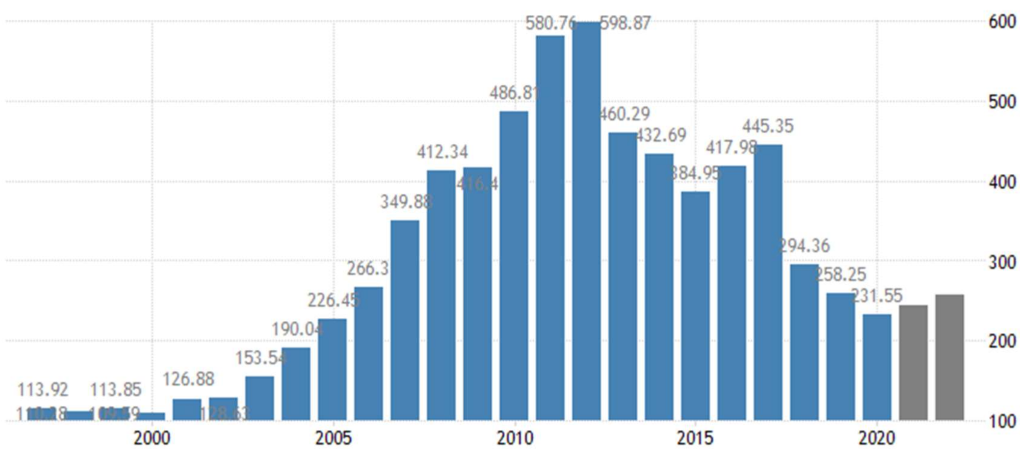

Figure 1: GDP (Current USD) Islamic Republic of Iran

The extent of which the gross domestic product has been affected by US sanctions is a subject hotly debated by both sides of the argument however it is outside of the scope of the current study. It should be noted however that as oil exports still make up nearly twenty percent of current GDP (according to world bank estimates), the effectiveness of the sanctions cannot be overruled.

We can further our understanding of the current economic climate of Iran by analysing the released annual (point to point) inflation rates published by the Statistical Centre of Iran.

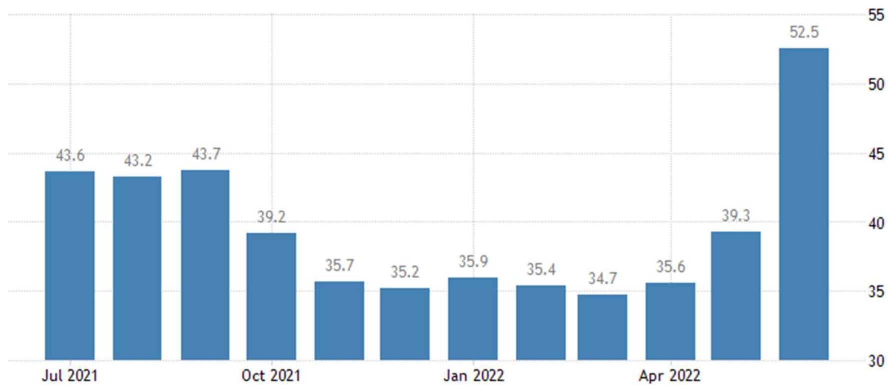

Figure 2: Annual Inflation Rate Islamic Republic of Iran

Figure 2 clearly outlines the high level of inflation that is currently rampant in Iran. By using the Consumer Price Index (CPI) published by the Statistical Centre of Iran it becomes apparent that transport, furnishings & home appliances, recreation & culture and food & beverage spending were hardest hit with increases of 61.3%, 60.2%, 59.5% and 57.4% respectively. (Statistical Centre of Iran, amar.org.ir).

In figure 3 we outline the fluctuation of the IRR/USD exchange rates which also highlight the devaluation of the rial and its effect on the inflation rates discussed above.

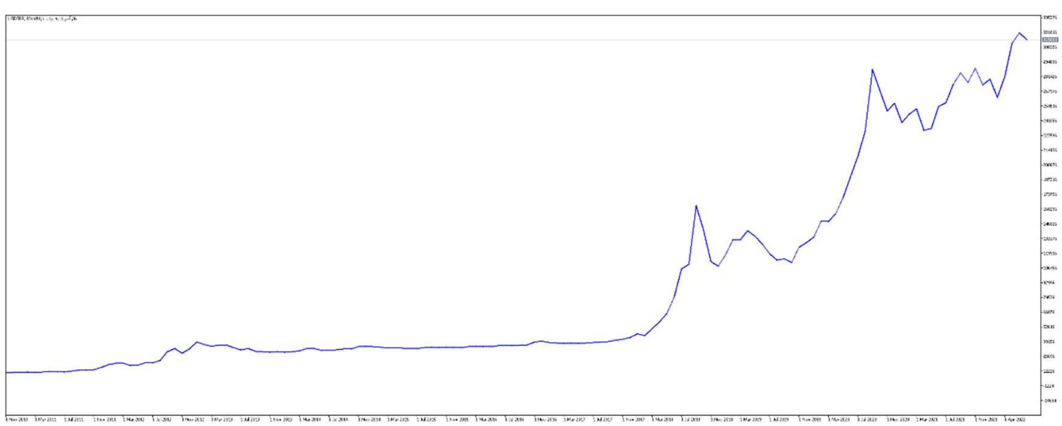

Figure 3: IRRUSD Exchange rate (Source: Mofid Securities Exchange Metatrader 5)

As can be seen from Figure 3, the IRR/USD exchange rate rises sharply from 14,335 IRR in 2011 to 315,000 IRR which is an increase of 2200 percent. In order to put this in context of the inflation numbers seen in figure 2 (July 2021-July 2022), the increase in the IRRUSD exchange rate during the aforementioned period is twenty two percent.

In contrast we can also look at the development of the Iranian capital market for further insight into the current economic climate of Iran.

The first factor that we can assess is the TSE (Tehran Stock Exchange) index which is portrayed in figure 4. As can be seen from the chart below the index has risen from 284,000 in September 2019 to a high of 1,965,000 in the June of 2020 before settling to a value of 1,487,000 in July 2022. Even without taking into the account the all-time high, the index has enjoyed a growth of nearly 420 percent within a two-year timespan. In conditions where there is an ever-increasing compound form of inflation taking hold in the economy, the Iranian stock market seems to be a relatively safe haven which could have been utilised by institutions and the general public to safeguard portions of wealth whilst maintaining spending power. Also, it quickly becomes apparent that this influx of liquidity could not have been a result of gradual foreign investment. As outlined in the beginning of this section sanctions imposed on Iranian financial institutions and markets by the United States act as a strong barrier against the inflow of new foreign capital into the country. So, in order to make sense of the huge increase we will need to look at the country internally.

One factor which could have led to high growth of the index could be attributed to the ever-increasing percentage of the population who have entered the market. This can easily be tracked by measuring the number of new trading account registrations at exchanges. According to figures released by Securities and Exchange organisation, the number of trading accounts in 2021 increased to 37,500,000 individual accounts from 11,664,000 in 2019. These figures become even more interesting when we look back further and find that in 2011 this number was only 1,900,000 accounts registered.

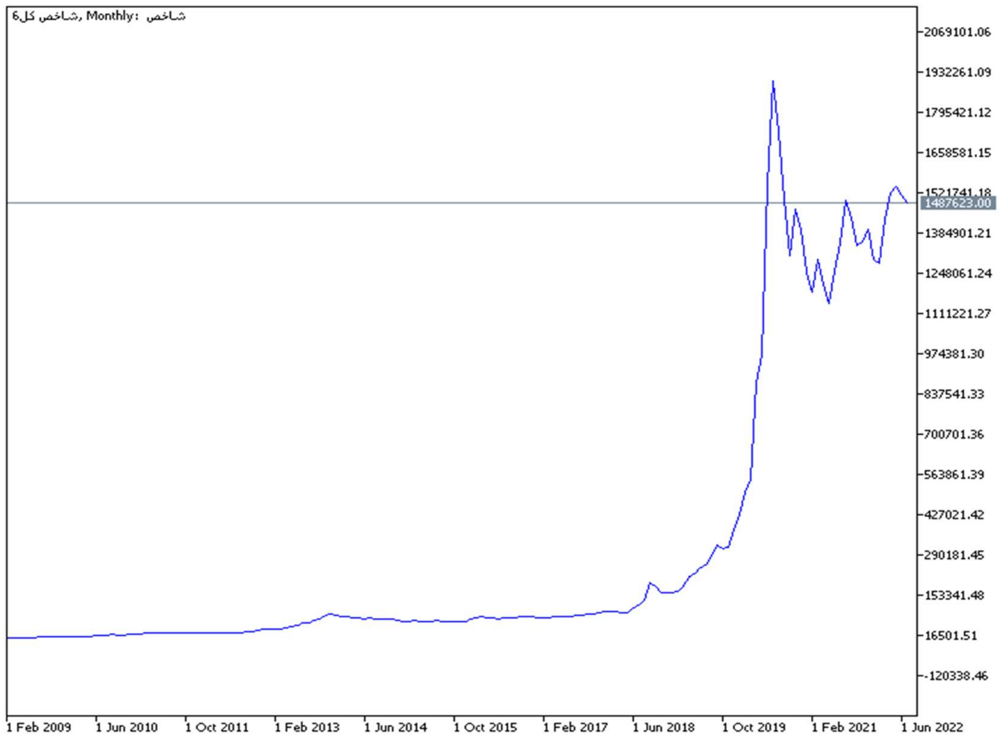

Figure 4: TSE index (source: Mofid Securities Exchange, Metatrader 5)

In other words, the number of registered trading accounts (which are legally bound to individual social security numbers for individuals) has increased by 1900 percent from 2011. (Bourse24.ir)

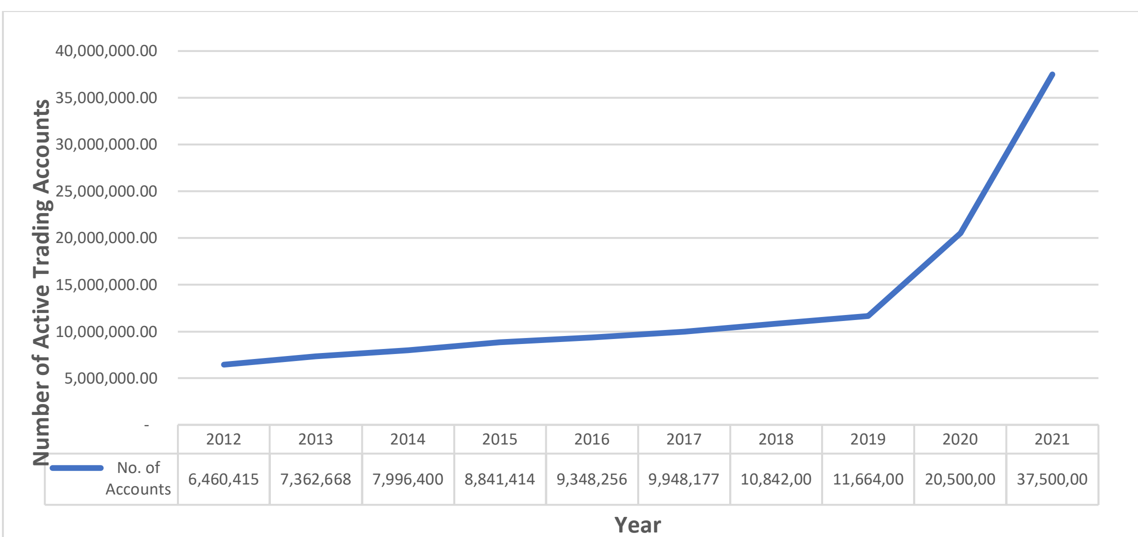

Figure 5: Number of Active Trading Accounts

In order to gain a clearer insight, we also need to address and identify the role of the government within this phenomenon. Governmental interests in the Tehran stock exchange can be traced back to 2006 when the initial law outlining the issuance and distribution of "Justice Shares" to low-income families was first passed within the national assembly (Framework for the Distribution of Justice shares Act, 2006). The aim of this law was the privatisation of parts of state-controlled industries whilst offering a subsidy/benefit program to lower income tiers of the economy, however subsequently it went through many changes during ratifications and only became fully active in its current state in 2016. According to numbers released by the National Privatisation Organisation there are currently 49.1 million "Justice" shareholders (Boursepress.ir). It should be noted however that many of these shareholders are not registered with any broker which could be due to lack of education regarding the program and its benefits.

The government and its leading figures also played a large part in the influx of new capital to the capital markets in 2019. President Rouhani explicitly lauded the benefits of capital markets and its possible rewards whilst maintaining that his government had a large part to play in the stability and prosperity of the market (dana.ir). This message was repeated by other influential members of the government giving people a certain sense of security whilst paving the way for more governmental privatisation programs and steering capital away from more inflationary assets such as foreign currency and gold at a time when the preservation of wealth had become a high priority for most of the general public.

As previously seen in figure 4, the combination of these events led to a massive spike in the TSE index which was well timed in order to act as buffer against increasing economic sanctions and high levels of inflation.

There is an argument to be made that the government has a duty to educate and empower these new shareholders who may have no previous experience of investment in the capital markets. Also, it could be argued that other market stakeholders such as brokerages and institutions could take advantage of this large influx of potential customers and capital in order to devise new innovative asset allocation strategies that can cater to the specific needs of their target audience.

The current economic climate of Iran presents a unique opportunity for these stakeholders to educate and engage a new generation of potential customers whilst managing their investment risks and attaining the highest rates of return possible. In the next part of this chapter, we study the capital markets whilst outlining the many challenges faced by potential investors.

1-3-Iranian Capital Market

Capital markets act as a foundation in the infrastructure of a healthy economy, acting as a magnet for the flow of capital whilst efficiently distributing them between active sectors of the economy based on their competitiveness (NadAli et al, 2017). This approach encourages public participation in national growth, advocates public ownership rights and also acts as a conduit for foreign investment, relieving the effects of governmental budget deficits whilst pushing much needed monetary system reform. The financial market in Iran has long been dominated by banking money markets, acting as the primary source of finance for businesses. The loan-to-deposit ratio of the banking sector at the end of 2017 and 2018 was reported at eighty-five and eighty percent respectively (Central Bank of Iran Report, 2018) outlining a high level of imbalance between the incoming and outgoing cash flow and disparity between expenditure and attraction of capital. Whilst trying to develop a healthy economy that is independent of governmental control, policy makers have often faced many obstacles such as the banking systems inability and inefficiency regarding the raising of capital and also their poor asset management protocols. These inefficiencies have resulted in an investment vacuum where excess capital is redirected away from productive markets towards inflationary non-productive investment vehicles such as gold and the hoarding of foreign currency, forcing the government to respond through the introduction of reactionary policies and highlighting the importance of capital markets and the role that they play in the economy.

Law makers have attempted to address these shortcomings by introducing legislature and regulations such as the Securities Market Act of 2005 (which paved the way for the development of new financial tools and products) and the Financial Institutions Act of 2009 which provided a standardised format of definitions for businesses that wanted to operate in the capital markets.

Governments have also utilised privatisation reforms with the aim of improving the qualitative and quantitative performance of the financial markets (Capital markets, banking & insurance sectors) whilst emphasising efficiency, transparency and proper ethics through the introduction of regulatory bodies and ombudsmen.

Although the above legislature and initiatives helped ease regulations and pave the way for a more prosperous capital market, there still exists substantial barriers and problems which are restrictive to growth and prevent the market from attaining its true potential. Low levels of market participation (in relation to the country population), short term outlook of trades, regulatory imposed market limitations, inability to attract foreign investment and low levels of leverage/finance offered by brokerages are only some of the limitations faced by these markets, each acting as a barrier in the future growth and further development of the capital market.

The points above become more important however when a review of global financial markets is conducted, with results highlighting a negative correlation between the rate of growth/maturity of the market and the level of trade restrictions imposed by regulatory and supervisory bodies. In other words, mature markets with higher levels of participation are less encumbered by restrictive regulations. In contrast when reviewing the history of the Tehran Stock Exchange it becomes apparent that low levels of market penetration and subsequent low investor base (previously around 8 percent of the country population) resulted in increased sensitivity from lawmakers which led them to introduce legislature primarily aimed towards the safeguarding of investor equity resulting in trade restrictions, limited daily price fluctuation ranges and volume dependant closing prices.

The market volatility and subsequent fluctuations in 2013 highlighted some of the problems that were a result of the above-mentioned regulatory restrictions. The high trade volumes that were the result of an in influx of new capital coupled with restrictions on the daily price fluctuations of a stock resulted in an accumulation of buy/sell orders at each extreme of the permitted daily range. The formation of these buy/sell order queues at each extreme of the 8 percent permittable fluctuation range were compounded by the escalation of market excitement and herding phenomenon, resulting in the lock up of liquidity and higher levels of risk for new investors that were experiencing FOMO (fear of missing out). The repetitive formation of these queues affects the trading trend in an adverse way as highlighted below:

- Formation of buy/sell queues downplays the role of analysis in the selection of suitable assets, creating a breeding ground for rumours and speculation whilst (wrongly) advocating herding behaviour and queue formation as a suitable selection criterion.
- Order queue formation gives institutional traders the opportunity to conduct market manipulation by utilising their disproportionate capital to steer the market by

initiating queues through the placement of large orders, making certain stock seem more or less attractive than they really are.

- Order queue formation serves as a braking mechanism for delaying market participants and their decision-making processes resulting in decreased levels of market liquidity.
- Repetitive formation of buy/sell queues increase the volatility and intensity of price fluctuations whilst increasing systematic risk. When reviewing more mature international markets and comparing the daily trends than those of the Iranian market it becomes apparent that the absence of fluctuation ranges actually results in less volatility. In other words, it could be implied that trade restrictions lead to an extent of overreaction and an unusual increase in the volatility of price fluctuations.

The Iranian Capital market is also plagued by a lack of diversity of the tools and institutions that cater to the needs and tastes of investors, resulting in lower levels of participation (adoption)and limited supply of liquidity entering the market.

The underdevelopment of proper risk management practices and underutilisation of hedging tools is another barrier to the effort of attracting foreign capital whilst trying to steer surplus liquidity away from bank deposits and towards the markets. As previously discussed, when studying mature markets, it becomes apparent that they facilitate the mitigation of risk through the trade of index-based products and their subsequent futures contracts. Unfortunately, the absence of these strategies and products is apparent in the Iranian Capital market.

To further complicate matters, policies such as privatisation and the distribution of "Justice Shares" among forty nine million of the general public (most of whom were unfamiliar with the rules and regulations governing the market) created a unique opportunity to increase public awareness on the advantages of the market whilst steering these novice investors away from short term speculation and more towards long term investment and growth through the use of innovative low risk investment vehicles which offer suitable returns. The emergence of such an investment culture could lead to less volatility in the market whilst offering resistance to the formation of trading queues as discussed above. Whether lawmakers capitalise on this opportunity is yet to be seen but the short-term effects of such reactionary policy making is clear. The subsidised distribution of these shares between an uneducated general public has already resulted in a series of mass selloffs by individuals who due to a lack of knowledge have no interest in long term investment and only see this privatisation strategy as a one-off subsidy payment by the government.

Finally, it could be argued that the further development and growth of the Iranian capital market will depend on the sustained and long-term support of the government and their regulatory bodies. Widespread adoption as a suitable investment vehicle by the various classes of society will rely heavily on educating the masses and the introduction of diverse investment products which can cater to the needs and requirements of the various market participants (in terms of risk mitigation and expected returns). Changes in the investment culture towards a more long-term investment approach can also help to attract surplus liquidity form the money markets whilst managing the effects of uncertainty and regulatory restrictions.

In order to address the need for new innovative products that are able to overcome the aforementioned problems of the Iranian capital market, a proper framework which addresses optimal asset allocation, the risk-returns proposition and uncertainty of the markets need to be devised which caters to the unique circumstances of the Iranian capital market. To better understand the efficient allocation of assets, a review of the historical evolution of investment strategies is offered in the next section of this chapter.

1-4 Evolution of Investment Strategies

The evolutionary path of the investment strategies can be divided into four distinct stages. (Bahr-oloom, Tehrani & Hanifi, 2012). The first stage began with the individual analysis of equities and the allocation of assets based on that analysis. This approach later evolved into a methodology which also assessed the relationship between individual equities and grouped them into funds or portfolios which worked in unison to mitigate risk and increase rewards. With the advent of the efficient market hypothesis, the third stage advocated the use and development of index tracking portfolios/funds which attempted to mirror the performance of the market indexes. The final and fourth stage could be viewed upon as a marriage of previous strategies in which the conflict between active and passive investment strategies is finally resolved and a new paradigm of investment strategies is offered (Bahr-oloom, Tehrani & Hanifi, 2012).

Generally, an investor is looking to adopt a strategy which achieves the highest returns with the lowest level of possible risk involved. However, these are not the only deciding factors which dictate the efficient allocation of assets within a portfolio or fund. Fund managers regularly utilise a collection of market factors and investor preferences (which are in line with expected risk and returns levels) to devise efficient asset allocation strategies (Ruiz-Torrubiano & Suarez, 2008). To achieve these goals and objectives managers utilise active and passive strategies of investment.

Active strategies rely on the experience, knowledge and expertise of managers to select equities and suitable entry/exit points with the objective of attaining the highest levels of returns possible (Beasley, Meade & Chang, 2003). As a result, the recurring restructuring of the portfolio in reaction to market shifts and changes leads to higher transactional costs and risk.

Passive strategies on the other hand are based on the efficient market hypothesis of financial markets where it is offered that the price of an asset reflects all of the inherent risk and desired returns of that asset. In this approach the main role of the portfolio manager is the selection of equities as a portfolio where the final objective of the resultant portfolio is to mimic the fluctuations of the market index and attain similar returns to that of the index. Some benefits of this approach include lower transactional costs and risk (due in part to the much less frequent reshuffling of the portfolio) whilst also bearing the disadvantage that investors utilising these strategies are in fact forsaking returns exceeding the market index. Some researchers however argue that over an extended period of time, passive management does in fact offer higher yields than that of active strategies (Sharpe, 1991).

The underlying concept which ties passive investment strategies into the efficient market hypothesis is the simple fact that the market cannot return higher yields than itself (surpass itself) with investors on average gaining the same returns than that of the market minus expenses. In other words, the more active they are in trading terms the higher the transactional costs, which in turn leads to lower returns (Hanifi, Bahr-Oloom & Javadi, 2009).

It could be argued that the Efficient Market Hypothesis removes any notions of "competitive edge" in the prediction of market fluctuations of a particular equity over an extended period of time. Faced with this dilemma active investors will also have to overcome higher transactional costs. Whether an investor believes in the Efficient Market Hypothesis or not, the logical approach to investment would be to manage all controllable variables through diversification and the utilisation of risk mitigation protocols, whilst aiming to minimise transactional and tax related costs. The underlying principle and logic of passive strategies is to achieve these goals through the introduction of index orientated investment products.

The counter argument also exists that there are recorded cases of exceptional managers with remarkable market insight who have been able to achieve higher returns than that of the market index. These arguments are expanded to introduce new innovative strategies which attempt to merge both active and passive disciplines into a new format of strategies which benefit from both previous approaches (Schoenfeld, 2004). One of the proposed hybrid models is the core-satellite approach in which investors utilise passive strategies to develop a core set of index-bound investment assets which are then supplemented with a series of actively managed satellite components. In this framework the core component takes the shape of an index tracking portfolio whilst actively managed portfolios of individually picked equites are utilised to supplement returns.

As previously discussed in this chapter, the Iranian capital market requires new innovative investment products to drive adoption within the mass population. Hybrid core-satellite portfolios could be utilised to attract new capital into the market from non-expert investors whilst allowing institutional smart money a new mode of diversification. These new products have the capacity to increase the investment outlook of investors whilst reducing the sensitivity to daily fluctuations (reduce risk). By developing products based on these hybrid frameworks we can compensate for the underdevelopment of risk management tools in the market by aligning them with regulatory policies which will not only increase the penetration and adoption of the market, but will also address market uncertainty.

It should be noted that the increased complexities of the market environment now mean that investors and manager will also need to take into account an additional factor of uncertainty. To this extent, robust optimisation techniques have been promoted by many in circumstances where a small irregularity in the input data could result in an incomputable solution. Thus, there is also a clear need to address, analyse and measure uncertainty as an input variable and offer new methodologies which attempt to manage uncertainty in a robust framework (Gharekhani, 2012).

1-5 Importance of study

A short review of the previous sections of this chapter reveals that there is a direct need for new innovative products in the Iranian capital market which are able to safeguard the wealth/buying-power of a growing population in an economy battling with inflation and the devaluation of the local currency. The design of a framework for the formation and optimisation of a portfolio of equities would allow financial institutions and fund managers to offer innovative products which can be tailored to the specific needs (risk and rewards) of different market stakeholders. By ensuring that the proposed framework is robust enough to handle the market uncertainty (which itself is a product of the current unique socio-political circumstances of Iran), the reliability of the final product is increased.

It is clear that due to governmental promotion and legislation the Iranian capital market is growing both quantitively (market value, number of participants, equities…) and qualitatively (new institutions & new products). This also become apparent when we review that that number of registered funds in the market increased from 7 in 2007 to 339 in a span of 14 years. The utilisation of the aforementioned framework would further increase the number of these products ensuring product diversification in the market.

Suitable allocation of financial funds in the capital market is one of the most important principal factors in the decision-making process of investors. When considering that governmental legislature and promotion have caused a massive influx of novice retail investors in the market, the need for suitable asset allocation strategies which can mitigate risk and maximise rewards becomes even more apparent. The adoption of the framework will not only benefit the investors but will also increase the efficiency of the market (less speculation- driven market, less order queue formation, longer term outlook on trades, higher liquidity).

It should also be noted that the lack of mathematical tools in the past lead to less inclination towards the adoption of quantitative optimisation techniques. These approaches inherently require the processing of massive amounts of data (specifically in an index tracking based approach) which made the problem more mathematically complex. However technological advances in the fields of heuristics and machine learning coupled with the development of user-friendly software are now pushing the adoption and utilisation of innovative quantitative optimisation techniques.

Furthermore, there has previously been a lack of inclination towards the use of such optimisation techniques, with studies having been conducted to show that results from such models may not have sufficient levels of reliability in real life situations. One example would be the various studies conducted comparing equal weight portfolios (EWP) and their mean variance counterpart (MVP) in which it has been shown that the EWP outperformed its MVP counterpart in different evaluation terms (Cai, 2021). These examples should not however be used as the basis for the outright rejection of risk-reward optimisation theories as they lack diversification. It quickly becomes clear that in when utilising such theories is practical situations which require reliable solutions a series of improvements and reforms have to be implemented on top of classical frameworks. This is further compounded when adding robustness into the model to protect the results from the uncertainty of the input data. The proposed framework is a step in this direction of improving classical models whilst adding robustness through design.

The utilisation of a core-satellite approach in this framework will also aim to reduce management and transactional costs whilst attempting to maximise returns. The development of this approach is also in line with the fifth development program outlined by the government advocating the transparency and performance of investment managers, which should eventually attract potential investors.

Finally, it is important to note that the development of index-based products in mature markets such as the United States is now nearly 40 years old and the market cap of these products dwarves the combined market cap of all investment funds in Iran (one trillion dollars vs 23 billion USD) (FIPiran.ir). It quickly becomes clear that the development of these financial products and the adoption of index-based strategies has not received much consideration in the Iranian market, highlighting an opportunity to offer a practical framework which can be utilised to develop and introduce a low cost, transparent investment product which not only manages risk but also operates in times of high uncertainty (Gharekhani 2012). Thus, the research beforehand is innovative in its problem-solving methodology whilst expanding the boundaries of the related scientific domain.

1-6 Research Questions

Whilst developing a roadmap to developing the framework highlighted in the previous section of this chapter the following research questions are proposed:

- Can a heuristic genetic algorithm be used in order to identify and select suitable stocks in the formation of an index tracking portfolio whilst also optimising the asset allocation (weights) of each constituting component?
- Can a heuristic RNN (Recursive Neural Network) in the shape of an LSTM (Long-Short Term Memory) model be utilised to predict future rates of return of the market

index using historical data, thus increasing the effectiveness and robustness of the aforementioned genetic algorithm model?

- Will the utilisation of mathematical robust optimisation methodologies reduce the tracking error of the developed index tracking portfolio? Is there a significant difference between the performance of a robust index tracking portfolio and the TSE index?
- Is the tracking error of the subsequent index tracking portfolio affected by the number of its constituents and further integer constraints?
- Is there a correlation between the number of assets in in the resultant portfolio and the realisation of performance on par with that of the index (Whereby the tracking error is a benchmark of the performance against the index)?
- What is the optimum number of assets which constitute the index tracking portfolio?
- What is the optimum interval at which the resultant portfolio should be re-evaluated and reshaped?
- Will the utilisation of a core-satellite strategy further develop the resultant portfolio to achieve higher rates of return when compared to the returns of the index?

1-8 Research Goals

The expected goals of this research are highlighted below:

1. Offer framework for the formation of an optimized robust index tracking portfolio in the Tehran Stock Exchange.
2. Utilisation of heuristic quantitative modelling methods of uncertainty analysis in the portfolio optimisation.
3. Develop hybrid core-satellite model to exceed returns when compared to market index.

4. Overcome existing limitations apparent in classical asset allocation investment frameworks in regards to model constraints and robustness.
5. Advocate the use of index tracking products and strategies to investors and other market participants by highlighting their risk reward benefits.

1-7 Research Methodology

When considering the objectives and methodology of this research we can ascertain that it falls into quantitative realm of research and its methodology can be considered as correlational research. Correlational research is a type of non-experimental research method in which a researcher measures two variables, understands and assesses the statistical relationship between them with no influence from any extraneous variables. Based upon objectives, correlational research can be categorised into three groups: dual variable correlation, regression analysis and covariance analysis or correlation matrix. Correlation is used to analyse the type and scale of relationships between variables whereas regression analysis is used to forecast future trends of a standard variable (dependant) based on the relationship between the dependant variable and one or more independent variables which have been recorded previously.

By reviewing the goals of this research, it becomes clear the objective of the study is to devise an optimised strategy for the formation of a hybrid core-satellite asset allocation model through the utilisation of heuristic algorithms and the implementation of mathematical robustness. Thus, we are looking to minimise the tracking error of the core component (in regards to the returns of the index) whilst supplementing returns with the satellite components. As we are able to calculate the returns of the portfolio and the tracking error of the core component and the returns of the hybrid core-satellite model, we can measure the effect of various optimisations (number of portfolio constituents, constituent weights, ratio of core-to-satellite components, degrees of uncertainty/robustness….) and compare them against the base index. This allows us to construct a correlational framework for assessing each optimisation method implemented, and verify whether the introduced changes have a meaningful relationship to the end results.

1-9 Scope of Research

1-9-1 Geographical Domain of Research

The research domain specified in this study encompasses all publicly tradeable companies and funds listed on the TSE.

1-9-2 Time Period of Study

The time scope of this study is a four-year period ending in 2018.

1-10 Statistical Population & Sampling

The establishment of a portfolio which tracks the index with high level of precision requires the selection of stocks which have the highest impact on the index itself. This requires that the initial sample pool is narrowed to only include assets which have the highest impact on the index. Therefore, filtration and cross-sectional selection methodologies need to be implemented to limit the possible viable solution space to the 100 top stocks by market capital and also stocks that experienced 100 days of active trading per year within the five-year studied period. Further characteristics can be utilised to further narrow the initial solution space and improve the efficiency and accuracy of the proposed model.

Due to the heuristic nature of the model, the data sets will also need to be divided into training and testing sets in order to train and evaluate the performance of the resultant model.

1-10-1 Data Collection

The data collected in this study includes the daily returns of stocks and funds traded on the TSE which will be collected from the Tehran Stock Exchange archives.

1-11 Data Analysis Method & Modelling

As previously stated, the objective of this research is the formation of a robust framework for the optimised allocation of assets, in conditions of uncertainty, utilising a hybrid cores-satellite format. Mathematical programming is utilised to model the problem whilst a heuristic genetic algorithm will be used to solve the gradient decent problem of index tracking (minimising tracking error of the base index), whilst supplementing returns (through the selection of suitable satellite components). To further this goal econometric frameworks are used to simulate missing values and standardize and initialise all data input into the model.

Software and programming languages such a Python, Lingo, MATLAB and their respective libraries will be used to solve the mathematical models whilst Excel is utilised for the formatting and initialisation of data.

1-12 Research Limitations

- Privatisation or insolvency of some companies within the defined time frame.
- Changes in index calculation method within the defined time frame
- Ignoring of trading and management fees for funds in the satellite components
- High rates of inflation have resulted in high prices for new IPOs (which may not meet filtering characteristics highlighted above) affecting the index.

1-13 Summary

This chapter attempted to give the reader an insight into the Iranian capital market and its fundamentals whilst offering a core framework for the research. Through the definition of the research problem and its importance a methodology for finding the solution was offered.

In the next chapter a conclusive literature review of the core categories of asset allocation and optimisation (including a specific study into machine learning practices) will be conducted whilst offering an insight into the fundamentals of Efficient Market Hypothesis and its relation to the Iranian capital market.

The third chapter will further expand the research methodology and extrapolate the data collection and filtering methodologies used whilst introducing the mathematical models developed for the research problem.

The fourth chapter will be used to convey the analysis of the data utilised, and the results of the proposed model.

Finally, the fifth chapter will be dedicated to the analysis of results of the proposed model for users of the research whilst offering suggestions for future research to be conducted in this subject fields whilst offering the contributions and limitations of the current research.

Chapter 2: Literature Review

2-1 Introduction

As stated in the previous chapter of this thesis, the main aim of this research is to develop a framework for the formation of an optimised robust index tracking portfolio whose returns are to be supplemented via the implementation of a core-satellite model, in which the core index tracking component is optimised using heuristic and exact programming. In order to meet this goal and develop a theoretical framework, a comprehensive literature review of the underlying investment paradigms and their subsequent optimisation methodologies was conducted.

The first part of this chapter will focus on the introduction and evolution of portfolio theory and the effects of the efficient market hypothesis on its subsequent investment strategies. The review will assess the conflict between active and passive strategies(index-tracking) whilst highlighting modern approaches such as the core-satellite approach, which aim to overcome the disadvantages of the aforementioned strategies, whilst reconciling both and benefitting from their advantages. A portion is also included to review research on the efficiency of the Iranian capital markets which is hoped to provide further insight into the domain of study.

The second part of the review will offer the results of previous research into various optimisation techniques (exact and heuristic) and the implementation of robustness when dealing with uncertainty and stochastics.

The final part of this chapter a summary of research conducted on portfolio optimisation methodologies is conducted with the aim of identifying the research gap that this thesis attempts to rectify.

2-2 Portfolio Theory

One of the key issues investors face, is how to allocate wealth among alternative assets. Almost all financial Institutions have the same problem with the added complication that they need to explicitly include the characteristics of their liabilities in their analysis (Elton, 1997). In other words, it could be argued that the most important factor in the decision-making process of investors are risks and returns. If stocks are risky by nature, then the primary focus of any investor is to compile a selection of stocks which has the highest desirability and returns, whilst maintaining the lowest level of risk possible. This is the same as deciding upon the best and most optimized portfolio out of all combinations possible which is known as the portfolio selection problem.

Th evolution of the portfolio theory can be thought of as three distinct phases: Traditional Portfolio Theory (TPT), Modern Portfolio Theory (MPT) and Post-Modern Portfolio Theory (PMPT). Traditional Portfolio theory was mainly concerned with the analysis of individual securities and was characterised by simple, non-systematic, subjective and insufficiently analytical approach to forming an optimal portfolio. MPT on the other hand focuses on the on the analysis of portfolio characteristics, enabling the optimisation of the relationship between risk and returns through the utilisation of an objective system-based approach. Finally, the PMPT approach phase was developed to address the lack of compatibility of the MPT assumptions and market reality (Lekovic, 2021).

Appearing at the beginning of the 20th century TPT played an important role in the field of asset allocation until the publication of Markowitz's portfolio selection article in 1952. TPT itself evolved from a subjective approach based on subjective assessment (without any scientific and analytical basis) to a more scientific approach based on the analysis of financial statements of companies and securities (partly due to stricter controls on financial companies listed on the stock exchange).

TPT emphasized the analysis of individual assets while the analysis of portfolio characteristics was ignored. By highlighting the inefficiency of the markets, followers of TPT believed that the fundamental analysis of a company's internal financial statements could result in higher returns. Also, practitioners of TPT introduced diversity into portfolios based on the law of large numbers resulting in portfolios consisting of securities with the highest rates of returns. This form of naive diversification (increasing the number of securities in order to reduce the overall portfolio risk) implied that if investors want to eliminate risk, it is enough to invest in a large number of securities (Lekovic, 2017).

Williams (1938) argues that the total portfolio risk could be eliminated by diversification. His study continued to claim that whereas future dividends are uncertain in nature, investment in a sufficient number of assets/securities in the form of a portfolio could reduce risk to zero. He further claimed that through the utilisation of the law of large numbers, actual portfolio returns are almost the same as the expected returns.

John Richards Hicks was another proponent of simple diversification. He argued that the risk factor is important because it affected the expected investment period and the expected level of return on the investment. He continued to claim that in conditions of risk, there are many probably outcomes and suggests the presentation of these outcomes using the expected value and an appropriate measure of dispersion, however he fails to indicate the measure itself (Hicks, 1935).

Leavens (1945) highlights the importance of diversification, however the author does not include correlation in his analysis but argues the assumption that return on securities is independent. After this analysis, leavens does state that this assumption is not always in line with the reality of the markets i.e. diversification of a portfolio to include companies from one industry cannot protect the investors from unwanted factors affecting the entire industry, however even diversification between industries cannot protect investors from cyclical market factors that adversely affect all industries at the same time.

In summary it can be concluded that in the TPT approach, no statistical measures are used to quantify risk and improve and complement the fundamental analysis based on the accounting methods. TPT was based on simple analysis and characterised by a subjective and insufficiently analytical approach (Lekovic, 2022).

Pioneered by Harry Markowitz in 1952, Modern Portfolio Theory (MPT) offers a mathematical framework for the optimisation of the risk return ratio of a portfolio, shifting focus away from the analysis of individual securities and simple diversification. Markowitz begins his approach by assuming that an investor has a specified amount of capital to invest. This capital is to be invested for an indicated amount of time also known as the investor's holding period. At the end of the investor's holding period all securities that have been purchased at the beginning will be sold and the returns of the sale will either be withdrawn or reinvested. In other words, this approach can be labelled as a "single period approach" to investment with t=0 indicating the start of the period and t=1 as the end of the period. At t=0 the investor must select which securities to purchase and keep till t=1. In traditional portfolio theories the investor must evaluate the returns of different securities at t=0 and then invest in those with the highest yields. Markowitz argues that this approach will be illogical as the investor will not only want to maximize returns but will also want to minimize risk and be assured of the returns as much as possible. To justify this argument Markowitz continues by stating that if investors were only looking to maximize expected returns, then they would only invest in a single asset which has the highest expected rate of returns, whereas a quick review shows that investors are in fact the owners of a selection of securities. To rationalize this behaviour, it could be said that investors consider both phenomena of risk and return simultaneously. Thus, an investor which is looking to maximize returns while minimizing risk is in fact faced with conflicting targets which need to be balanced against each other. As a result, the investor must diversify their portfolio by purchasing a number of different types of securities.

Markowitz (1952) argues that returns on securities are correlated mutually and that instead of investing a in a large number of securities, investors should in fact invest in securities with low return correlation. He continues to advocate the implementation of efficient diversification (in contrast to the simple diversification approach advocated by Hicks), providing mathematical proof that appropriate diversification can minimise portfolio variation at the given return level and quantify the trade-off between risks and returns, thus creating a set of efficient portfolios that maximise returns at the given risk level (Markowitz, 1959).

In his 1999 paper, Markowitz also accredits Roy (1952) as an equal contributor to the field of modern portfolio theory. Roy also developed an independent set of efficient portfolios similar to that of Markowitz with the distinction that Markowitz required exclusively non-negative investments and proposed allowing the investor to choose a desired portfolio from the efficient frontier whereas Roy allowed the amount invested in any security to be positive or negative and recommended a choice of a specific portfolio (Markowitz, 1999). Furthermore, it could be argued that the most important aspect of Markowitz's work is that he has shown that the risk of individual securities in a portfolio is not as important to the investor as their contribution to the variance of the overall portfolio that depends on their covariance with other securities in the portfolio (Rubenstein, 2002).

Sharpe and Lintners work on the capital asset pricing model (CAPM) can be thought of as the complimentary second part of the macroeconomics of the capital market (Markowitz, 1991).

The CAPM is a model that describes the relationship between the expected return and risk of investing in a security and is calculated using the below formula:

$$ER_i = R_f + \beta_i(ER_m - R_f)$$

Formula 1: CAPM Calculation

Where ER_i is the expected return of investment, R_f is the risk-free rate, β_i is the beta of the investment and (ER_m-R_f) is the market risk premium. Two important factors to note are that investors expect to be compensated for the risk undertaken and the time value of money. By utilising the formula above we can evaluate whether a specific security is fairly valued when comparing its expected returns to its risk and proportionate time value of money.

Although the introduction of the CAPM is attributed to Sharpe (1964) and Lintner (1965), Tobin (1958) is also considered to have laid the groundwork for CAPM. Markowitz highlights the fact that the models introduced by Tobin and Sharpe are similar in postulating a model with n risky and one riskless security however they differ in their assumptions. Tobin assumed that an investor can invest at the risk-free rate whereas Sharpe argues that the investor can either borrow or lend at the same rate. The second major difference in assumptions is the fact that Sharpe postulated that his model applied to all securities whereas Tobin argues that his model is only applicable to "monetary assets". To elaborate further, Tobin's more cautious assumptions expressed doubts that cash itself should be considered risk free (Markowitz 1999). Tobin (1958) agreed with Markowitz's theory that it was more beneficial to "not place all eggs in single basket" and that portfolio diversification would in fact lead to lower levels of risk. While looking to adjust and lower portfolio risk without changing the composition of the stocks within, Tobin demonstrated that an investor can attain desirable level of risk by changing the ratio between cash and stocks in a portfolio.

Furthermore, he proved that using cash was a more efficient way of lowering the risk of a portfolio than altering the composition of stocks.

It should be noted that MPT over simplifies the reality of the financial markets as it ignores a series of different market factors: 1) transactional costs 2) information asymmetry 3) inefficiency of the markets 4) volatility of correlation (correlation of return on securities is changing daily, so it needs to be observed dynamically 5) irrational behaviour of investors 6)individual risk averseness of investors (Lekovic, 2021). By taking into account the above factors it can be argued that MPT suffers from the following shortcomings (Radivojevic, 2009):

1) Choosing an optimal portfolio is not viewed as a continuous process of tracking changes and adjusting portfolio over time, but as a decision to be made on a one-time basis.
2) The Assumption about the infinite divisibility of securities, i.e., the possibility of buying or selling securities in unlimited proportions, does not stand in practice.
3) In conditions of a financial crisis, the correlation coefficients converge to one, so the benefits of diversification are reduced or even completely disappear. The portfolio risk becomes equal to the simple weighted sum of the individual risks of securities of which it is compiled.

Despite these limitations MPT is relied upon by market participants to make effective timely investment decisions when structuring a portfolio or assessing the performance of a portfolio. By allowing managers to make reliable decisions MPT contributes to the depth, liquidity and efficiency of the market whilst its mathematical base and precise nature of results give a sense of security and comfort (Vyas, 2014).

Developed in the nineteen eighties at the Pension Research Institute (USA) Post Modern Portfolio Theory (PMPT) attempted to overcome the limitations of MPT and make it more in line with market Reality. Designed to offer a stronger more precise framework PMPT went against the existing assumptions that variance and standard deviation were reliable risk measure, that investors all shared the same expectations and most importantly the fact that returns on financial assets followed normal distribution. PMPT introduced minimum acceptable returns (MAR) for each investor, acting as a benchmark that could be used to gauge the performance of a portfolio making it more customisable to the specific needs of each individual investor.

PMPT defines risk as the total returns' volatility around the mean value and is measured by variance or by standard deviation of return where deviations above or below the mean value are treated in the same way. Whereas MPT associates' risk with achieving an average return, PMPT argues that the investment risk should be linked to the MAR for each individual investor and that only volatility of the returns below MAR represent risk. Return above the target creates uncertainty which is nothing but a risk-free opportunity to achieve higher than expected returns (Rom, 1993). Furthermore, with the implementation of MAR (which is distinct to each individual investor) and given that it is used to determine an efficient frontier there are now an infinite number of efficient frontiers (based on an infinite number of investors and their relevant MAR). This is in direct contrast to the MPT approach which discusses an infinite number of efficient portfolios along a singular efficient frontier which are defined with three variables: standard deviation, correlation coefficient and returns (Lekovic, 2021).

Finally, it should be noted that although Post Modern Portfolio Theory is thought of as an upgrade to Modern Portfolio theory, it shares some characteristics with TPT in that they both advocate greater diversification of portfolios when compared with MPT. Also, both TPT and

PMPT can be customised to the needs of the individual investor whereas in MPT individual investor goal is not explicitly taken into account.

2-3 Efficient Market Hypothesis

The notion that financial markets are efficient is one of the underlying principles of modern portfolio theory. In order to fully understand this principle, there is a need to review the CAPM of Sharpe (1965) in which he attributes the fluctuations and uncertainty behind stock returns to systematic (undiversifiable) and unsystematic (idiosyncratic or diversifiable) risk. Systematic risks or market risks are threats which the investor faces for having invested in the market, including inflation and the threat of war. Unsystematic risks on the other hand are limited to a single stock with examples including rumours on change of management, potential takeovers or a new product failure. As it is possible to eliminate unsystematic risk in a portfolio by compiling a portfolio which covers all the different aspects of a market (while also negating any rewards), investors can only expect a return when faced with market or systematic risks.

The original work submitted by Bachelier in 1900 argued that price speculation should be a “fair game”, where expected profits to the speculator should be zero (Fama, 1970). Although this would be the first instance where notions of a random walk and efficient markets are introduced his work would go unnoticed by many for forty years.

Fama (1965) attempted to answer the question whether the past history of a stock’s price can be used to make meaningful predictions concerning the future of its price. In his study he tests the random walk model of stock price behaviour and continued to make the argument that although Markowitz defines an efficient portfolio as one that has maximum expected return for given variance of expected return, it is in fact the case that if yields on securities follow distributions with infinite variances however, the expected yield of a diversified portfolio will also follow a distribution with an infinite variance. In this situation the mean-variance concept of an efficient portfolio will become meaningless (Fama, 1965). Fama concludes that this is not a challenge to diversification however and concludes that factors other than variance should be used to develop portfolio analysis models.

Samuelson (1965) also studied the random nature of stock prices. In his study he outlined the inherent information in stock prices. He argues that the intrinsic value of a stock was nothing more than their market value at any given moment. Furthermore, constant fluctuations are the result of continuous disagreements between buyers and sellers on the intrinsic value of a stock and only the value which is agreed upon as a result of the aforementioned disagreement, is the true market value.

Fama (1970) conducted a historical review on the theoretical and empirical literature on the efficient market model and introduced the "Efficient Model Hypotheses". He introduced the notion that markets can be categorised as having three states of efficiency: Weak, Semi-strong & Strong. The weak-form EMH implies that the market is efficient, reflecting all market information. This hypothesis assumes that the rates of return on the market should be independent; past rates of return have no effect on future rates. The semi-strong form EMH implies that the market is efficient, reflecting all publicly available information. This hypothesis assumes that stocks adjust quickly to absorb newly released information. The semi-strong form EMH also incorporates the weak-form hypothesis. Given the assumption that stock prices reflect all new available information and investors purchase stocks after this information is released, an investor cannot benefit over and above the market by trading on new information. The strong-form EMH implies that the market is efficient: it reflects all information both public and private, building and incorporating the weak-form EMH and the semi-strong form EMH. Given the assumption that stock prices reflect all information (public as well as private) no investor would be able to profit above the average investor even if he was given access to exclusive information.

Malkiel (1973) identifies flaws in the fundamental and technical analysis of securities and liken the price action movement of prices to a random walk, where previous price movements have no bearing on the future values of a security. He is widely accredited for introducing concepts of efficient market theory and modern practical risk and return concepts to general investors. His summary however came at a shock to most investors who had finally been shown that although the stock market is generally in growth, it acts in an unpredictable accidental way which is akin to the slurred movement of a drunk who is in search of the way home. The publication of information which is itself unpredictable defines the value of a stock. Malkiel's resulting analysis was that active management of mutual funds was a waste of time and that a cheap and diverse index tracking fund would be the best choice for long term investors (Schoenfeld, 2004).

To summarise the Efficient Market Hypothesis, it could be said that prices reflect information in an unbiased way. This reflection can include past price history, publicly available information and in some cases all available information. This view is in conflict with the behavioural finance school of thought which argue that there are irrational pockets of investors that drive prices to be irrational and the size of this group is such that the market cannot take advantage of them in terms of pricing. Below we will review some of the literature that highlight market anomalies and irrational participant behaviour in contrast to the EMH.

Basu (1977) attempted to test the EMH by conducting a study on the effects of P/E ratios on future prices of a stock in order to highlight pricing bias. Although his results failed to unequivocally reject the EMH, it showed that there are some lags and frictions in the adjustment process where security prices are impounded by publicly available information.

Banz (1981) identified an anomaly in the NYSE which presents a challenge to the EMH. In his study of a forty-five-year period he identified that on a risk adjusted basis, the fifty smallest stocks outperformed the fifty largest by an average of one percentage point per month giving rise to the small firm effect which was corroborated in many different countries.

De Bondt & Thaler's (1985) work on the over-reaction hypothesis showed that most people over-react to unexpected and dramatic news events. Strikingly their results showed that when comparing two different portfolios consisting of "winners" and "losers" in a three-year period, the resultant portfolio containing the "loser" stock outperformed the prior "winners" by twenty five percent. Although many aspects of their findings remain without adequate explanation, it is important to note that even after a five-year period of portfolio formation the excess returns gained by the "losers" portfolio remains consistent.

Lo and McKinlay (1988) test the random walk hypothesis for weekly stock market returns by comparing variance estimators derived from data sampled at different frequencies. By utilising a simple volatility-based specification test, they rejected the random walk hypothesis for weekly stock data. This rejection however does not imply the inefficiency of stock price formation.

There exist other such market anomalies which could be used to counter the efficient market hypothesis such as those presented by Benartzi & Thaler (1995) regarding the performance of stocks vs treasury bills over a 100-year period or anomalies attributed to forward discount bias in the foreign exchange markets (Froot & Thaler, 1990), however to categorically reject it would not be acceptable as there are just as many occurrences in the literature supporting EMH.

For example, the review conducted by Fama (1991) highlights an extensive quantity of work in the defence and proof of EMH. Interestingly, his findings on private information again mirror the work presented by Samuelson and Malkiel in which it is re-iterated that private information is very rare, and that analysis of returns using 2- and 3-portfolio benchmarks that are consistent with multi-factor asset pricing models show that investment managers do not necessarily have access to privileged information.

2-4 Birth of Indexing

Although the birthplace of Index Investing models can be traced to the early 1850's, the first stock indexes were first established at the end of the 19th century. Charles Dow (Founder of the Wall Street Journal) and Edward David Jones first established the DOW JONES index in 1884 which was comprised of the stock price average of 9 railroad and 2 industrial companies. By doing this Dow hoped that investors could have a general view of the market performance of the market at any given day. In 1986 the DOW JONES Industrial Average (comprised of 12 stocks) was established and the railroad stocks where registered as a separate index named the DOW JONES Transportation Average. Daily Publication of the DOW JONES Industrial Average (DIJA) commenced on the 26th of May 1896 in the Wall Street Journal. The DIJA was expanded to include 20 stocks in 1916 with a further 4 being added in 1928 bringing the total number of stocks in the DIJA to 30. This index is currently made up of the moving average of 30 stocks with the only member of the original 12 being General Electric. Although the DOW JONES index is still used by investors as a measure of market performance, most agree that the S&P500 is the best standard to measure market performance in the United States. Alfred Cowles conducted a comprehensive study in 1913 which later became the foundation of the S&P stock indexes. Before the 1960's there were no technologies available to calculate up to the minute indexes such as the S&P which relied on the moving average of the market value of all stocks in the markets. Since then, many index based products ranging from index tracking portfolios and electronic tradeable funds to options and futures contracts which are linked to these portfolios have been devised and offered in different markets internationally (Schoenfeld, 2004).

An artificial growth in the market continued till the mid 1960's causing investors to believe that specialist experts are able to beat the market easily by identifying high performing stocks and it was only after the subsequent market crash that competent experts were distinguished from those who had chosen stock and securities based on luck.

The fundamental theory which links index-based investments to the EMH is the simple fact that the market cannot outperform itself. On average, investors will be rewarded the market returns minus transactional costs. In other words, by being more active they will incur more costs.

This factor should be combined with the underlying principle of the EMH which is the fact that all information regarding the present and future performance of a stock are reflected in its current price. Thus, the price of a stock balances itself like a scale with an equal possibility of going up or down based on the emergence of new information. Therefore, active investors must not only overcome transactional costs but the EMH states that it is impossible of forecast the correct fluctuations of a single stock during a long-term period of time. This observation acts as the underlying philosophy behind index orientated investment principles. Regardless of whether an investor believes in an efficient market or not, the most logical method of investment is to first manage controllable variables through the reduction of risk via diversification, and then minimize the number of transactions and the costs that they incur. Accomplishing this objective is the principal logic behind index orientated investment and index based financial products (Schoenfeld, 2004).

After its introduction the index-orientated investment approach adoption faced numerous delays due to a lack of trust resulting from unfamiliarity within investment circles. The Wells Fargo research and analysis department began a new initiative in regards to portfolio management headed by William Fouse, John McQuown and James Vertin. On the 1st of July 1971 Wells Fargo established the first index tracking portfolio with a 6-million-dollar contribution from Samsonite fixed to the NYSE index. The resulting portfolio invested an equal proportion of the funds in each of the 1500 existing companies listed in the New York Stock Exchange, however it quickly became apparent that the management of such a portfolio was extremely difficult due to the unique properties and divergence of each stock in relation to others. In other words, the portfolio needed constant re-arranging in order to maintain the balance. High transaction costs lead to a change in portfolio strategy where instead of spreading the investment between members equally it was distributed by order of market value. The resulting portfolio grew in sync with the performance of the market and there was no longer a need to continuously re-arrange the structure thus becoming self-balancing.

Index based funds however were benefitting from these conditions and enjoying stable growth such that total assets under management grew from six million USD in 1971 to 10 24 billion USD in 1980. This immense growth was also supplemented by other changes in the investment environment (such as the changes in rules and regulations governing commissions and also the acceptance of index tracking products as a low-risk investment vehicle) brought about by lawmakers. On the 1st of May 1975 regulations regarding stock trading commissions where reviewed. The previous charge of two percent commission on trades greatly affected index-based products as such portfolios had a high number of constituents.

The change that brought about the most transparency however was the announcement that prudence in mutual funds would be calculated based upon the diversification of the fund and not the merits of each individual contributing element.

As more and more research highlighted asset allocation as the most important factor in the return of portfolios, index-based investment theories garnered more acceptance and confidence.

Ellis (1975) conducted a study reporting that 85 percent of active managers had been unable to gain higher returns than that of the SS&P500 index in a 10-year period. He further argued that investment in the stock market was a "zero sum game", where all investors would eventually receive the same returns than that of the market. There exists a loser for every winner in the market and that not all investors can have greater returns than that of the market, as they are the market. He summarized that in order to beat the market investors should gain a reflection of the market through an index tracking portfolio with the lowest cost.

A study conducted by Brinson, Hood & Beebower (1986) revealed that investment policies such as asset allocation were of higher importance than investment strategies such as market timing and individual stock selection. They argued that investment policies (main groups being: stocks, bonds, cash and real estate) could account for 93 percent of the fluctuations in returns experienced by the portfolio in the timeframe. In other words, they summarized that investors should spend the vast majority of their time in asset allocation rather than predicting market movements and unease in the selection of individual stocks.

John Bogle from the Vanguard Group introduced the first index tracking portfolio for retail investors in 1976. The new portfolio tracked the S&P500 index without any pretentions of beating its performance, later becoming one of the biggest mutual funds in the United States with nearly 82 billion USD under management in 2003. The Vanguard S&P500 however had humble beginnings with only 11 million USD attracted in the funds initial public offering in 1972.

As further theories regarding index-based investments arose, natural comparisons were made with the returns from active managers. Sharpe (1991) proved that without taking into account transactional costs, the mean of all dollars invested actively must have the same performance of the mean of dollars invested in index-based portfolios. For example, Sharpe asked investors to focus on all managers who had committed to have their performance rated and compared to the index of country x. He continued to argue that an emerging market such as Brazil should be used as an example (as an emerging inefficient market could be used to show that the theory could be expanded to all other markets). All investors in the aforementioned market could be categorized into two groups: active managers and index-based managers (who could be internal or external). The performance of these two groups would either be similar to that of the index (passive managers) or there would deviate from it (active managers). In this case the presumption is that the index is a suitable benchmark of the market performance. We know that if we group all of the active managers together (whether external or internal) in a single national portfolio then they will have the same returns than that of the index (Sharpe, 1991).

The assets under management of the passive index-based managers will be similar to that of the index as they will be made of the same stock with similar ratios. Thus, we are left with the active managers which as a group should have similar performance to that of the index. This is because these two groups of investors make up the base index and keeping in mind that the index-based groups closely mirror the base index then the active group must also fall into the same category. Also, if the active and passive managers both keep the same assets then they will have the same returns. When we take into account transactional costs then it is the active managers group which are affected negatively. Statistically it has been shown that those active managers partake in two to four times the number of trades that a passive manager will initiate resulting in higher tax deductions as well as a less efficient portfolio. These added costs act as barriers to higher returns for active managers, which when compounded with an absence of reimbursement for unsystematic risk has led to the evolvement of active management principles to a point where higher returns to that of the index are sought whilst controlling costs and risk. This structured approach in portfolio management is a direct result of the reinterpretation of the founding principles of an index-based methodology.

2-4-1 Advantages of Index-Based Investments

Four important factors which contributed to the expansive growth of an index-based approach include: ease of risk budgeting, lower transactional costs, simplicity of manager performance evaluation and competitiveness of returns in respect of actively managed approaches.

2-4-1-1 Risk Budgeting

The risks faced by institutional investors can be categorized into two groups: 1. The overall risk of the investment portfolio (which can be comprised of many different types of assets with different managers overlooking each different asset group). 2. The relative risk or active risk which is a result of the deviation of performance and returns of the manager from the performance and returns of the base index which they have committed to being evaluated against. The degree of deviation from the base index is directly linked to the frequency of 'bets' that a manager undertakes rather than tracking the index itself. These 'bets' depend upon a number of different factors and can be intentional or unintended in nature. Selection of a stock portfolio comprised of high market capital value equities or the purchase of stocks from outside of the base index are clear examples of such risky bets These bets can also be implied or unobvious; for example, the management of a diverse portfolio comprised of stocks from different sectors of the market which are all sensitive to changes in the interest rates (sectors such as the automotive industry and financial sector). The only apparent way to reduce risk would be to track the index which is exactly what an index-based managing approach does. It is of importance to note that the study conducted on portfolio risk does not aim to pit active management and index-based methodologies against each other. In truth these two approaches could (and should) be used in tandem to achieve higher rates than that of the index whilst allowing managers to control risk and lower transactional costs (Schoenfeld, 2004).

2-4-1-2 Lower Costs & Commissions

Management commissions for portfolio managers (whether active or index based) can vary greatly depending on different factors such as previous performance, fame, ability in creating new investment opportunities, overhead and personnel costs and the complexity and depth of research and investment processes of managers. Inherently the commissions for an index-based manager are much lower than that of an active manager due to the lower underlying costs of the investment approach. However, it should be noted that even index-based approaches require a level of resources to achieve the same returns than that of the index in an efficient and continuous way. The components of the portfolio will not be stable in nature and in the course of a year they will have to be shuffled and changed. The management of these changes in a way which does not affect the performance of the portfolio requires expertise and experience. Index based managers must also keep an eye on the inflow and outflow of liquidity to the portfolio and maintain reserves for the management of these cash flows. Finally, part of the returns from most stocks includes dividends paid out to stock holders or interest, which needs to be managed in a such a way where cash reserves are not accumulating in the portfolio and are reinvested accordingly. Although the aforementioned complexities require index-based managers to have a certain level of expertise and experience, they also benefit from lower research costs and also gain relief from the inherent economies of scale of their portfolios. On the other hand, active managers must continually develop tools and opportunities in order to gain higher returns than that of the index, and in a competitive field where many managers are looking for an edge over their competitors, any competitive edge soon becomes nullified once identified by others. The inherent nature of active managers also means that as a result of their decision-making process, they will undertake a higher number of trades resulting in higher transactional costs whereas index-based managers will only restructure their portfolio when it is needed to track the index better. The efficient management of the volume of trades of a portfolio is an important factor in the success of any investment manager. (Schoenfeld, 2004).

2-4-1-3 Ease in Manager Selection

An index-based management approach has removed the need to analyse many different criteria in the selection of managers by investors looking to reach their objectives. In comparison to active managers (which look to utilise a wide range of assets and allocation strategies), index-based managers operate in a more transparent method, removing the need for a detailed analysis of investment processes.

• Nature of Risk: Index-based managers which track the base index have the same risk as that of the base index, removing the need for risk factor analysis.

• Important factors in Performance: The performance of index-based managers is very easy to gauge as their overall objective is to have the minimum deviation possible from the returns of the base index. Three factors which contribute to tracking error include:

1. Excessive amounts of cash in the portfolio (un-invested capital)

2. Transactional costs

3. Assignment of suitable weights to each component of the portfolio

In contrast, the evaluation of active managers is a very hard process including analysis of:

- Different components of the portfolio
- Market timing
- Asset allocation

Furthermore, each of the aforementioned processes have numerous unidentified factors and we also have to keep in mind that in order to react to different market triggers, managers may change their methodology and style throughout time (Schoenfeld, 2004).

2-4-1-4 Performance in Relation to Active Managers

As previously noted, one of the contributing factors to the expansive growth of adoption of index-based approaches was investor discontent at the performance of classical active managers. The artificial bullish market of the 80's and 90's in the United States presented a clear problem to managers which were looking to achieve returns greater than the market through fundamental analysis. By studying the table 6 below, we can compare the performance of active large capital funds against a common market index such as the S&P500. We find that both in short-term and long-term periods the index has consistently matched (And sometimes outperformed) the median of active managers. These results also debunk the common misconception that in bear markets active managers perform better.

Active Large Cap Funds versus S&P 500 Index

Funds versus the S&P 500	Period						
	One Quarter	Year to Date	One Year	Three Years	Five Years	Seven Years	Ten Years
S&P 500 index (official) return	−3.15	−3.15	−24.77	−16.09	−3.77	5.60	8.53
Percentile ranking	58	58	63	68	74	66	66
Number of funds in universe	406	406	405	377	334	283	217

Source: BGI, Wilshire data as of 3/31/03.

Figure 6: Active Large Cap Funds vs. S&P 500 Index (Schoenfeld, 2004)

Further analysis shows that in the economic downturn of early 2000 active managers enjoyed relatively higher returns than that of the index. In this timeframe some expert managers upheld their commitment of gaining consistent returns. However, this success becomes more meaningful when we take into account the survivorship bias that arose out of the conditions of that time. This bias parameter is taken into account by only considering managers which are still active. Managers which had to stop their activities based on poor performance are removed from the pool and new managers take their place (Schoenfeld, 2004).

2-4-2 Barriers of Adoption

Although the reasoning behind the expansive growth of index-based approaches looks convincing, it has still been unable to garner the support of all its opponents. In the next section we will outline some of its most common criticisms.

Many opponents to the index-based approach argue that the only markets where active managers face problems in gaining desired returns higher than that of the base index are artificially bullish markets. They continue to stress that in a time of market downturn and high volatility active managers are more cautious and so are able to achieve better results than a blind strategy of index tracking. As can be seen from the diagram below, the performance of stocks selected by active managers is not bound to market direction in a timeframe.

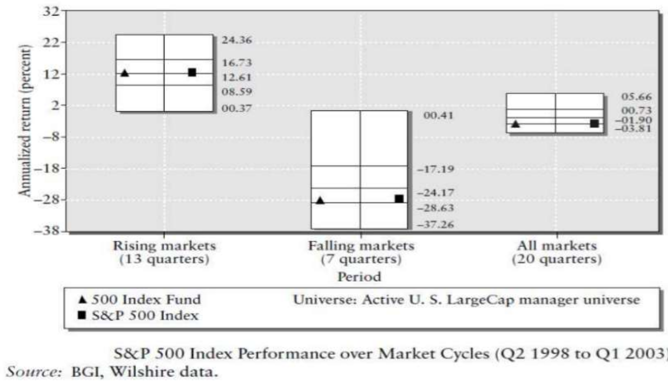

Figure 7: S&P 500 Index Performance over Market Cycles (Schoenfeld, 2004)

Figure 7 also shows that in bullish markets, the index-based approach enjoys better performance than that of active management, however in bearish markets the difference in performance is of lower magnitude. Different theories have risen to explain this phenomenon with the most widely accepted being the difference in cash reserves being held by active managers when compared with index-based approaches. In theory index-based managers should not have cash reserves but in practice in order to fulfill settlement of fund investors and also the payment of interest and dividends, small cash reserves need to be managed. In contrast active managers generally hold considerable cash reserves and mostly use a tactical liquidity strategy as part of their broader investment plans (Schoenfeld, 2004).

When active managers deem the value of assets to be no longer profitable, they will liquidate large parts of their portfolio. In a bearish market these cash reserves act as a safe haven until market inertia is achieved and new investment opportunities present themselves.

When active managers deem the value of assets to be no longer profitable, they will liquidate large parts of their portfolio. In a bearish market these cash reserves act as a safe haven until market inertia is achieved and new investment opportunities present themselves.

2-4-2-1 Index Based Approaches are only practical in Efficient Markets

Although on principle index-based approaches are feasible outside of the United States, international adoption of these methodologies has been slow. Sharpe's Active Managers Mean is also applicable to markets other than that of the United States meaning that disregarding trading costs and commission active managers will have similar returns to that of the index. This becomes important when considering that in some international markets the difference in costs between active management and index-based approaches is even higher than the United States (Schoenfeld, 2004).

The lack of support could in part be explained by the performance of classical active managers throughout the 90s who expertly managed to consistently gain higher returns than international indexes such as the Morgan Stanley Capital International (MSCI) and Europe Australasia Far East (EAFE), bringing about the misconception that index-based approaches may not be suitable due to the inefficiency of non-American stock markets. Meanwhile in a 5-year timeframe ending on the 30th of September 2001 the performance of the EAFE index in relation to all active managers had reached its lowest ranking .The turnaround started in the last 3 months of the 2001 and well into 2002 where the EAFE managed to improve its ranking and place itself above the average of all active managers. This template shifted towards a normal distribution in 2003. Three factors that contributed to the improvement of performance of the EAFE and brought about changes in the methods of investment in international markets include: The weight of Japan in the index, higher levels of efficiency in international markets and the reduction of stock collections outside of the base index.

Due to the direct relation of the last point to the Iranian stock exchange we will attempt to explain this in more detail.

2-5 Efficiency of the Iranian Market

Shooshtar & Namazi (1995) conducted an experiment to show that the Iranian Stock market had a weak form efficiency. Their result showed that the consecutive changes in the price of securities in the TSE do not follow a random walk model. Their results also highlighted the fact that the average returns of stocks using the filtering rule, was more than that of buy-hold approaches and concluded changes in price are not independent and coincidental. Their reasoning continues to state that there exists some form of trends and templates in the price movements of securities that investors could leverage to gain higher returns than that of the market. Finally, they suggest that as more companies are accepted in the TSE and public knowledge regarding investment paradigms increases the TSE will move towards a more efficient frontier.

Allahyari (2009) tested the weak form efficient of the TSE using three methods: Serial Correlation, Run Test & Spectral Analysis. His results highlighted many flaws and limitations in the market that hinder the efficiency of the market. Below some of the main factors leading to the inefficiency of the Iranian market are discussed (Allahyari 2009):

1. Flow and dissemination of information in the TSE is flawed and there needs to be an extensive overhaul in the technologies and methodologies used to enable participants to receive information in a more efficient manner.
2. Low trading activity also hinders the efficiency of the market and there is a need to educate the general public on the advantages of investing in the TSE which will in turn lead to more market penetration. The study also highlights the fact that high interest rates on savings accounts offered by banks create a barrier to entry to the market with most of the population choosing guaranteed returns in contrast to the market risk factor.
3. Lack of rules and regulations regarding misinformation and market manipulation.

4. Lack of a knowledgeable Analysts at the institutional level which can compile all relevant information and offer decision making frameworks to high level management. This essentially means that as there not enough knowledgeable analysts working through the available data, therefore the prices of securities cannot inherently reflect their value.
5. The under-development of self-regulating professional bodies (such as the chartered accountant's association, National Bar Association and Investment Banking Consortium) in relation to the capital market.
6. Lack of foreign investment (which at the time limited foreign investment to 10% of a single securities market cap).
7. High level of market participation from pseudo-governmental investment companies which not only adhered to the political policy of the time but also destabilised the market.

Dehdar & AliAghayi (2009) continue this work and examine some evidences indicating the lowest level of the Weak Form of Efficient Capital Market Hypothesis in TSE. The results of the Parametric and Non-parametric tests indicate that the time-series of stock returns are not consistent with the Random Walk Theory and that the correlation coefficient rejects the Null Hypothesis at different statistical intervals and levels of assuming a Weak Form of the Efficient Capital Market Hypothesis.

Ahmadzadeh et al (2014) utilised the windowed testing procedure of Hinich & Patterson to highlight the weak form efficiency of the TSE. Although they argue that studies conducted by Grossman & Stiglitz substantially lowers the level of efficiency achievable, their study also makes some further recommendations regarding improving the efficiency of the Iranian capital market:

1. Utilisation of electronic trading platforms
2. Development of regulations which ensure the reliable and verified dissemination of information.
3. Abandoning the use of predetermined daily price variation ranges
4. Reducing the effect of “noise traders” which contribute to the inefficiency of the market by creating hype and speculation through herding behaviour

Barasoud & Zomorodian (2019) conducted a study of 354 investors (including institutions, brokerages and funds) to identify whether they were followers of classical investment paradigms or behavioural finance. Their results highlighted that although the vast majority of the research population reflected a desire towards classical strategies they were in fact bound by the hybrid nature of the market and often partook in speculation and based their decisions on rumours rather than hard facts. The researchers attempt to address this phenomenon by highlighting the research conducted by Hunton (2005) which studied the decision-making process of investors in the high technology sector of the Chicago stock market. The study showed that weak EPS and financial variables in the high-tech sector did not dissuade investors but in actuality the expectation of high returns in the future fuelled speculation and lead to high demand in the market. Barasoud & Zomorodian (2019) conclude that although there is eagerness between the study participants for classical strategies based on the analysis of information there needs to be more development of tools which cater to the needs of the market participants.

In contrast, Jouzbarkand & Panahian (2020) utilise the Kalman approach to linear filtering to ascertain the efficiency of the TSE. They argue that the TSE is so inefficient that the prediction of future returns is possible and attribute this inefficiency to a variety of different factors including:

- Low liquidity of the market
- Untimely dissemination of information

The study concludes by re-iterating the fact that the market will only be able to achieve a state of efficiency if there is a substantial increase in the volume of trades leading to higher liquidity and the proper flow of information between different stakeholders is addressed.

Global growth and competition have led to a lessening of the gap in efficiency between developed international stock and the United States. With the advent of the global village national boundaries and borders have been put aside as investors from all over the world are willing to invest in international markets. Although this is most prevalent in the United States and United Kingdom its affects can also be seen in the European Union as well. Foreign investment in the Dutch stock market has increased from 50 to 80 percent during the 1998-2003 timescale. One of the clear signs of these changes is the 21 percent drop in trading costs of the EAFE from 1998 and the rapid and timely release of earnings, price and economic information of stocks (Schoenfeld, 2004).

Gaps of inefficiency are rapidly being narrowed and eliminated putting more and more pressure on classical managers and the importance of index-based approaches and risk control strategies are becoming more apparent, highlighting the benefits that they can bring about to developing markets.

We can see traces of these changes in the Tehran Stock exchange as well. By taking an overall look at the number of companies accepted in the TSE we find that the number of analyst institutions active in the capital market is also growing. Amin and Novin Investment banks were first established in Iran in 2008 and with the growth of the sector, it is only inevitable that market depth, information dissipation/access and in result efficiency will also grow exponentially. Also, low trade costs (currently at 0.5 percent) are an important factor in attracting foreign investment in the TSE. Finally, it could be argued that index-based approaches will form an inevitable part of the future investment paradigm of the Iranian stock market.

2-6-Union of Active investments and Index Based Approach

There has long been a debate on whether passive, index-based strategies are better performing than their active counterparts. Index-based investment proponents argue that on average, active market participants will only realize returns similar to that of the market. In their argument this is due to the fact that investment is a "zero sum game" in nature and after a deduction of trading costs and overheads active investors will have a lower return when compared to their index-based counterparts. On the other hand, opponents to passive investment strategies often argue that this approach only offers average returns whilst disregarding the potential for inefficiencies in the market. Taking into account the recurring theme, that for every investor that has the necessary skillset and expertise to achieve higher returns than that of the market, there also exists an unsuccessful investor, market equilibrium is achieved when the resultant trades of both investors vector towards the movement of the market itself. The very notion that there are investors who are able to navigate the various complexities of the market and achiever higher returns than that of the index, opens up the possibility for a truce between the two schools of thought. Thus, it can be seen that investors are making extensive use of index-based approaches in tandem with active investment strategies as part of a larger strategy (Schoenfeld, 2004).

One of the hybrid applications of these two strategies is the Core-Satellite approach. Many investors utilize an index-based approach to construct a core investment position on one or many groups of assets and then add active managers as satellite constituents. In this structure the index-based core portfolio should achieve returns similar to that of the index whilst supposing that active managers will be able to supplement these returns. In terms of risk, the index tracking core portfolio will have the same level of risk as the market whilst giving the investor the option to give managers more leeway in their satellite investment selections. Core-Satellite investments have matured into a standard between investment professionals whilst guaranteeing relative stability for all investors (whether institutional or individual).

2-7 Core Satellite Investment

Many investors have turned to core-satellite approaches to improve upon their investment strategies. The core-satellite approach divides the portfolio into two different groups; a passive group called the core and an active part called the active or satellite component. The core aspect of the portfolio is usually managed by a single manager and the active or satellite component is dedicated to equities with higher returns that require a manager with specific skillsets and expertise. The aim of the core component is to manage risk and increase efficiency whilst limiting costs whereas the objective of the satellite component is diversification and achievement of higher returns through higher efficiency. The core-satellite approach is a cost-effective strategy for managing the relative risk (also known as the tracking error) of a portfolio. This approach does not deal with the absolute risk of the core portfolio (reduction in portfolio value). By having the majority of assets in the core component of the portfolio, the portfolio becomes susceptible to severe economic cycles in the market. Tracking error (which is defined as the deviation of an active portfolio from the base index) is also known as relative risk and is the opposing factor to absolute risk. Absolute risk is a concept coined by Markowitz in 1952 in answer to the portfolio selection problem. By forcing severe investment limitations in active strategies (As a result of tracking error limitations) investors lose out on chances of achieving higher rates of return whilst limiting risk in specific market conditions (bearish markets), where active strategies are thought to be more efficient than their passive counterparts. Tracking error is not necessarily a bad thing. Good tracking error points to a higher rate of efficiency in the portfolio when compared with the base index and bad tracking error highlights a lower efficiency in relation to the base index. In core-satellite portfolios' assets which are expected to have better performance than that of the index are placed in the satellite component. However, if the economic conditions are not favourable, it may be the case that the satellite component underperforms the index. Usually, investor expectations are never symmetrical. In other words when stock market indexes are performing well investors are happy to partake in relative returns strategies. On the other hand, when indexes are performing poorly there is a strong sentiment for an absolute returns strategies.

In the late 1980's many investors had started to use the core-satellite paradigm. One of the first studies conducted was the book published by Scherer (2002) who focused on the quantitative portfolio selection methodologies. In his book Scherer divides the management of assets into the two core and satellite segments continuing to state that the core component be managed passively and the satellite component to be managed actively, wherein the optimized allocation of assets between the core and satellite component depends greatly on the level of risk accepted by the investor. Singleton (2004) further expands upon the academic and practical principles of core-satellite strategies and offers the best strategy in gaining higher returns than that of the base index. Later on, Hilary Til (2004) presents to investors a framework for measuring risk when utilizing core-satellite strategies with the main focus of her book being the measurement of risk for hedge funds. She later expanded upon these principles by offering benchmarks for risk which investors should be aware of when implementing a core-satellite strategy. One of the standout studies conducted in the subject of core-satellite strategies is the work of Noël Amenc in 2004 where he introduces a dynamic model in where the allocation of assets to the satellite component is moderate depending on its performance. In other words, if the satellite component is outperforming the index, then more assets are allocated to it accordingly. The main idea is that the accumulation of higher than index performances in the past results in a margin of error for the tracking error, allowing for more risky strategies in the future. On the other hand, if the satellite component underperforms the index, then this approach limits the tracking error and risk of the satellite so that relative performance is guaranteed. This study was expanded upon to show the importance of the use of ETF's (Electronic Traded Funds) in benefitting from dynamic risk management approaches (Amenc, Goltz & Grigoriu, 2010).

We can describe the core-satellite problem as the optimized allocation of assets to the core and satellite components where the resultant portfolio outperforms the base index greatest whilst minimizing the tracking error. We can formulate the problem as follows:

$P = wS + (1-w)C$ *Formula 2*

Where w is the weight of investments in the satellite component (S) and (1-w) is the weight of investments in the core component. In order to calculate the tracking error in relation to base index (B) first we must formulate the excess returns of the portfolio as follows:

$P - B = wS + (1-w)c - B$ *Formula 2-1*

$= w(S-B) + (1-w)(c-B)$ *Formula 2-2*

If the core portfolio tracks the index exactly, we will have B=c, thus the excess returns in relation to the index could be formulated as:

$P - B = w(S-B)$ *Formula 2-3*

Thus, the tracking error in relation to the index can be calculated as:

$TE(P) = \sqrt{var(P-B)} = w\sqrt{var(S-B)} = wTE(S)$ *Formula 2-4*

Meaning that the tracking error of the portfolio will be w times the tracking error of the satellite component.

The next step is to calculate an optimum w for investment in the satellite component in relation to the core. This problem can be solved in the context of a simple mean-variance analysis. The target function is constructed as:

U = E(P-B) – λσ2(P-B) = IR(P) x TE(P) – λTE2(P) *Formula 2-5*

Where IR(P) is the informational ratio of the portfolio P in relation to the base index where:

$IR(P)=\frac{E(P-B)}{\sigma(P-B)}=\frac{E(P-B)}{TE(P)}$ *Formula 2-6*

It should be noted that when the core component tracks the base index exactly, then the information ratio of the portfolio as whole IR(P) is independent of the ratio of assets invested in the core component e and equals the information ratio of the satellite component IR(S). We deduce:

$IR(P) = \frac{E(wS + (1-w)C-B)}{\sigma(wS+(1-w)C-B)} = \frac{wE(S-B)}{wTE(S)} = IR(S)$ *Formula 2-7*

We rearrange the target function as:

U(w) = IR x w x TE(S) – λw2TE2(S) *Formula 2-8*

By calculating the derivative of w, the optimum amount of assets which can be allocated to each component of the portfolio is as follows:

$\frac{\sigma U}{\sigma w}(w^*) = 0 \longrightarrow w^* = \frac{IR}{2\lambda TE(S)}$ *Formula 2-9*

By expanding upon the above analysis in the condition that the satellite component

$S = \sum_{i=1}^{N} wiSi$, is the result of investment in N number of active portfolios (with risk) we can solve the problem on a larger scale. In this case the excess returns of the satellite in relation to the base index equals:

$S - B = \sum_{i=1}^{N} wi(Si - B)$ *Formula 2-10*

And the tracking error of the satellite component can be calculated as:

TE(S) = $(\sum_{i,j=1}^{N} wiwj\sigma ij \;\; - 2\sum_{i=1}^{N} wi\sigma\beta + \; \sigma_{\beta}^{2})$ *Formula 2-11*

Where σij is the covariance between risky assets of the satellite component (Sj & Si), σiβ is the covariance between risky assets and the base index and σβ is the standard deviation of the base index returns.

Index based approaches will undoubtedly continue to grow as one of the cornerstones in future investment and asset allocation strategies and the marriage of these approaches with classical active management styles will also play a huge part in their adoption by institutional and retail investors. Although there are criticisms to this strategy for investment, its importance in today's investment environment cannot be denied. The obvious advantages of these strategies (which include better performance evaluation, lower transactional costs and frequency and flexibility to market environments) will continue to attract managers who are looking to for the most efficient way to which attain market returns (frequently referred to Beta). One of the principles of investment is that investors need to select managers who have the ability to achieve higher returns than that of the market by conducting research and information analysis whilst choosing the most optimum basket of assets. Investors which are unable to select such managers should view index-based products as a suitable alternative. Although some level of inefficiency is a pre-requisite for the selection of active management styles, there is more to it than that. Managers with a sufficient skill level are also required to be able to take advantage of the market inefficiencies.

Over the past thirty years the index based approached has gone through an evolutionary stage which will surely continue. One important factor that should be taken into account however is that investors should not look upon index-based approaches as a complete substitute to active management strategies. It should however be part of a bigger investment strategy which looks to minimize risk through diversification whilst achieving higher returns than that of the market index.

Having identified a suitable investment paradigm for our framework, we will now delve into the second part of the literature review which specifically focuses on the optimisation of the framework and the way that the model will deal with stochastics.

2-8 Optimisation Theory

The continued strive of humans in achieving perfection is a perfect example of optimisation theory. Mankind wants to define the best and achieve it (Beightler, Phillips & Wilde, 1979). However, as it becomes apparent that not all the contributing factors can be recognized and quantified in most cases, the absolute perfect answer is relegated in favor of the most satisfactory solution (Warner, 2002). Also, when judging the performance of others, they are evaluated in relation to others (Goldberg, 1989). Thus, it could be argued that due to humanity's inherent inability to optimize, a special significance is given to betterment.

Beightler et al (1979) believe optimisation to be above improvement and argue that optimisation strategies include the quantitative study of the optimal and how to achieve it. Expanded further it could be argued that optimal as a technical term embodies quantitative measurement and mathematical analysis whereas the term "best" has a lower degree of precision and is mostly used in daily activities rather than a scientific endeavor. In most cases most of what is done in the name of optimisation results in improvements with the aim of reaching an optimum point (Beightler et al, 1979). This definition is comprised of two parts:

1. Improvements in the search of the…

2. Optimum point

The difference between the improvement process and the optimum point should be clear. Due to the mathematical roots of optimisation the implied performance of the process is still considered secondary to its results on convergence (Can the optimum point be reached?). In practice however this lack of focus on the implicit performance is not natural or logical (Goldberg, 1989). This fact should not take anything away from the value of convergence however, as it is an important basis which allows the comparison of different optimisation techniques.

When comparing optimisation algorithms two factors of convergence and performance are studied. Some algorithms may result in convergence but have weak performance, meaning that the process of improvement is slow and inefficient. Likewise, some algorithms may have very good performance but not result in convergence. We can define the objectives of the search process in three main categories:

1. Optimisation

2. Finding a practical solution

3. Pseudo-optimisation

In conditions where we are happy to find a solution neighboring the optimal solution, the target of the search is defined as the pseudo-optimal point. If the objective is to find a suitable practical solution within a defined confine of the optimal point the process is defined as near-optimisation. On the other hand, if we remove the condition of proximity to the optimal point as a factor and only consider the highest probability of finding a solution near to the optimal point as the objective, then the resulting optimisation process is identified as approximation.

Due to the inherent complexity of most practical situations, pseudo-optimisation is applied to balance the cost and quality of a solution. The calculations required to solve compound optimisation problems naturally reach astronomical figures leading to the elimination of optimisation conditions as an economic necessity. Pseudo-optimisation can be utilized to offer algorithms which guarantee a level of balance between the number of calculations needed and also the proximity to the optimal point. The algorithms must have configurable parameters which the user can use to adjust the balance between number of computations and the precision of the resultant solution (Pearl, 1984). Many optimisation techniques have been studied in previous research, with the aim of attaining the optimized solution (or as close to it as possible), some of which we will outline in the next part of our study.

2-8-1 Quadratic Programming

Quadratic programming (QP) is the problem of optimizing a quadratic objective function and is one of the simplest forms of non-linear programming (Frank & Wolfe, 1956). The objective function can contain bilinear or up to second order polynomial terms (Floudas & Visweswaran, 1995) whilst the constraints can be both equalities and non-equalities. Quadratic programming problems are encountered in many real-life scenarios which include some form of linear constraint and have a quadratic cost function such as signal processing, scheduling in chemical plants and portfolio optimisation. The methodology was first pioneered by Princeton University's Frank and Wolfe in the 1950's who developed the theoretical background and was then utilized in the field of portfolio optimisation by Markowitz. A general form of quadratic programming is displayed as:

$$\min f(x) = c\,x + \frac{1}{2}x^T Q x$$

$$s.t.\,Ax \leq b$$

$$x \geq 0$$

Formula 3

Where c is an n-dimensional row vector describing the coefficients of the linear terms in the objective function, and Q is an (n×n) symmetric matrix describing the coefficients of the quadratic terms. If a constant term exists it is dropped from the model. As in linear programming, the decision variables are denoted by the n-dimensional column vector x, and the constraints are defined by an (m×n) A matrix and an m-dimensional column vector b of right-hand-side coefficients. We assume that a feasible solution exists and that the constraint region is bounded (Jensen & Bard, 2015).

Quadratic programs have been widely used since its development in the 1950s because it is a simple type of non-linear programming that can accurately model many real-world systems, notably those dependent on two variables. Problems formulated this way are straightforward to optimize when the objective function is convex.

2-8-2 Robust Optimisation

With the advent of new technologies and techniques associated with IOT (Internet of Things), big data and artificial intelligence, we are now able to capture and process vast varied volumes of data. However, when optimizing big-data driven processes we are usually faced with the problem of data uncertainty. To solve complex optimisation problems, we must first consider that they are generally multivariable problems with inherent limitations and constraints. When considering robustness of an optimisation strategy we are usually determining if the proposed solution is feasible to apply for any parameter scenario and stochastic uncertainty and that this feasibility remains close to the optimality. (García & Peña, 2018).

Each practical optimisation problem suffers from some form of uncertainty in measurements or changes in the environment of the system (also referred to as microscopic and macroscopic uncertainties). In order to solve the optimization problem, we can analyze a nominal scenario which would describe the most typical (or average) case. This however cannot be generalized to all real-world situations and in some cases finding the most typical (or average) case is a problem in itself with a magnitude of difficulty (Jenkins, 2000).

Another form of uncertainty in optimisation problems could be classified as the uncertainty in the optimality of the solution. In other words, it may be the case that depending on the set of uncertainty chosen, the optimality of the solution could be altered. Due to the fact that not all scenarios require the same treatment in respect to optimality, a series of different concepts of robustness have been studied in the literature including: Strict Robustness (Ben-Tal, Ghaoui & Nemirovski, 2009), cardinality constrained robustness (Bertsimas & Sim, 2004), adjustable robustness (Ben-Tal, Goryashko, Guslitzer & Nemirovski, 2004), lightweight robustness (Schöbel, 2014), soft robustness (Ben-Tal, Bertsimas & Brown, 2010) & lexicographic robustness.

If taking into account an optimisation problem with constraints to be formally written as:

$$minf(x)$$

$$s.t.\ F(x) \leq 0$$

$$x \in S,$$

Formula 4-1

Where F:$R^n \rightarrow R^m$ describes a problem of n dimensions with m constraints. F:$R^n \rightarrow R$ is the objective function and S ∈ R^n is the search domain. Supposing that Ɛ ∈ R^K corresponds to a scenario that could occur in our practical problem, we can formalize the uncertainty of the optimisation problem and rewrite the above equation to consider uncertainty scenario Ɛ:

$$min f(x, \Sigma)$$

$$s.t. F(x, \Sigma) \leq 0$$

$$x \in S$$

Formula 4-2

In most practical problems the exact value of Ƹ is unknown, but if it is clear that the problem falls on uncertainty set u $\in R^K$, which itself is a subset of scenarios that are enough to consider, we reach a family of optimisation problems represented by the pair (P(Ƹ),Ƹ∈u). The underlying objective of robust optimisation is to turn this family of problems into a single optimisation problem where the choice of uncertainty subset directly affects the result and complexity of the problem. For the adequate treatment of the problem of uncertainty it is fundamental to give structure to the set (Garcia & Pena, 2018).

2-8-2-1 Robustness Models

This section aims to formally define the main concepts of robustness used to solve optimization problems with uncertainty. In each of the ways to approach robustness, the intuition that exists behind the definition is described; later, the sets that model the uncertainty are characterized and then the problem is written in its robust version.

1. *Strict Robustness*

Considering an instance of x to be a solution to the optimisation problem with uncertainty (P(Ƹ),Ƹ∈u). The solution is strict if x is feasible for all possible scenarios of u. Thus if F(x, Ƹ)≤0 for all Ƹ∈u. This method is considered the most intuitive way of trying to solve optimisation problems in a robust way. Formally, consider the set of all possible strictly robust solutions with respect to the uncertainty set U given by:

$$f(\Sigma) = \{x \epsilon S : F(x, \Sigma) \leq 0\}$$

$$R(U) = \bigcap_{\Sigma \in U} f(\Sigma)$$

Formula 5-1

Then the strict robust problem corresponds to the problem formulated as

$$\min \sup f(x, \Sigma)$$

$$s.t. x \in R(U)$$

$$x \in S$$

Formula 5-2

To our knowledge, the first practical application of the strict robustness methodology was by Soyster in 1973, applying uncertainty to convex sets and solving the problem by utilizing linear programming. His work was further expanded upon by Ben Tal and Nemirovski, offering a theoretical framework where they argued that the essence of strict robustness is that all possible scenarios can occur and all are critical. In practical settings this methodology is most suitable for critical systems where failure is not tolerable such as airplanes and nuclear power plants. However, in the wider spectrum of problems this robustness can be relaxed.

2. *Cardinality constrained robustness*

One method of relaxing the robustness as outlined in the previous section is to restrict the space of uncertainty. This restriction can be achieved through the use of various methodologies however in the case of cardinality constrained robustness it is considered that it is unlikely that all the uncertainty parameters will fluctuate and change at the same time except when the scenario in question is the worst case possible. Thus, we can limit the cardinality of the uncertainty space by varying only some parameters and others are modeled with their representative values (Bertsimas, Sim, 2004).

If we consider X = $\{x1, \dots, xn\}$ and b1x1, … , bnxn ≤ c be a solution and restriction of the optimisation problem then:

u = $\{b \in R^n : \hat{b}_i \in [\hat{b}_i - d_i, \hat{b}_i + d_i], i = 1, \dots n\}$

Thus the cardinality constrained robustness can be described as:

$$\sum_{i=1}^{n} \widehat{b_i}\, x_i + \max_{R \in \{1,\dots,n\}, |R|=y} \left(\sum_{i \in R} d_i\, |x_i| \right) \leq c$$

Formula 6

As conceptualized by Bertsimas and Sim the approach is utilized for continuous problems and extended to combinatorial problems (Goetzmann, Stiller, & Telha, 2012).

3. Adjustable robustness

Another method of relaxing the robustness introduced by Ben Tal is to divide the space of uncertainty into groups of variables. The primary group can be identified as “here and now” variables corresponding to variables that must be evaluated before scenario $\xi \in u$ is determined and “wait and see” variables which are determined once scenario ξ is known.

Let X be one point of our search space; then, X=(u,v) can be divided into u∈S1⊂Rn1 and v∈S2⊂Rn2 where n1+n2=n. Then the variable u corresponds to the “here and now” group of variables and the variables v to the “wait and see” group of variables. Formally, this is written as:

$$minf(u, v, \mathcal{E})$$

$$F(u, v. \mathcal{E}) \leq 0$$

$$(u, v) \in S^1 \times S^2$$

Formula 7-1

In the next step we ensure that for all the $\mathcal{E} \in$ u scenarios, there is v $\in S^2$ such that (u,v) is feasible for $\mathcal{E}$. Let P_{S^1} (F($\mathcal{E}$)), (defined in the formula 7-2) be the projection of F($\mathcal{E}$) over S1.

$$P_{S^1}(f(\mathcal{E})) = \{u \in S^1 : \exists v \in S^2 \ s.t. (u, v) \in f(\mathcal{E})\}$$

Formula 7-2

Where f($\mathcal{E}$) corresponds to the solution space that complies with the constraints defined in equation 3. To represent the set of solutions for the split robustness it is studied that:

$$R = \{u \epsilon S^1 : \ \forall \mathcal{E} \in \mathrm{U} \exists \mathrm{v} \in S^2 \ s.t. (u, v) \in f(\mathcal{E})\}$$

$$\bigcap_{\mathcal{E} \in \mathrm{U}} P_{S^1}(f(\mathcal{E}))$$

Formula 7-3

Given a u, the worst-case w for some specific u with respect to the set of solutions R, is given by:

$$w^R(u) = \sup \inf f(u, v, \mathcal{E})$$

$$\mathcal{E} \in U^{v:(u,v) \in f(\mathcal{E})}$$

Formula 7-4

And finally, the split robustness is formalized as:

$$\min \{w^R(u): u \ \in R\}$$

Formula 7-5

Ben Tal et al, applied the above methodology to uncertain problems in linear programming however the concept has now matured and developed and has been adapted for uses in portfolio selection (Fliedner & Liesiö, 2016), power systems and capacity extension planning (Mejia-Giraldo & McCalley, 2014) and aperiodic timetabling (Goerigk & Schöbel, 2014,).

4. Light robustness

Another method of relaxing the concepts of strict robustness is to focus on the constraints instead of the space of uncertainty. Schöbel, considers the fundamental hypothesis that by solving the optimisation problem in the nominal (or average) scenario adequately, the resulting solution cannot be inherently bad and thus we can then concentrate on finding relatively close solutions of the fitness that also fulfill the constraints of the problem considering all Ƹ∈u. In other words, by relaxing the constraints in favor of the quality of solution, we can relax the concept of strict robustness on an optimisation problem. This new concept called light robustness (Schöbel, 2014) can be formalized as:

$$min \sum_{i=1}^{k} w_i \lambda_i$$

$$s.t. f(x, \hat{Ƹ}\ ) \leq f^*(\hat{Ƹ}) + p$$

$$F(x, Ƹ) \leq \lambda, \forall Ƹ \in \mathrm{U}$$

$$x \in S, \lambda \epsilon \mathbb{R}^k$$

Formula 8

Introduced by Fischetti and Monaci in 2009 as a solution to problems in linear programming, a tradeoff between robustness and quality of solution was gained by adding a constraint with parameter p, forcing the solution to have a certain fitness to the solution for the nominal case represented by $\hat{\Sigma}$. Then the λ factor is implemented to find the best set of coefficients that relax the original constraints keeping in mind the tradeoff between quality and robustness whilst achieving the fit of the nominal case.

The concept was later developed to determine the best route of travel in the German public transport network (Carrisoza et al, 2014) and later generalized to take into account any optimisation problem and any set of uncertainty (Schöbel, 2014).

5. Regret Robustness

As outlined by Kouvelis & Yu (1997), regret robustness attempts to relax the strict robustness of the problem through the analysis of the objective function. Let f*(Ƹ) be the best target value in the scenario Ƹ∈u. Instead of minimizing the worst-case performance of a solution, it minimizes the difference to the objective function of the best solution that would have been possible in a scenario. The regret robustness formulation is shown as:

$$\min sup \ \left(f(x, \Sigma) - f^*(\Sigma)\right)$$

$$s.t. F(x) \leq 0$$

$$x \in S$$

Formula 9

The concept is now utilized in a wide variety of environments including portfolio optimisation problems (Xidonas, Mavrotas, Hassapis, & Zopounidis, 2017), evacuation planning models (Goerigk, Hamacher, & Kinscherff, 2018) and safety investment models (Aven & Hiriart, 2013).

6. Recoverable Robustness

Recoverable robustness utilizes a recovery algorithm to obtain the solution of an optimisation problem in two stages (as is the case in adjustable robustness). Given a family of algorithms u, A solution x is recovery robust with respect to u if it exists for every scenario Ƹ∈U, an algorithm A∈u such that A applied to the solution x and a scenario Ƹ allows you to build a solution A(x,Ƹ)∈F(Ƹ) (Cicerone, D'Angelo, Di Stefano, Frigioni, & Navarra, 2007). The robust form of the optimisation problem can then be written as:

$$\min_{(x,A)\in(f(\Sigma)\times \mathrm{u})} f(x)$$

$$s.t. A(x,\Sigma) \in \mathrm{f}(\Sigma), \forall \Sigma \in \mathrm{u}$$

Formula 10

In the original study conducted by Cicerone and partners, the recoverable robustness was applied to shunting problems and later developed and adapted to railway problems with linear programming (Liebchen, Lübbecke, Möhring, & Stiller, 2009). Today the widespread adoption of the methodology has seen practical instances and use in the fields of location planning (Carrizosa, Goerigk, & Schöbel, 2017), logistics of supply (Cheref, Artigues, & Billaut, 2016), allocation and network design problems (Kutschka, 2016) and transit network design (Cadarso & Marín, 2012), among others.

2-8-2-2 Portfolio Management & Robust Optimisation

The portfolio optimisation solutions can be divided into two categories: absolute and random. Random approaches such as robust optimisation are increasingly utilized to develop suitable strategies in the management and allocation of assets and a significant amount of research has been conducted in the literature, some of which we will now expand upon further.

The available literature depicts various approaches in dealing with uncertainty in mathematical programming including stochastic programming and robust methodologies. Stochastic approaches utilize a decision tree approach to the requisite problem and analyze all possible scenarios leading to difficulty and complexity in finding suitable solutions. This is due to the fact that the dimensions of the solution space expand exponentially as the size of the problem increases. Robust optimisation is not a new technique but new developments in the field has paved the way for dealing with uncertainty in optimisation problems. As we have seen the optimal answer to a linear programming problem lies on an extreme point or the edge of feasible solution area. A small convulsion in the data entered could easily make changes in the end solution or even render it infeasible. Therefore, there is a clear need to analyze uncertainty as an input parameter and also take into account modern techniques to manage the level of uncertainty.

Soyster introduced the concept of robust optimisation in 1973, however many researchers did not pay attention due to the fact that his concepts were so pessimistic in nature (worst case scenario analysis only) that their practical use would not be feasible. Ben Tal and Nemirovsky expanded upon his initial ideas in 1998 to present a methodology that presented a more optimistic solution. Their approach relied on the use of an interior-point mathematical basis where a linear or nonlinear programming method is used to go through the middle of the solid plane defined by the problem rather than around its surface. They then used their model to solve a series of practical real world portfolio optimisation problems to show that the final solution is feasible against range of different uncertain data inputs. Their proposed Ω parameter is used as control mechanism to limit the probability of deviation from nominal constraints. The implementation of their 2002 approach however, normally transforms a general linear problem into a non-linear convex problem, the results being that many users who had become accustomed to conventional optimisation techniques were wary of adopting such an approach. Bertsimas and Sim (2004) developed a variety of different robust optimisation approaches with the objective of maintaining the original problem structure. Although their approach does not have the same level of optimism than that of the Ben Tal methodology, it has garnered much support due to the fact that it attempts to maintain the linear structure of the optimisation problem. Robust optimisation should not be viewed upon as a new science in operational research however the aforementioned approach is an innovation in conical programming which affects robust optimisation in many ways.

Robust Optimisation should be viewed upon as an innovative modeling tool with the ability to analyze the inherent chaos prevalent in many problem parameters of a decision-making process. Generally, the objective of robust optimisation is to find solutions to specified optimisation problems where optimal target values for all alternative non-linear parameters are attained and recently has been utilized to reduce the sensitivity of the optimal portfolio towards statistical errors in estimating the parameters employed in portfolio optimisation. For example, Goldfarb and Iyengar (2003) assume that the return vector follows a linear factor model as in:

$$r = \mu + V^T f + \varepsilon$$

Formula 11

where μ is the expected return vector, $f \sim N(0,F) \in \mathbb{R}^{m}$ is the returns of the factors that are assumed to drive market returns, $V \in \mathbb{R}^{m\times n}$ is the factor loading matrix of the n assets and $\varepsilon \sim N(0, D) \in \mathbb{R}^{n}$ is the model residual. They also continue to introduce uncertainty sets Sv and Sm for v and μ where Sm acts as a box and Sv is a Cartesian product of a number of ellipticals. Their study continues to prove that by utilizing a robust optimisation approach to counteract the inherent confusion and chaos of problem parameters, it is possible to transform the VaR portfolio selection problem into a second order cone program (SOCP), keeping in mind that the uncertain parameter sets are elliptical in nature.

Halldorson and Tutuncu (2003) sought to minimize the objective function in the worst-case realization of the input parameters and expressed their results as saddle-point problem solved in polynomial time. Further developing on their work Tutuncu and Koenig (2004) computed “robust efficient frontiers” by using real market data and found that the robust portfolios offered significant improvement in worst-case scenarios when compared to nominal portfolios (although there were extra costs in expected returns).

El Ghaoui et al (2003) argue that classical formulations of the portfolio optimization problem, such as mean-variance or Value-at-Risk (VaR) approaches, can result in a portfolio extremely sensitive to errors in the data, such as mean and covariance matrix of the returns. Thus, they propose a way to alleviate this problem in a trackable manner assuming that the distribution of returns is partially known, in the sense that only constraints on the mean and covariance matrix are available. They continue to define the worst-case Value-at-Risk as the largest VaR attainable, given the partial information on the returns' distribution. Computation and optimisation of the worst-case VaR are considered and it is finally shown that these problems can be cast as semi-definite programs. The robust VaR problem is implemented on the distribution of returns using market data and shows that conditional VaR problems can be transformed into linear second order cone programs to take into account simple uncertainty as a result of returns distribution. The CVaR approach also known as the expected shortfall, is a risk assessment measure that quantifies the amount of tail risk an investment portfolio has. The measurement of coherent risk can be classified as a trailing risk and thus can be considered an efficient hybrid risk assessment tool which is able to overcome the limitations of VaR.

The structure of the uncertainty set is an important part of the formalization of the robust portfolio selection problem. The literature points to the widespread use of independent uncertainty sets as outlined in the research composed by Tutuncu et al and Goldfarb. For example, Tutuncu and Koenig (2004) use two sets of uncertainty in the form of Sm=$\{\mu: \mu_L \leq \tilde{\mu} \leq \mu_U\}$ & Sv=$\{\sum : \sum_L \leq \sum \leq \sum_U \}$ for the mean of μ & Σ for the asset returns vector r, where A>0 outlines a symmetrical and positive semi-definite matrix.

Huang et al (2008) further develop the worst case VaR approach and formalize relatable problems as semi-definite programming problems. DeMiguel et al (2009) provide a general framework for finding portfolios that perform well out-of-sample in the presence of estimation error, solving the traditional minimum-variance problem subject to the additional constraint that the norm of the portfolio-weight vector be smaller than a given threshold. In their approach the robust estimation and optimisation of the portfolio is solved using an independent non-linear program.

Quaranta and Zaffaroni (2008) extended a robust optimization of conditional value at risk (CVaR), minimizing the CVaR of portfolio assets. They studied the implementation of double factor minimization models alongside robust linear models.

Up to now, various definitions of risk have been offered and formalized in mathematical terms. The definition of risk is mostly dependent on the viewpoint of the actor. The most famous definitions of risk are: Standard deviation (Markowitz, 1952), Safety first (Roy, 1952) and semi-variance (Markowitz, 1959). In practical terms some actors may be worried about specified loss levels and the probability of them happening, whereas others are interested in all levels of loss applicable and their probabilities. Huang (2008) offered a new definition of risk where ξ is the random return of a portfolio and b is the target returns of the investor. Then they continue to state that $f(r) = pro\{(b - \xi) \geq r\}, \forall r \geq 0$ is the portfolio investment risk where r outlines the severity of losses. The bigger the value the more severe the loss. The curve of $f(r)$ is the possibility that return ξ is less than target returns b by r amount. This is similar Roy's (1952) safety first approach where he argued that by setting a minimum required return for a given level of risk, investors could compare potential portfolio investments based on the probability that the portfolio returns will fall below their minimum desired return threshold. Some investors feel that Roy's safety-first criterion is a risk-management philosophy in addition to being an evaluation method. By choosing investments that adhere to a minimum acceptable portfolio return, an investor can sleep at night knowing that their investment will achieve a minimum return, and anything above that is surplus gains. The difference in Huang's (2008) approach is that he surmised that investors are worried about all possible outcomes (and not a singular bad outcome). For example let us consider that $\xi \sim N(\mu, \sigma^2)$ where μ and σ are real numbers which identify the normal distribution of stochastic returns of a portfolio. In this case the risk curve can be formalized as:

$$f(r) = pro\{(b - \xi) \geq r\}, \forall r \geq 0$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{b-r} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx, r \geq 0$$

Formula 11

Consider that ξ is the return of the portfolio.

Benchmark 1: According to Markowitz's (1952) standard deviation benchmark we say that A is secure id V[ξ] ≤ a, where V is the variance operator and a is the maximum variance that the investor can stomach.

Benchmark 2: According to Markowitz's (1959) semi-variance benchmark it is considered that A is secure if SV[ξ] ≤ b where SV is the semi-variance operator and b is the maximum semi-variance given that the investor can bear.

Benchmark 3: In Roy's (1952) safety first approach A is considered secure where $pro\ \{(b - \mathcal{E}) \geq r\} \leq \alpha_0$, where b is the target returns of the investor, r is the possible level of losses and $\alpha(r)$ is the assurance curve given. Huang's (2008) approach is a generalization of the Roy's benchmark where it is theorized that $\mathcal{E}_i$ represents the stochastic returns of the ith asset. The annual returns are formalized as $\mathcal{E}_i = \frac{\acute{p}_i + d_i - p_i}{p_i}$ where $\acute{p}_i$ is the final price of the asset, p_i is the previous year's final price of the asset, and d_i is the dividend paid through the course of the year. Consider that b is the target returns of the investor and x_i is the level of investment in each asset, we can calculate the loss of this investment as $b - (\mathcal{E}_1 x_1 + \mathcal{E}_2 x_2 + \cdots + \mathcal{E}_n x_n)$. Consider that r is the loss severity index and $\alpha(r)$ is the assurance curve of the investment. Thus, the portfolio selection problem can be formulized as:

$$\text{Max E } (\mathcal{E}_1 x_1 + \mathcal{E}_2 x_2 + \cdots + \mathcal{E}_n x_n)$$

$$\text{s.t.}$$

$$\text{pro}\{(\ b - (\mathcal{E}_1 x_1 + \mathcal{E}_2 x_2 + \cdots + \mathcal{E}_n x_n)) \geq r\} \leq \alpha_0,$$

$$x_1 + x_2 + \ldots + x_n = 1,$$

$$x_1 \geq 0$$

Formula 12

If r is depraved to the fixed value of r_0 then the above model is the same as Roy's (1952) model, however Huang (2008) argues that is the returns distribution is not simple then the above problem becomes complicated and he utilizes a genetic algorithm methodology to solve it.

Bertsimas and Pachamanova (2008) consider the multi-period portfolio optimization problem suggesting robust optimization formulations of the multi-period portfolio optimization problem that are linear and computationally efficient whilst proving advantageous when complex additional requirements need to be imposed on the portfolio structure, e.g., limitations on positions in certain assets or tax constraints.

Moon & Yao (2011) further develop the concepts of robust portfolio optimisation by considering parameter uncertainty through the control of the impact of estimation errors on the portfolio strategy performance. They construct a simple robust mean absolute deviation (RMAD) model which leads to a linear program and reduces computational complexity of existing robust portfolio optimization methods, testing their model using real data consisting of 100 randomly chosen stocks from the AMEX, NASDAQ and NYSE indexes. Their results showed that the proposed robust optimization generally outperforms a nominal mean absolute deviation model whilst offering some precautions in circumstances where robust optimisation would not be considered. In their study Moon et al (2011) argue that despite most of literatures concluding that a robust portfolio optimization outperforms a nominal one, that sometimes the robust portfolio might not be successfully constructed, concluding that a robust optimization approach could result in poor performance under such certain conditions where parameters are estimated over long time horizons in a declining market and that for a low beta sorted portfolio, the difference between a robust and a nominal approach is very small.

Zymler, Rustem, & Kuhn (2011) tackled the robust portfolio optimisation problem by considering derivatives contracts. The main objective of robust portfolio optimisation is to maximize the worst-case portfolio returns in circumstances where asset returns are permitted to fluctuate in a predetermined uncertainty set. If the uncertainty set is not too large then the resulting portfolio will perform well in normal market conditions. However, in bad market conditions (market crash) the resulting portfolio will have a much degraded performance (due to the asset returns materializing far outside of the uncertainty set). Thus, Zymler et al (2011) propose a novel robust optimization model for designing portfolios that include European-style options, trading off weak and strong guarantees on the worst-case portfolio return. The weak guarantee refers to circumstances when the asset returns materialize within the predetermined uncertainty set and the strong guarantee refers to circumstances where all possible outcomes of the market are analyzed. The resulting model is a convex second-order cone program which is tested against real world and simulated data.

Gregory, Darby-Dowman, & Mitra (2011) approach the problem in another way to evaluate the cost of robustness for the robust counterpart to the maximum return portfolio optimization problem. The uncertainty of asset returns is modelled by polyhedral uncertainty sets as opposed to the earlier proposed ellipsoidal sets deriving the robust model from a min-regret perspective and examining the properties of robust models with respect to portfolio composition. Robustness can be viewed upon as a guarantee of portfolio performance and as with any guarantee there will be costs attached. In robust portfolio optimisation these guarantees means that the resultant portfolio returns will have the same performance than that of the robust solution. Gregory et al (2011) conclude that their proposed model is best applied in situations where parameter values are unknown, variable, and their distributions are uncertain. In the case when distributions can be precisely estimated one should consider other methodologies such as stochastic programming.

Huang, Zhu, Fabozzi, & Fukushima (2010) also conducted research on the matter, considering the relative robust conditional value-at-risk portfolio selection problem where the underlying probability distribution of portfolio return is only known to belong to a certain set, the resulting model takes into account the worst-case scenarios of the uncertain distribution, but also pays attention to the best possible decision with respect to each realization of the distribution (Huang, Zhu, Fabozzi & Fukushima, 2010). The study then continues to show that a robust optimized model can be formalized using a sequence of linear models or a second-order cone program which allows for multiple experts (priors) to be utilized in the decision-making process. In order to familiarize ourselves with the CVaR model, consider y to be the uncertain returns vector and x to be vector of portfolio component weights, then the loss function can be formalized as$f(x, y) = -x\acute{y}$. By assuming that x is constant and the distribution of y is known, the probability that the loss is more than threshold Σ can be formalized as:

$$\varphi(x, \Sigma) = \int_{f(d,y) \leq \Sigma} \pi(y) dy$$

Formula 13-1

By assuming confidence level α, the value at risk (VaR) associated with portfolio x is defined as:

$$VaR_{\alpha}(x, \pi) = min\{\Sigma \in R: \varphi(x, \Sigma) \geq \alpha\}$$

Formula 13-2

Rockafellar and Uryasev (2000) define the conditional value-at-risk (CVaR) as the conditional expectation of the loss of the portfolio exceeding or equal to VaR i.e.:

$$CVaR_{\alpha}(x, \pi) = \frac{1}{1-\alpha} \int_{f(x,y) \geq VaR_{\alpha}(x,\pi)} f(x, y) \pi(y) dy$$

Formula 14

Also Guastaroba et al (2011) studied the effectiveness of different robust optimisation techniques on the portfolio optimisation problem. By using the methods conceived by Bertsimas and Ben Tal, they managed to solve the CVaR portfolio optimisation problem and tested their system by using real world data from the London stock exchange, evaluating each proposed approach against different market conditions. They concluded that in general, the nominal and the robust portfolios tend to have a similar behavior, with a generally better performance of the Ben-Tal and Nemirovsky robust portfolios (Guastaroba, Mitra & Speranza, 2011).

Fastrich and Winker proposed a new approach to solving the robust portfolio optimisation problem 2012 where they used an amalgamation of techniques pioneered by Tutuncu & Koenig (2004) (robust optimisation with regards to correlation) and Ceria and Stubbs (2006) (independent robust optimisation) to propose a more general version of Markowitz's model where real constraints of the market (in the form of uncertainty sets) could be taken into account. Because of the complexity caused by real integer constraints, Fastrich and Winker use an innovative hybrid approach to solve the problem using German DAX100 data inputs.

2-8-3 Machine Learning

Coined in 1959 by Arthur Samuels, Machine Learning is the scientific study of statistical models and algorithms that computers use to perform certain tasks without the use of explicit instructions, instead relying on patterns and inference. Using sample data (known as training data), machine learning algorithms build a mathematical model, in order to make predictions or make decisions without being explicitly programmed to perform the task (Koza, Bennett, Andre, & Keane, 1996).

Closely related to computational statistics, machine learning is utilized in a wide variety of circumstances where it is difficult or infeasible to develop a conventional algorithm to perform the task. Sometimes referred to as predictive analysis in practical business settings, machine learning is used in conjunction with mathematical optimisation techniques in order to solve complex problems in situations where conventional methodologies are unable to process the vast amount of information and data. Many learning problems are formulated as minimization of some loss function on a training set of examples. Loss functions express the discrepancy between the predictions of the model being trained and the actual problem instances. however, it should be noted that there is a distinct difference in their generalization. Whilst optimisation algorithms generally minimize the loss on the training set, machine learning algorithms are concerned with using their training to minimize the loss on yet unseen data sets (Le Roux, Bengio, & Fitzgibbon, 2012).

Mohri, Rostamizadeh & Talwalkar (2012) argue that a core objective of a learner is to generalize from its experience. In other words, it is the ability to perform accurately on new, unseen data based on the experiences of the learning data set. The training examples come from some generally unknown probability distribution (considered representative of the space of occurrences) and the learner has to build a general model about this space that enables it to produce sufficiently accurate predictions in new cases.

In summary Artificial intelligence can be any technique that enables computers to mimic human behavior. Machine learning is considered a subset of artificial intelligence that actually focuses on teaching an algorithm on how to take information and without explicitly being programmed to do so, learn the sequence of patterns from the data itself. Taking things one step further, deep learning is a subset of machine learning where the system extracts patterns automatically from raw data without human intervention (No need for a human to annotate the necessary rules or relationships).

2-8-3-1 Genetic Algorithms

Since the 1960's researchers have looked to emulate evolutionary processes in complex algorithms as a solution for solving difficult optimisation problems. Genetic algorithms were first introduced by John Holland (Michigan University) and the approach was further developed by Rechenberg, Schewefel, Fogel and Koza.

Optimisation techniques based on the evolutionary process are significantly different to standard methodologies in the fact that whilst in standard practices every new solution candidate is only selected if they improve the objective function, in evolutionary algorithms all solution candidates are given a chance at selection.

Genetic Algorithms are one of the most important innovative algorithms which are utilized in the optimisation of different functions. The algorithms utilize past information extracted from the algorithms (due to their hereditary nature) in the search for an optimal answer.

In his pursuit to understand how systems adapt to their surroundings, Holland first offered an initial concept of genetic algorithms which was then further extrapolated by Goldberg. Holland's version involved a simulation of Darwinian 'survival of the fittest,' as well as the processes of crossover, recombination, mutation, and inversion that occur in genetics. In his 1975 book "Adaptation in Natural and Artificial Systems", Holland presents genetic algorithms as "an abstraction of biological evolution" whilst giving a theoretical framework for the utilisation of genetic algorithms thus becoming the first to propose a foundation for computational evolution. In the same year one of Holland's doctorate students, Kenneth De Jong, presented the first comprehensive study of the use of genetic algorithms to solve optimization problems as his doctoral dissertation. His benchmark examples have become the gold standard for all research on genetic algorithms.

Computer Scientist quickly found that the unique nature of genetic algorithms meant that they could be used to find solutions to problems that other methodologies could not solve. Their ability to simultaneously test many points from all over the solution space, optimize with either discrete or continuous parameters, provide several optimum parameters instead of a single solution and be robust enough to work with a wide range of data types allowed genetic algorithms to produce stunning results when traditional optimisation methods (calculus based, exhaustive search and random) failed miserably.

Goldberg (1989) further developed Hollands work by arguing that genetic algorithms had a distinct set of differences to standard search and optimisation techniques summarized as:

1. Genetic Algorithms works on a premise of coded answers and not the answers themselves.

2. Genetic algorithms searches in a domain of solutions and not a singular solution.

3. Genetic Algorithms implement a fitting function and not derivatives or other supplementary shapes.

4. Genetic algorithms adhere to probabilistic inheritance rules rather than absolute inheritance.

2-8-3-1-1 Advantages of Genetic Algorithms

When considering a calculus-based optimisation technique there are two approaches, direct and indirect. The direct method follows the objective function gradient towards a local maximum or minimum value and is sometimes referred to the gradient ascent/descent method. In the indirect approach the gradient of the objective function is set to zero and then the resultant set of equations is solved. In exhaustive search algorithms an exhaustive search is performed requiring a finite search space or a discretized infinite search space of possible values for the objective function to be tested against. All values in the search space are tested to find the optima, resulting in a very vast search space which in turn is very inefficient. Random search algorithms became increasingly popular over exhaustive search approaches by choosing some representative sampling from the search space and then attempting to find the optima in that sampling. Both these types of algorithms essentially mean that the proximity of the sampling to the optimal point is completely left to chance.

Although these approached have been studied extensively and improved significantly, they still only search for local optima. In a situation where the global optima in a solution space is unknown, they are rendered useless. By progressing from a population of candidate solutions instead of a single value, the likelihood of finding local optima instead of the global optima is greatly reduced.

Secondly traditional calculus-based techniques rely on the use and formalization of derivatives which is never the case in practical applications. Genetic Algorithms only rely upon the numerical fitness value of the candidate solutions (based on the creator's definition of fitness) to guide their search making them suitable for problem applications with multiple local optima. This allows genetic algorithms to function in a noisy, nonlinear search space where derivatives do not even exist whilst allowing designers to adjust them in each practical situation based on the accuracy vs efficiency needs of the problem.

The designer may make more decisions on the objective function, sample size, number of iterations, constraints and restrictions, ending conditions and solving methodology. Each choice affecting the accuracy of the model. Overall, the important thing to note is that a genetic algorithm is a solution domain independent approach which efficiently searches the search domain for points of better fitness.

2-8-3-1-2 Disadvantages of Genetic Algorithms

Like any other optimisation approach, genetic algorithms also suffer from a series of limitations. One limitation is that although traditional techniques may not be applicable to all problem sets, they may perform better in specific problems (with a higher degree of precision) than their genetic algorithm counterpart. In other words, although genetic algorithms perform with a high level of efficiency in finding acceptable solutions, they do not guarantee the resultant solution will be the global optima. This has been studied by various researchers with results showing that although the global optimal is not guaranteed, the probability of finding the global optima in genetic algorithms is relatively high and secondly that the resultant solution in most cases is found to be neighboring the global optima.

2-8-3-1-3 Overview of Natural Genealogy

In the natural environment the evolutionary process is triggered once the following four preconditions are met:

1. A being possesses the ability to reproduce itself.

2. A population of above beings exist.

3. The population is diverse.

4. Members of the population are distinguishable from others according to some parameter or ability

John Holland proposed an overall framework for all adjustable systems portraying how evolutionary processes could be utilized in artificial systems. Generally, any adjustable problem can be coded in genetic terms. In other words, any problem that can be formalized as genetic terms can be solved using genetic algorithms.

GA are mathematical algorithms which use reproduction, natural selection (survival of the fittest) and natural genealogy models to transform a set (population) of individual mathematical objects (usually fixed length character strings as chromosomes) to a new population (e.g., next generation) with specific fitness. All anatomical and behavioral information of the objects are coded into the chromosome and in mimicking the natural transformations and evolutions of chromosomes, genetic algorithms attempt to achieve the expected results through multiple generations of inheritance and mutations in the population individuals. This objective is achieved by creating a preliminary population and creating the necessary conditions for evolution in next generations.

By studying their natural counterparts we can deduce the following notions on genetic algorithms:

1. Just as all inherited characteristics of creatures are encoded in their chromosomes by genes, every possible solution can also be encoded by numerical strings in a binary system. A collection of these possible solutions can be regarded as solution candidates in the preliminary population.

2. The survival capability of each string in the population is measured against the degree of fitness of that string against problem conditions. In genetic algorithms the problem and its conditions are tantamount to the natural environment in evolution theories. The problem conditions and information are formalized into an assessment function where each population string is allocated as a function variable resulting in the output of a number which shows the survival rate of that string in respect to the problem conditions (environment). As the survival rate increases (fitness increases) so does the probability of the string being present in the next generation. On the hand if the survival rate is low then that means that there is a low possibility that the selected string will be present in the next generation. As a result, the next generation will be made up of strings that have displayed a higher level of fitness whilst those that resulted in lower fitness are excluded.

3. Crossovers are one of the factors which lead to change in creatures and result in different specimens (diversity) within a population. In natural crossovers, matching chromosomes are bound together exchanging genes. Crossovers in genetic algorithms are used in order to create diversity in the population strings and allow for the passing of information between strings. First the population strings are matched together in groups of two and then spliced into two parts at a random point. Each corresponding half is then exchanged with the matched string.

4. Another factor contributing to diversity is mutation. Emulating the natural genetic mutation, mutation in genetic algorithms on a string is achieved by choosing a random

number between 1 and L (length of string) and then naming the number K. Then the corresponding key number K is selected in the string, substituting 0 for 1 and vice versa.

In the mathematical model of the genetic approach, each genetic operation is simulated and their amalgamation results in the natural selection optimisation process.

Thus after a few iterations and generation of multiple populations, the most appropriate population which will be the optimum solution to the problem is produced.

Natural Systems	**Genetic Algorithm**
Hereditary characteristics are encoded on chromosomes	Possible solutions to the problem are encoded onto numerical strings in binary
Environment	Fitness Function Problem is formalized as mathematical relationship
Natural Selection Generally the measure for survival of a creature is its ability to adapt to the environment	Reproduction Every string is considered as a function variable. The resultant fitness is calculated and based on them strings for the new population are selected.
Crossover Chromosome matching and splicing genes between chromosomes	Crossover Population strings are matched in twos and spliced at a determined location.
Mutation Replacing a gap with another gap in the DNA chain	Mutation Random Selection of string key and exchange of value.
Evolution and creation of new generations	Repetition of reproduction steps onwards

Figure 8: Natural Genetic Operations & their mathematical Counterparts

To summarize, the genetic algorithm searches the solution domain to find the string resulting in the highest level of fitness resulting in a solution which through generations of evolution satisfies the objective function and constraints.

2-12-4-4 Components of a Genetic Algorithm

Since genetic algorithms are designed to simulate a biological process the basic components to almost all of them are:

- A fitness function for optimisation
- A population of numerical strings (chromosomes)
- Selection of which chromosomes will reproduce
- Crossover to produce next generation
- Random mutation of chromosomes in new generation

The first step in implementing a genetic algorithm is to devise a mechanism to transform every solution to a chromosome. Then an initial population is produced containing a set of the chromosomes, each representing a possible solution. The initial population size is defined by the user and is usually selected randomly.

In the next step new chromosomes known as offspring are generated via the use of genetic operations (crossover and mutation). In order to select the parent chromosomes a crossover rate and mutation rate are utilized by the user.

After the generation of new offspring, the most suitable chromosomes are selected using an evolution operation. The evolution operation handpicks suitable chromosomes from the offspring and parents in such a way that their number is equal to the initial population. The selection process itself is based on the fitness value of each chromosome. In fact, the evaluation process is the most pivotal factor in the selection process.

After all the above steps have been taken it could be said that a single iteration has been processed resulting in a new generation of the algorithm. After many iterations the algorithm will gradually converge towards the optimized solution. The stop condition of the problem is defined as the number of iterations (generations)(that the algorithm will go through and is predetermined by the user.

We can write the pseudo code for a genetic algorithm process as:

Step 1: Set t=0

Step 2: Generate initial population, p(t)

Step 3: Evaluate p(t) to create fitness value

Step 4: While (not terminated) do{

Step 5: Recombine p(t) to yield c(t), selecting from p(t) according to the fitness values

Step 6: Evaluate c(t)

` Step 7: Generate p(t+1) from p(t) and c(t)

Step 8: set t=t+1

}

Step 9: End

Where p(t) is the parents of generation t and c(t) is the offspring of generation t.

When designing a genetic algorithm, the following factors are highlighted as being significant:

1. Introducing a coding system for the chromosomes

2. Initializing the population

3. Genetic Operands definition

4. Fitness function

2. Constraint Handling

2-8-3-1-3-1 Coding System

As discussed previously a fundamental concept in the implementation of genetic algorithms is the representation of problem solutions as a chromosome. We can categorize this process into string and none string coding. In string coding practices, the objective is to transform problem solutions into a string of numbers (sometimes in binary format). If the string is composed of other non-binary elements it is known as a non-binary string which is rarely used in optimisation and is exclusively used in very specific cases. The number of bits which are used to encode the variable is dependent on solution accuracy, range of parameter fluctuations and the relationship between parameters. A string of bits which are an encoded representation of the problem solution is known as a chromosome. The bits in a chromosome act in the same way genes do in a natural environment. One of the main characteristics of genetic algorithms is that they alternate between the coding domain and solution domain. Genetic operations are applied to the encoded domain whilst selection and evaluation are applied to the solution domain.

2-8-3-1-3-2 Population Initialisation

After selecting a technique to encode each solution as a chromosome, the genetic algorithm model requires that an initial population of chromosomes is produced. In this step the initial solution is usually generated randomly although in specific cases innovative approaches can be used to increase efficiency and speed up convergence towards an optimal solution.

2-8-3-1-3-3 Genetic Operations

The genetic operands in a genetic algorithm mimic the transfer of genes for the reproduction of offspring in each generation by parent chromosomes. The two main types of genetic operands are crossovers and mutation. Generally, the genetic operands are defined based on the type of problem and are experimental in nature. The efficiency of these operands in reaching an optimal solution varies depending on the problem itself. Some operations only target a single gene and its residual information in order to create a new chromosome whilst other operations target multiple genes or even the entire chromosome. Based on their modus operandi they are categorized as:

Mutation:

Operators which select one or more genes in a chromosome and change their values. These operators target one or multiple points in a fixed length string and change the characters at those points. Factors which are of importance in these operations include:

1. Number of points to be changed

2. Point selection method

3. Mutation method

By defining the three factors above we are able to construct a specific mutation operator where information from one solution is used to create another solution. Depending on the amount of information used the mutation is considered small or extensive. The more extensive the mutation the less previous information has been used. Thus, it is argued that more extensive mutation (the less previous information utilized) will lead to more randomness in the resultant solution which is considered beneficial in introducing new genetics into the population. When the population is converging on a specific solution the probability of mutation needs to be increased to prevent convergence on a local optimal, and inversely when a population has varied solutions the mutation probability needs to be lowered in order for it to converge. Based on this, the mutation probability can be set to fluctuate between the two values. Gen & Cheng (1999) argue that in order to find the optimum point of mutation a trial-and-error approach should be utilized.

There are various mutation methodologies which can be utilized, some of which we will outline below:

- Swap

In this method two genes are selected randomly and substituted. For example

If we consider chromosome P to be P: 1 2 3 4 5 6 7 8

And the two genes selected randomly to be the 3rd and 6th genes then the swapping mutation would result in:

P: 1 2 6 4 5 3 7 8

- Inversion

In this method we consider a chromosome with length L and two random point are selected in [1,L], dividing the chromosome into 3 parts. The middle segment is then inversed. For example, in chromosome P: 1 2 3 4 5 6 7 8

If we consider the inversion segment as [2,5] then the result of the mutation would be:

P: 1 2 5 4 3 6 7 8

- Insertion

In the insertion method of mutation, a gene is selected randomly and moved to a new position within the chromosome, for example if we consider P: 1 2 3 4 5 6 7 8 and the randomly selected gene is 6 with an insertion point of 4 then the result of the mutation would be

P: 1 2 3 6 4 5 7 8

Crossover:

Crossover operations handle the transfer of genes between possible solutions. In a crossover operation one or multiple points of two or more possible solutions are selected and then substituted resulting in the creation of new solution. One important concept of crossovers to consider is the crossover rate which defines the ratio of new offspring created in relation to the population effectively regulating the number of chromosomes that are changed in the crossover operation. A higher crossover ratio will allow a larger portion of the solution space to be searched, however if the crossover ratio is too high it will lead to inefficiency as it could cause the algorithm to be guided towards undesirable solution space instead.

Below we will discuss some examples of crossover operations.

- One-cut Point

Proposed by Davis (1991) a single-point is selected randomly on both parent chromosomes and the chromosomes are divided on that point and then the corresponding pieces are substituted.

If we consider Parent A as A: 1 1 0 1 0 1 1 0 and Parent B as B: 0 1 0 0 1 1 0 0 select the point after the 3rd gene as the crossover point then the resultant offspring would take form:

A: 1 1 0/ 1 0 1 1 0 Offspring$_A$: 1 1 0 0 1 1 0 0

B: 0 1 0/ 0 1 1 0 0 Offspring$_B$: 0 1 0 1 0 1 1 0

- Two-cut Point

The two-point crossover operation is more complex in the way that genes from each chromosome are substituted. Two points are selected randomly on each parent chromosome effectively dividing each into 3 segments. The middle segment on each chromosome is kept while the other two segments are substituted. Continuing with the above example of parents A and B the resultant offspring from a two-cut point crossover would be:

A: 1 1/ 0 1 0/ 1 1 0 Offspring$_A$: 0 1 0 1 0 1 0 0

B: 0 1/ 0 0 1/ 1 0 0 Offspring$_B$: 1 1 0 0 1 1 1 0

There also exists a third variation of the crossover function known as a Multi-cut Point approach which behaves as the two-cut point technique but with more than two cut points.

Selection Operation:

Selection is the stage of a genetic algorithm in which individual genomes are chosen from a population for later breeding. In other words, the selection operation guides the algorithm towards hopeful solution spaces. Parent selection is very crucial to the convergence rate of the GA as good parents drive individuals to better and fitter solutions. However, care should be taken to prevent one extremely fit solution from taking over the entire population in a few generations, as this leads to the solutions being close to one another in the solution space thereby leading to a loss of diversity. This taking up of the entire population by one extremely fit solution is known as premature convergence and is an undesirable condition. There are various different selection operations some of which we will outline below:

- Roulette Wheel Selection

Proposed by Holland (1975), the roulette wheel selection (also known as fitness proportionate selection) method is one of the most suitable random selection approaches based on selection probability. The probability of selection of each chromosome is calculated based on its fitness which can be formalized as:

$$P_k = \frac{f_k}{\sum_{i=1}^{k} f_i}$$

Formula 15

Where f_k is the fitness of the k^{th} chromosome and n is the population size. The chromosomes are ordered by P_k and a sum of all probability values is calculated and assigned as q_k.

$$q_k = \sum_{i=1}^{k} P_i$$

Formula 16

Every individual can become a parent with a probability which is proportional to its fitness. Therefore, fitter individuals have a higher chance of mating and propagating their features to the next generation. Therefore, such a selection strategy applies a selection bias to the more fit individuals in the population, evolving better individuals over time.

Consider a circular wheel (Roulette wheel). The wheel is divided into n pies, where n is the number of individuals in the population. Each individual gets a portion of the circle which is proportional to its fitness value. A fixed point is chosen on the wheel circumference and the wheel is rotated. The region of the wheel which comes in front of the fixed point is chosen as the parent. For the second parent, the same process is repeated.

- Tournament Selection

Tournament Selection is another selection method in genetic algorithms where randomly chosen chromosomes from the population are pitted against each other in a tournament of fitness. The winner of each tournament (the one with the best fitness) is selected for crossover. By changing the size of the tournament, we can adjust the probabilistic measure of a chromosome's likelihood of participation in the tournament.

- Remainder Stochastic Selection

A variation to roulette wheel selection is known as 'Remainder Stochastic Selection with Replacement'. The fitness probability of individuals is multiplied by the population size to calculate the number of times the individual will reproduce in the mating pool, ie the expected number of copies. The expected number of copies is a fractional number. An exact fraction of the expected number of copies of the individual is sent to the mating pool. It is also determined whether it can go back to the mating pool for the fraction part (Brindle, 1981). According to Goldberg (1989), remainder stochastic sampling with replacement has a greater probability of diversity in the population than the roulette wheel technique

Fitness Operation

As discussed previously one of the main concepts in Genetic Algorithms is the selection of the most suitable chromosomes. Thus, the need for some measure of suitability is of the upmost importance. In optimisation problems the objective function is considered the measure of suitability or fitness, meaning that each chromosome is fed into the objective function and the corresponding result is assessed for fitness. The fitter the result of the objective function, the more suitable the solution chromosome. In some complex problems we will need to define the fitness function with various strategies, some of which we will discuss below:

- Constraint handling strategy

An important factor in the implementation of Genetic Algorithms is the methodology which is used to handle problem constraints. If left unchecked the genetic operations could result in the production of undesirable chromosomes. Michalewicz (1995) presents a series of standard techniques in approaching the constraint handling problem:

1. Modification of Genetic Operations

One method of countering constraints is altering the definition of the genetic operations in a way which ensures the production of desirable chromosomes after having undergone transformation. This methodology has some issues in that it increases complexity and the alteration of the genetic functions is a difficult proposal and differs on a case-by-case basis.

2. Rejection Strategy

In this approach each resultant chromosome is tested for desirability and rejected if set criterion are not met, making it a simple and efficient methodology to implement.

3. Repairing Strategy

Instead of rejecting an inappropriate chromosome, this strategy transforms it into a suitable chromosome. Like the first strategy described above, this strategy is problem specific and finding a suitable transformation model could be difficult.

4. Penalizing Strategy

This strategy differs from the previous approaches discussed above in that it includes undesirable chromosomes to some extent. Previous techniques focused fully on only the desirable solution space, ignoring the fact that in optimisation problems undesirable solutions make up a large portion of the population. This results in an inefficient and complex approach to optimizing the solution. The penalizing strategy operates by first ignoring the problem constraints completely, for each violation of constraints a penalty is considered and placed within the objective function. The main complexity in this approach is assigning a suitable value for penalties which could direct the algorithm towards the desirable solution. A fundamental principle in this approach is that undesirable solutions are not simply relegated as they could contain genes which with little alteration could lead to the optimum solution. This concept is illustrated in the diagram below.

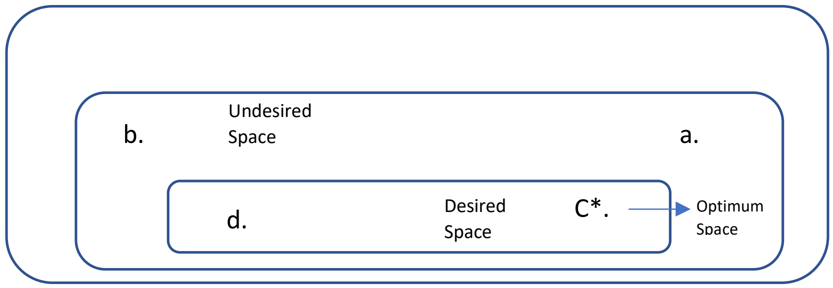

Figure 9: Optimal Solution Search Utilising Penalization

As we can see above points c and d are considered desirable solutions and points a and b undesirable. Point C is considered the optimum answer and thus the penalty given to point a should be lower than the penalty given to point b (resulting in less distance between a and c when compared to b and c). Thus, it is inferred that when considering the optimum point c point a contains more useful information than point d (which is again highlighted by the closeness of point a to point c) which would have been totally disregarded if a different strategy was utilised (Haupt & Haupt, 2004).

2-8-3-2 Artificial Neural Networks

Three important factors which are of part of the underlying importance of Artificial Neural Networks can be summarized as below:

1. Big data: We live in the age of big data and large data sets have become prevalent in today's society. With the technological advances made in the field of memory and chip technologies it has become easier and exponentially cheaper to collect and maintain data. The amount of raw data needed to act as training data for these models was previously consider infeasible or in some cases was not even available.

2. Neural Network models are considered massively parallelizable at their core lending well to processing by modern specialized graphics processing unit (GPU) technologies developed in modern times such as NVidias' CUDA and OPENCL from AMD.

3. Open-source toolboxes such as TensorFlow have radically improved upon the implementation models and techniques available to coders and developers, streamlining development and deployment while also acting as sandbox environments which lend themselves well to testing.

The elementary building blocks of the human nervous system are known as neurons. Similarly, the underlying components of an artificial neural network are also known as neurons based on Rosenblatts single layer perceptron (Izenman, 2009). The idea of a perceptron or a single neuron is that for a vector of multiple real valued inputs X=$(X_1, \ldots.., X_r)^T$ and a single output Y. The connection between input value X_i and an output Y is indicated with a connection weight W_i. The output is is then obtained by computing the activation value U as the sum of X, with their respective weights in the vector $W = (W_1, \ldots, W_r)$ and a bias term W_0 (which allows us to shift our activation function):

$$U = g\left(W_0 + \sum_{i=1}^{r} X_i W_i\right)$$

Formula 17

Where g is the activation function. Using linear algebra, we can rewrite the above formula using vectors, dot products and matrices as:

$$Y = g(W_0 + X^T W)$$

$$\text{Where } X = \begin{bmatrix} X_1 \\ \vdots \\ X_r \end{bmatrix} \text{ and } W = \begin{bmatrix} W_1 \\ \vdots \\ W_r \end{bmatrix}$$

Formula 18

From the input to each neuron, an output is generated through a transfer function known as the activation function (Du & Swamy, 2014). The purpose of activation functions is to introduce non-linearities into the network. This is extremely important in deep learning because in real life data is almost always very non-linear. Linear activation functions produce linear decision boundaries no matter the network size or depth. These functions can collapse an infinite input to a finite output, allowing us to approximate arbitrarily complex functions and likewise draw arbitrarily complex decision boundaries in the feature space. There are three popular activation functions:

1. Sigmoid Function: The sigmoid function is useful for modelling probabilities, collapsing the input to between (0,1). Since probabilities are also modelled between 0 and 1 this makes sigmoid functions perfect for predicting probability distributions at the end of the neural network.

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$g\acute{}(z) = g(z)(1 - g(z))$$

Formula 19

2. ReLU Function: The rectified linear unit is piecewise linear making it easy to compute however it has non-linearity at Z=0 so at Z<0 the function returns 0 and at Z>0 the result is

just the same as the input. Because of this non-linearity it is still able to capture all of the great properties of activation functions whilst still being extremely simple to compute.

$$g(z) = \max(0, z)$$

$$g\acute{(z)} = = \begin{cases} 1, & z > 0 \\ 0, & otherwise \end{cases}$$

Formula 20

3. Hyperbolic tangent Function: tanh function is another sigmoidal (s-shaped) activation function with a range of (-1,1). The advantage over the sigmoid function is that the negative inputs will be mapped strongly negative and the zero inputs will be mapped near zero in the tanh graph.

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

$$g\acute{(z)} == 1 - g(z)^2$$

Formula 21

It is important to note that when updating the curve, differentiation is used to know in which direction and how much to change or update, making it a vital part of almost every machine and deep learning system.

In general, the architecture of an Artificial Neural Network consists of multiple neurons that are connected with weights W_{ij}. Depending on how these neurons are interconnected, one can obtain many different network structures. In a fully connected network, $B_{ij} \neq 0$, for all i, j. If there exists a $W_{ij} = 0$, the network is considered partially connected.

A single layer Neural Network is a shallow network composed of the perceptron model discussed above with an added hidden layer added between the inputs and outputs. The extra layer is defined as hidden because its states are not directly enforced and observable. The artificial intelligence designer typically only enforces the inputs and outputs. Since we now have a transformation from our input space layer to our hidden space layer and a transformation from our hidden space layer to our output space layer, we need two weight matrices corresponding to the weight matrices of each layer (θ_1 and θ_2). Each single unit in the hidden layer is again composed of one perceptron.

A deep neural network is composed of a stack of layers with fully connected neurons. There are several different forms of Artificial Neural Networks, with the most popular models categorized as Feed-Forward Neural Networks (FNN) and Recurrent Neural Networks (RNN). The main difference between the two models is the way each handles information flow. In FNNs, there is no flow of information between nodes in the same layer and the signal travels directly from input to output. In RNNs however signals are travelling in both directions and between nodes in the same layer.

The process of calibrating or fitting an ANN to data is often referred to as training (or learning). Learning algorithms are used to assign weights and other parameters and a complete run of a learning algorithm is defined as an epoch. Typically learning methods are split into three categories:

1. Supervised learning: Typically used in regression studies, this method is a closed loop feedback system where the network parameters are adjusted by minimizing the error function which generally is some variation of the difference between the network output and the desired output.

2. Unsupervised learning: In this method no target (desired) values are used. Instead, the network attempts to draw information from the input data using correlation-detection with the

objective of finding patterns or features without external input. Typically, this method is used in clustering.

3. Reinforcement learning: Typically used in artificial intelligence this method specifies how an artificial agent is meant to operate and learn from the given input data whilst adhering to a set of rules aimed to maximize the reward.

The loss of the network measures the cost incurred from incorrect predictions. In other words, it is a measure of the distance of the predicted output from the known outputs. The empirical loss measures the total loss over an entire dataset. The empirical loss is also known as the objective function. A neural network is trained by quantifying its empirical loss and then optimizing the network by finding the network weights that achieve the lowest loss. Since there exists a desired output for every input, the loss can be computed with the error signal then being back propagated into the network and the weights adjusted by a gradient –descent-based algorithm resulting in a closed loop system.

Loss functions are generally categorized into two groups: Regressive and Classification loss functions. Regressive loss functions are typically used in the case of regressive problems where the target variable is continuous. The most widely used regressive loss function is the Mean Squared Error which is the sum of squared distances between the target variable and predicted values:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - y_i^p)^2}{n}$$

Formula 22

Another regressive loss function is the Mean Absolute Error which is the sum of absolute differences between our target and predicted values. In other words, the MAE measures the average magnitude of errors in a set of predictions without considering their directions.

$$MAE = \frac{\sum_{i=1}^{n} \left| y_i - y_i^p \right|}{n}$$

Formula 23

The MSE is considered easier and more efficient to solve but the MAE is more robust to outlier data as the MSE loss model will give more weight to outliers. One big problem with MAE loss however is that its gradient is the same throughout, which means that even for small loss values it will portray a large gradient. This however can be resolved by utilizing a dynamic learning rate.

Classification loss functions are used in classification problems and their output is usually a probability value $f(x)$, called the score for the input x. Generally, the magnitude of the score represents the confidence of our prediction. The target variable y, is a binary variable, 1 for true and -1 for false. In contrast to regressive functions most classification losses mainly aim to maximize the margin (the margin being a measure of how correct the network is). Some popular methods of classification loss functions include Binary Cross Entropy and Margin Classifier approaches.

The loss optimisation of the network as discussed above can be formalized as:

$$\breve{\theta} = argmin \frac{1}{n} \sum_{i=1}^{n} \tau(f(x^i; \theta), \, y^i)$$

$$\breve{\theta} = argmin_{\theta} J(\theta)$$

Formula 24

Where $J(\theta)$ is the empirical loss which takes the set $\theta = \{\theta_1, \dots, \theta_n\}$ as input. In order to train the network, we start with a random guess and pick a point $\theta_0 = \{\theta_1, \dots, \theta_n)$ as a starting point and compute the gradient of this point on the loss landscape portraying how the loss is changing in respect to each of the weights. This gradient tells us the direction of the highest ascent and not descent, so we take a small step in the opposite direction by negating the gradient and readjust our weights such that we step in the opposite direction of the gradient, moving continuously towards the lowest point in the loss landscape until we finally converge at a local minimum.

Modern deep neural network architectures are extremely non convex meaning that there is a possibility that when back-propagating through the network, the network converges on one of the many local minimums instead of the true global minimum. This problem can be overcome by utilizing a learning rate variable to define the size of the step taken in the direction of our loss gradient. If the learning rate is set too low then the model may get stuck in a local minima (due to the fact that at a local minima the gradient is 0). On the other hand, if the learning rate is set too large, then the model could actually diverge and overshoot. In order to find the right balance, we can design an adaptive learning rate that adapts to the loss landscape. In practical terms this means that learning rates are no longer a fixed variable and will adjust itself to different situations such as how large the gradient is, the size of particular weights and so on. Due to the importance of this factor many adaptive learning rate algorithms have been developed some of which include Momentum (Qian, 1999), Adagrad (Duchi, Hazan, & Singer, 2011), Adadelta (Zeiler, 2012) and Adam (Kingma & Ba, 2014).

The final issue that we address in this part of the study is overfitting and regularization. Ideally in machine learning we design a model that accurately describes the test data and not the training data. In other words, the aim is to build models that can learn representations from the training data whilst still generalizing well on unseen test data. Under fitting describes the situation where the complexity of the resultant model is not high enough to capture the nuances of the test data. Overfitting describes the situation where the resultant model is too complex, having memorized the training data and not generalizing well to the test data. The ideal fit is somewhere in the middle of the two aforementioned situations where the model is not too complex but still contains the capacity to learn some of the nuances of the test set. Regularization is a technique that constrain our optimisation problem to discourage complex models improving generalization of the model on the test set.

One approach to regularization is the dropout technique, in which during training in each iteration we randomly drop some proportion of the hidden neurons with probability. When neurons are dropped from the hidden layer this forces the network to not rely on a single node but to devise alternative paths through the network. In other words, the weight distribution between nodes is more normalized and essentially discourages memorization.

The second technique of regularization is early stopping, where the network training is stopped before the probability of an overfitting situation. The early stopping algorithm identifies the point where the loss of the testing data set and the loss of the training data set start diverging. This means that although the loss from the training data set is decreasing this is in fact causing overfitting through memorization. Stopping at the ideal stop point ensures that the model performs well towards unseen data.

2-8-3-2-1 Recurrent Neural Networks

Recurrent Neural Networks (RNN) are a category of artificial neural networks devised as an answer to sequence modelling problems. In order to model sequences, there is a need to:

1. Handle variable length sequences

2. Track long term dependencies

3. Maintain information about order

4. Share parameters across sequences

RNNs differ fundamentally in their network architecture when compared with standard neural networks. As previously stated, in standard neural networks data propagates in one direction from input to output making this form of network architecture unsuitable for the modelling of sequential data. RNNs in contrast are networks which allow for information to persist through the utilization of loops where information is being passed internally from one time step to the next. Recurrent networks apply a recurrent relation at every time step to process a sequence:

$$h_t = f_w(h_{t-1}, x_t)$$

Formula 25

Where h_t is the internal cell state and f_w is a function parametrized by a set of weights W to update this state based on the previous state h_{t-1} and the current input x_t at time step t. It is important to note that the same function and same set of parameters are used at every time step satisfying the previously mentioned need to share parameters in the context of sequence modelling. More specifically the RNN computation includes both a state update as well as the calculation of the output. The state update is a standard neural network operation consisting of a multiplication by a weight matrix and then applying a non-linearity with the added complexity that there are now two weight matrices (input plus the previous cell state). nonlinearity is added to the sum of these two terms. Finally, an output is generated at a given time-step which can be formalized as:

$$\hat{y} = W_{hy} h_t$$

$$h_t = \tanh\left(W_{hh} h_{t-1} + W_{xh} x_t\right)$$

Formula 26

RNNs can essentially be thought of as copies of the same network where each copy passes information to its descendant creating a chain like structure with each link in the chain resulting in a computable loss. To define the total loss, the losses from all individual times steps is added up and since the resultant total loss consists of the individual loss contributions over time, training the network will also require some time component. The training algorithm used to address this component is a variant of the back prorogation approach studied in the previous part of this study called BPTT(Backpropagation Through Time). In BPTT errors are propagated at each individual time step and then across time steps all the way to the beginning of the time series. In order to be able to consider the gradient flow across the chain of repeating modules, we have the factor W_{hh}. Computing the gradient with respect to h_0 involves many factors of W_{hh} and a derivative of the activation function could result in an exploding gradient problem where the gradient becomes extremely large and optimisation cannot occur. Gradient clipping is used to address this problem which entails scaling back the gradient when it becomes too large. If the W_{hh} matrix values are too small then the network could face the vanishing gradient problem which is a result of multiplying many small numbers together. The vanishing gradient problem is of importance in RNN architecture as it creates a bias in the network to capture short-term dependencies whilst losses caused by nodes further back in time develop smaller and smaller gradients. There are three main approaches to the vanishing gradient problem: Changing the activation function (The ReLU activation function operates better than sigmoid functions), initializing parameters in the network by initializing weights based on an identity matrix and the third most robust solution which uses a more developed complex recurrent unit with gates to control what information is passed through time. This last approach is the driving principle behind recent gated cell technologies such as LSTM and GRU.

LSTM (Long Short-Term Memory) networks rely on a gated cell to track information throughout many time steps which are good at learning long-term dependencies and overcoming the vanishing gradient problem. As the RNN is being unrolled across time, each repeating module contains a simple computation node. In LSTMs the repeating unit contains different interacting layers which maintain a cell state c_t and structures formally called gates are used to add or remove information from cell state c_t. Gates are devised in such a way to optionally let information through, via a sigmoid neural network layer and pointwise multiplication. The sigmoid function is important as it enforces the input to the gate to constrain between 0 and 1, if the value is 0 then none of the information passes through the gate and if it equals 1 all of the information is passed through, regulating the flow of information through the gate.

The first step of the operation of the LSTM decides what information is to be discarded from the prior cell state, or in other words forgetting irrelevant history. The next step takes previous information and also the current input, selectively updating the current cell state. The final step returns an output via an output gate returning a transformed version of the cell state. In summary the key three LSTM operational steps can be summarized as 1. Forget 2. Update 3. Output.

In the forget step, The LSTM utilizes a sigmoid layer called the forget gate f_t parametrized by a set of weights and biases (like all neural network layers). The sigmoid layer considers previous information h_{t-1} as well input x_t and outputs a number between 0 and 1 corresponding to whether information is completely forgotten or retained. The forget gate can be formalized as:

$$f_t = \left(W_f . \sigma[h_{t-1}, x_t] + b_f\right)$$

Formula 26

In order to decide upon what information should be stored in the updated cell state another sigmoid layer is utilized as:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

Formula 27

And then a tanh layer is used to generate a new vector of candidate values that could be added to the state.

$$\tilde{C}_t = tanh\left(W_c\left[h_{t-1,}x_t\right] + b_c\right)$$

Formula 28

For the final phase of updating the old cell state c_{t-1} to the new cell state c_t the old cell state is first multiplied by the forget gate to decide what should be forgotten or kept. Then the set of new candidate values scaled by how much is decided to update.

$$C_t = f_t * \ C_{t-1} + \ i_t * \ \tilde{C}_t$$

Formula 29

For the final output step, another sigmoid layer is utilised to decide what parts of the state to output

$$O_t = \sigma(W_O[h_{t-1}, x_t] + b_O)$$

Formula 30

With the result multiplied by tanh of the cell state transforming the output to a filtered version of the cell state:

$$h_t = O_t * \tanh\ (C_t)$$

Formula 31

Finally, it should be noted that because in LSTMs Backpropagation from C_t to C_{t-1} requires only elementwise multiplication (and addition), no matrix operation is needed, which ultimately avoids the vanishing gradient problem. When these repeating LSTMs are linked in a chain the resultant gradient flow will be completely uninterrupted increasing training efficiency whilst being immune to the vanishing point problem.

2-9 Previous Research

This part of the literature review will focus on the previous research conducted on portfolio optimisation methodologies. It should be noted that due to the inherent interest of this study on hybrid passive and active investment strategies, the scope of the literature review emphasises a variety of different optimisation techniques which include (but are not limited to) linear and quadratic programming and machine learning.

Rudd introduced a single parameter model with a heuristic framework with the aim of tracking the S&P500 index. In his study Rudd argues that portfolio selection by stratification gives no indication of the relative benefit of adding or deleting a stock (something which he believed to be important when revising and rebalancing the portfolio as the cost of the transaction must be weighed against the benefit of tracking the index more closely) and so came up with a model which sought to minimize variance by constraining the index tracking portfolio beta to one. The model also takes into account the transaction costs through a weighted parameter in the objective function (Rudd, 1980)

Brinson, Hood and Beebower (1986) showed in their study that investment policy or in other words the allocation of assets, was of a higher importance and more influential than other facets of the investment strategy such as the selection of individual stocks or market scheduling. They found that investment policy accounted for 93% of the fluctuations of a portfolios' returns in a given time frame and concluded that investors must prioritize investment policy over investment strategy. In other words, investors must spend more time on the allocation of funds and less time on market speculation and the selection of each individual stock

Meade & Salkin studied the selection of an index fund which minimizes expected tracking error by devising a multivariate model of returns on shares (which itself was a development of a univariate model by Taylor) to formulate the selection problem as a quadratic program. In the resultant portfolio the weight of each represented market segment closely mimics the index (Meade & Salkin, 1989).

A hybrid passive/Markowitz approach was introduced by Roll where he introduces the TEV Criterion: minimization of tracking error variance for a given expected tracking error, which can be used to select an efficient portfolio which intentionally does not produce a mean/variance Markowitz efficient portfolio. Furthermore, he proposed that that it is possible to minimize tracking error variance conditional on a specified beta (Roll, 1992).

Tabata & Takeda propose that the main two problems faced when devising an index tracking portfolio are a) the minimization of securities in the index fund and b) the minimization of tracking errors between the benchmark index and the index-tracking fund. They then introduce an innovative two stage optimisation model with the aim of minimizing tracking error. In the first stage the portfolio components are selected and in the second stage an optimal solution is presented which looks to simplify the complex quadratic programming involved in minimizing the tracking error under the given number of securities included in the index fund (Tabata & Takeda, 1995).

Rohweder also proposed a Markowitz based model where portfolio segmentation is offered as an alternative to tracking error optimisation where the total portfolio is divided into active and passive accounts. The passive sub account is handed over to a manager tracking the benchmark (index tracking) and the active sub account remains with the active manager who invests according to style and preferences. (Rohweder, 1998)

Larsen & Resnick further developed the framework specified by Rudd to construct an index tracking portfolio while studying the effects of portfolio composition on the tracking performance of the portfolio. The study included the effects of rebalancing and restructuring frequency on tracking performance and found that the rebalancing frequency of the optimal portfolio weights for the stocks in the indexed portfolios does not appear to be critical factor affecting tracking performance (Larsen & Resnick, 1998).

Rudolf, Wolter & Zimmerman use linear programming to investigate four different models for minimizing the tracking error between a portfolio and the benchmark index and compare the results against a quadratic optimization technique where squared deviations are used. By arguing that fund managers are compensated based on the linear performance of the fund they utilize minimized absolute deviations and find that the optimisation model should be targeted to the specific investment objective and that linear tracking error optimization is equivalent to expected utility maximization and lower partial moment minimization (Rudolf, Wolter, & Zimmermann, 1999).

Jansen & Van Djik (2002) focused on the approach of minimizing tracking error with a relatively small number of stocks. The simple heuristic algorithm used is composed of two steps, a) selection of the best weights of the stocks in the target index to minimize tracking error and b) removal of a set number of stocks that are weighted smallest in the previous solution and then solve the problem of finding the best portfolio out of the remaining stocks to minimize tracking error using quadratic programming. These two steps are continued until only a set number of stocks remain in the portfolio. Coleman & Li (2006) continue upon this by proposing an adaptation where a convex programming problem without the cardinality constraints is solved and its global minimizer is calculated.

Beasely, Meade and Chang use a population heuristic (genetic algorithm) to search for an efficient tracking portfolio by explicitly imposing cardinality constraints. Their formulation of the problem explicitly includes transaction costs (associated with buying or selling stocks) and a limit on the total transaction cost that can be incurred whilst also including a constraint limiting the number of stocks that can be purchased. (Beasley, Meade, & Chang, 2003).

Focardi & Fabozzi discuss two issues in their paper, a) defining suitable performance objectives and tracking error that scale properly over the whole management period and b) implementation of an optimized investment strategy when it is not feasible to replicate the index in its entirety. They then continue by introducing a clustering approach to optimisation where it is possible to identify optimal portfolios, restricting the search to representatives from each cluster, thereby reducing the computational complexity of the optimization task. As clustering is instinctive, it allows for the interaction of judgement and quantitative optimization (Focardi & Fabozzi, 2004).

Oh, Kim and Min propose a scheme exploiting genetic algorithms providing the optimal selection of stocks using fundamental variables: standard error of portfolio beta, average trading amount and average market capitalization. Their approach consists of two steps. First the stocks for the index fund are selected based on the fundamental variables in each industry sector and secondly the relative weights of each individual selected stocks are optimized using a genetic algorithm process. Again, this study aims to show that the efficient portfolio selection process easily replicates the benchmark index while consisting of a small number of stocks (Oh, Kim, & Min, 2005).

Corielli & Marcellino argued that many common index tracking methodologies do not take into account the dynamic properties of the index component which leads to low frequency or integrated (persistent) components in the tracking error. They continued their argument by proposing a new approach by introducing a dynamic factor model where the price of each stock in the index is driven by a set of common and idiosyncratic factors which can be integrated or stationary. This model is used to construct a replica portfolio of the index which is driven by the same persistent factors as that of the index, whose behavior is as close as possible to that of the index itself (Corielli & Marcellino, 2006).

In their study Fernandez and Gomez further develop the standard Markowitz mean/variance model by including cardinality and bounding constraints. By limiting the total number of assets to be included in the portfolio and also the share of capital to be invested in each asset they transform the portfolio selection problem from a quadratic to a mixed quadratic and integer programming problem. By utilizing artificial neural networks to solve this problem they found that a Hopfield based neural networks provides better solutions in relation to genetic algorithm (Fernandez & Gomez, 2007).

Chang, Yang, & Chang (2009) offer a new approach to portfolio optimisation using genetic algorithms by utilizing three different risk measures based on the Markowitz mean/variance model. Mean-variance, semi-variance, mean absolute deviation and variance with skewedness are used to measure risk in a cardinality constrained efficient frontier. The study results show that genetic algorithms can be used as a robust solution to portfolio optimisation problems (including cardinality constraints) with different risk models. The proposed solution also re-affirms the fact that a portfolio of smaller size could have a better performance than those of a bigger one (Chang, Yang & Chang, 2009).

Hejazi et al propose a framework for the implementation of genetic algorithms in optimizing classical approaches to asset allocation. Their results show that by utilizing genetic algorithms they were able to reduce deviation from the index returns and their resultant portfolio was able to achieve higher returns than that of the index in the proposed research timescale. (Hejazi et al, 2012).

In their paper Chen & Kwon adopt an integer program that selects a given number of securities for a tracking period over a single period. By applying a robust discrete optimization which allows for the specification of ranges of uncertainty in the object function, this framework protects against worst case realizations of the object function whilst maintaining computational tractability. Their results showed that incorporating robustness in the model can provide improved performance (Chen & Kwon, 2012).

Razeghi & Hanifi utilized a Hopfield neural network in order to solve the asset selection problem for an index tracking fund. They implemented their proposed strategy on the top five global markets and finally performed a comparison of results from the Genetic algorithm strategy finding that the neural network methodology was more precise. (Razeghi,2012).

Gharekhani conducted research on how to apply robustness to the Markowitz mean-variance model whilst taking into account integer constraints and unknown returns. A multi-stage decision machine was used to solve the asset allocation problem whilst adhering to robustness constraints in the form of robust asset valuation (Gharekhani, 2012).

To mark the 60th anniversary of the Markowitz's paper "Portfolio Selection", Kolm, Tutuncu & Fabozzi review a variety of different approaches developed to address the challenges faced when optimizing portfolios. While taking into account various parameters such as transaction costs, management constraints and sensitivity of expected returns forecasts, the study covers concepts which have been instrumental in the development of financial markets and the financial decision-making process, while highlighting some of the new trends and developments (Kolm, Tütüncü, & Fabozzi, 2014).

Najafi & Mooshkhian (2015) studied the multi-period portfolio selection problem with the following assumptions: 1. Investment outlooks are short-termed 2. Transactional costs of trading are of importance 3. Problem parameters are known values. They introduce a model based on the mean variance theory to overcome Markowitz's existing models. Their results highlight the efficiency frontier whilst pointing out the fact that there is a need to address uncertainty.

Guastaroba, Mansini, Ogryczak, & Speranza (2016) attempt to address the problem of selecting a portfolio of securities able to outperform a market index while bearing limited additional risk. This problem is also known as the Enhanced Index Tracking Problem and in order to formulate a suitable solution the omega ratio is used to account for asymmetry in return distributions. Two models are portrayed in the study, the first computes the omega ratio in regards to a benchmark represented by the mean rate of return of the market index (OR model). In the second model, multiple scenarios are used by substituting the mean rate of return in the previous model with a random variable (EOR model). Results showed that the portfolios selected by the EOR model clearly outperform those found by the OR model in terms of out-of-sample performance. Furthermore, the portfolios selected by the EOR model track the behavior of the benchmark over the period very closely while yielding significantly larger returns in some variations. Results suggest that considering the market index as a random variable is a valuable choice.

AkbariFard et al (2017) conducted an experiment in which the results of three different portfolio optimisation methodologies were compared. They utilised Genetic Algorithms, Symbiotic Organism Search (SOS) and Particle Swarm Optimisation (PSO) techniques to formulate portfolios with the highest returns and risks with an initial population of 50 securities from the TSE utilising the Markowitz Model. Their results showed that all three approaches were suitable with SOS being the most efficient and bests performing of the three methodologies studied.

Asoroosh et al (2017) performed a study in which TLBO (Teaching Learning Based Optimisation) was used to optimize the Markowitz model. In this model variance, absolute deviate from average, semi-variance and CVAR were used as a measure for risk. When comparing the performance of the resultant portfolio to that of Imperialism Competition Algorithm (ICA) the researchers showed that the TLBO approach had better performance (study of Sharpe's ratios) whilst being more efficient (in terms of calculation).

Tehrani et al (2018) studied the performance of the Krill Herd Metaheuristic Algorithm when used to optimise the Markowitz model. Their results showed that in comparison to ICA and PSO optimisation methodologies the Krill herd algorithm performed better (Sharpe's ratio) and benefitted from a higher stability when obtaining the efficient frontier. The measures of risk used in the study included variance, semi-variance and expected shortfall.

FallahPour et al (2018) identified the shortcomings of Markowitz's mean-variance model and argue that adding constraints and portfolio members to the existing model makes portfolio optimisation an NP-hard problem which makes it impossible to solve using derivative-based methodologies. They continue to implement a Whale optimisation algorithm (introduced in 2016 based on the behaviour whales) to offer a solution to this problem and conclude by showing that the Whale Optimisation methodology performed better than their ICA and PSO counterparts.

2-10 Conclusion

A review of the current literature on asset allocation and portfolio optimisation clearly highlights the competition between active and passive strategies. Each strategy has their own set of benefits and disadvantages that managers must understand in order to identify a best form of practice for their individual goals. The advent of hybrid paradigms such as a core-satellite approach however acts as a bridge between both strategies allowing investors and managers to benefit from both while addressing some of the shortcomings of each individual strategy.

Once a suitable asset allocation strategy has been selected, the end user is then faced with a variety of different optimisation techniques which can be used to add further complexity and constraints to existing asset allocation problems. These methodologies are able to model uncertainty whilst leveraging the computational power of heuristic algorithms in order to find optimal or near-to-optimal solutions that would otherwise be impossible to calculate due to the inherent vast size of data sets and stochastic nature of data.

Although there is an extensive library of works dedicated to the optimisation of the classical Markowitz Mean-Variance model, there seems to be less focus on the optimisation of the index tracking problem. When taking into account the stochastic nature of the data and the resultant uncertainty, the implementation of robustness into a suitable model becomes important when attempting to minimise risk and maximise returns. The Use of a Genetic Algorithm in the reduction of tracking error has been studied before however this study was unable to find any models which handled robustness explicitly whilst acting in a compound manner to address the needs of a core-satellite approach.

One aspect of reducing uncertainty in the initial data would be attempting to model and make approximations of future index returns using an artificial neural network in an RNN configuration. Although this is mentioned in previous literature this study was unable to find conclusive results pertaining to the performance of such a model.

The next chapter of this study will outline the research methodology of this research whilst presenting the innovative heuristic model used in the framework.

Chapter 3: Research Methodology

3-1 Introduction

The choice of research methodology depends on the objectives of the research, nature of the subject matter and the resources required. In other words, the implementation of a suitable research methodology will allow for an efficient, cheaper, easier and more exact research process.

This chapter of the thesis will focus on the analysis of the utilised research methodology in this research whilst outlining the steps undertaken. After an overview of the statistical population and data gathering methodology used is given, the calculation and approximation of research variables is offered. The last part of the chapter will attempt to explain the heuristic algorithms utilised in the study.

3-2 Research Problem

As outlined in the first chapter of this study the proposed research problem can be split up into three distinct phases each with their own unique methodology and application.

The initial phase of the research problem can be defined as the formulation of the optimised index tracking problem using an exact algorithm. The resultant equation and constraints can then be further optimised and enforced by introducing robustness through the utilisation of an adjustable conservatism ratio. The third phase includes the development and utilisation of a Genetic Algorithm to solve the resultant equation and compile a portfolio of stocks with similar returns to that of the market index. The fourth phase concludes the research problem, formulising the core-satellite hybrid portfolio selection problem and using the Genetic Algorithm to solve the resultant equation to assign investment weights for both the core (index tracking) and satellite (secondary actively managed funds) components.

It should be noted however that this study will aim to improve the robustness of the index tracking model and resultant portfolio by utilising a Recurrent Neural Network in LSTM (Long-Short Term Memory) configuration to attempt to make predictions on future index returns based on long term historical data.

3-3 Research Method

Research in general is conducted with one of two objectives: 1. Resolving a specific set of problems which exist in an institution or specific industry 2. Contribution of knowledge to a field of expertise which is of specific interest to the researcher.

In its most abstract form, basic research is conducted with the aim of creation and refinement of theories. In this format of research, problems are not born from real-world applications and their objective is in fact to test theories, clarify relationships between phenomena and contribute to a specific field of expertise. In contrast, applied research is aimed at finding a solution for an immediate problem facing a society, or an industrial/business organisation (Kothari, 2008). Applied research could be further defined as research in which theories, relationships and principles which have been developed in basic research are utilised to solve real-world practical problems. As a result, this format of research is more focused on effective solutions rather than the cause.

By reviewing the objective of this study in its most basic form, it is possible to define the aim of proposed framework as the minimization of tracking error, which itself is a function of the resultant portfolios expected rate of return in respect to that of the index. Thus, it is argued that the present research is quantitative in nature and can be classified as inductive within the applied realm of research.

3-4 Research Steps & Processes

This research is comprised of five different steps as outlined below:

1. Collection of data regarding daily market index values and stock prices listed on the Tehran stock exchange within the proposed research timeframe and the preparation of this data to be used in data processing and analysis.
2. Formation of an index tracking portfolio of stocks using exact algorithm to be used as a point of comparison to portfolios utilising robust optimisation.
3. Formation of an index tracking portfolio using a genetic algorithm and then comparing the performance of the resultant portfolio against portfolios utilizing robust portfolio optimisation.
4. Utilization of an LSTM neural network to attempt to predict index returns into the future using historical data; providing an outline which can be used to not only add robustness to the final design of framework but also enforce the genetic algorithm utilized in the next step.
5. Implementation of a core-satellite approach investment approach and evaluation of the resultant portfolio performance against index returns.

3-5 Population & Data Gathering

The initial population of this study includes all companies listed on the Tehran Stock Exchange.

In order to minimize tracking error a collection of stocks has to be selected which have the most impact on the index, resulting in similar performances and fluctuations than that of the index itself. Filtering and cross-sectional techniques are used to identify a suitable sample from the statistical population. All selected stocks are filtered to have 400 trading days in the four-year time frame (100 trading days per year). Active trading days were defined as days that did not have a value of zero for returns, as stocks with daily returns of zero (as a result of minimum trading quantities not being met or no changes in daily stock price) could not impact the index. The top 100 stocks which held the highest market value in the time period were excluded from this filtering process. In order to ensure that the ensuing model is successful in formulating an index tracking portfolio (with performance similar to that of the index), the listed daily stock prices have to be prepared and initialised so that the daily returns reflect and take into account the earning per share payments and capital increases applied within the studied time series. On the other hand, many heuristic algorithms developed to formulate an index tracking portfolio require that daily returns are present and available throughout the time series being studied. In other words, gaps in daily returns which could be a result of variety of different factors such as absence of trading transactions within a given date, market exits or non-trading notices from regulatory bodies need to be mitigated. To this extent, missing values are simulated using econometric methodologies in order to attain a standard set of data, which allows all stocks which are effective in the index to be given a chance at selection in the final portfolio.

3-6 Research Problem Formulisation

The formulation of an investment portfolio with the aim of attaining similar returns than that of the market index is known as the index tracking problem. In order to reduce transactional costs only a subset of stocks which make up an index are selected. (Jeurissen, 2005). Thus, the research problem of this particular study can be defined as the selection of a subset of K stocks listed on the Tehran stock exchange with optimum weights for each individual stock in the fund at time T which has a similar performance in terms of returns within a time period (T, T+ε). The first objective function of this research is the minimization of tracking error for this subset of stocks in relation to the index. It should be noted that the final objective of the investor is to assign optimum weights to each stock in a way as to which minimize tracking error.

There exists a myriad of solutions in calculating the tracking error each with their own advantages and flaws. One of the most common techniques is calculation of the variance difference between the portfolio returns and returns of the base index which Beasely et al (2003) argue could pose some problems. Consider that the portfolio always underperforms the index ($r_t = R_t$-M) where M is a fixed positive value. In this situation the resultant standard deviation ($r_t - R_t$) will always equal to zero which would result in a tracking error of zero. (Beasely, Meade & Chang, 2003).

In order to solve this issue Beasley et al (2003) propose that tracking error should be considered as a function of the difference between the index portfolio returns and the base index. In the next section we will highlight a few of these linear models such as the Mean Absolute Deviation (MAD), Mean Absolute Downside Deviation (MADD) and MiniMax whilst introducing a quadratic model to formalize tracking error.

Consider an investor which chooses n stocks from a base index and their optimum corresponding weights in order to compile a fund. Variable β corresponds to the weight of the individual stocks, Y the index returns vector and X a returns matrix for n stocks within each timeframe where returns are compounded. T as the number of observations and ε the deviation between returns of index tracking portfolio and the base index.

3-6-1 Mean Squared Error Optimisation

Similar to Rudolf, Wolter, & Zimmerman (1999), the objective function for minimizing tracking error is based on asset weights β. The deviation between returns of the index tracking portfolio and the base index is commonly defined as:

$$\varepsilon = Y - X\beta, Y\epsilon\mathbb{R}^T, X\epsilon\mathbb{R}^{T\times n}, \beta\epsilon\mathbb{R}^n, \varepsilon\epsilon\mathbb{R}^T$$

Formula 32

Thus, the tracking error based on the mean squared approach proposed by Roll (1992) can be formalized as:

$$\min_{\beta} \acute{\varepsilon}\,\varepsilon \equiv \min_{\beta}(Y - X\beta)'(Y - X\beta)$$

Formula 33-1

It can be shown that the vector of optimum weights of assets in the means squared error minimization problem to be:

$$\beta = (\acute{X}X)^{-1}\acute{X}Y$$

Formula 33-2

Furthermore, a set of constraints can be added to the above model to disallow leveraged sales which translates to the asset weights vector being positive. Also, the sum of all weights should be equal to one. Thus, the mean squared error minimization problem can be rewritten as a quadratic program as follows:

$$\min_{\beta} \acute{\varepsilon}\,\varepsilon \equiv \min_{\beta}(Y - X\beta)'(Y - X\beta)$$

$$s.t$$

$$1.\beta' = 1,$$
$$\beta \geq 0.$$

Formula 33-3

3-6-2 Linear Models for Tracking Error Minimization Problem

The following models minimize the absolute deviation between the returns of the index and the tracking portfolio. MAD minimizes the mean absolute deviation. The second model MiniMax minimizes the maximum deviation between the returns of the tracking portfolio and the base index. The other approach focuses on the negative deviations between the index and the tracking portfolio which according to Harlow (1991) is significant because from the viewpoint of the investor risk is a result of situations where the returns of the portfolio become lower than those of the index. This is also known as downside risk in investment.

The functions for each different model can be summarized as:

$$TE_{QD} = \min_{\beta}(Y - X\beta)\acute{}(Y - X\beta)$$

$$TE_{MAD} = \min_{\beta} \acute{1}\,(|X\beta - Y|),$$

$$TE_{MADD=} \min_{\beta} \acute{1}\,(|\bar{X}\beta - \bar{Y}|) \quad \text{For some t where } \bar{X}_t\beta \leq \bar{Y}_{\acute{t}}$$

$$TE_{MiniMax} = \min_{\beta}\left\{\max_{t}|X\beta - Y|\right\},$$

$$TE_{DMiniMax} = \min_{\beta}\left\{\max_{t}|\bar{X}\beta - \bar{Y}|\right\}, \text{ For some t where } \bar{X}_t\beta \leq \bar{Y}_{\acute{t}}$$

Formula 34

Where X_t is the tth row of matrix X and Y_t is the t[th] element of vector Y considering that Matrix $\bar{X}$ and vector $\bar{Y}$ only consider rows where the resultant returns of the index are lower than the returns of the tracking portfolio. In order to portray the proposed model in a classical linear format a supplementary variable $z \geq 0$ is introduced as an upper limit of the absolute deviations as prescribed by Rudolph (1998). Thus, for the MinMax model we can consider:

$$z \geq |X_t\beta - Y_t|, t\epsilon\{1, \dots, T\}$$

Formula 35-1

Where T can have two states:

$$if\ X_t\beta \geq Y_t \Leftrightarrow X_t\beta - z \leq Y_t$$

$$if\ X_t\beta \leq Y_t \Leftrightarrow X_t\beta + z \geq Y_t$$

Formula 35-2

Which leads to the MinMax model being formalized as:

$$\min_{z} z, s.t. X_t \beta - z \leq Y_t, X_t\beta + z \geq Y_t$$

Formula 36-1

For downside risk only the second state applies.

Furthermore, for the DMiniMax model it can be deduced:

$$\min_{z} z, s.t. X_t \beta + z \geq Y_t$$

Formula 36-2

In the MAD model variable $z_t^+ \geq 0$ is introduced which signifies that the deviation is positive and $z_t^- \geq 0$ which is the absolute vale a deviation between portfolio returns and those of the index.

$$min \sum_{t=1}^{t} (z_t^+ + z_t^-)$$

$$s.t. X_t\beta - z_t^+ + z_t^- = Y_t$$

Formula 37

For the MADD model z_t^+ is eliminated and the problem is formalized as:

$$min \sum_{t=1}^{T} z_t^-$$

$$s.t. X_t\beta - z_t^+ + z_t^-$$

Formula 38

The four introduced models and the quadratic model use historical data in order to form an index tracking portfolio and suppose that the composition of the portfolio and their individual weights (which is based on the solution of the subsequent minimization problem) to hold true for a foreseeable future. On the other hand, robust optimisation is utilized in situations which arise from uncertainty, so in the suggested model we are forced to use expected rates of return rather than absolute historical returns as the basis for solving the tracking error minimization problem. This study attempts to overcome this problem by utilising a Recurrent Neural Network (in LSTM configuration) to anticipate and predict expected rates of returns thus adding robustness to the overall solution.

In order to develop a suitable robust implementation, the following objective function is defined as the minimization of deviation between the expected returns of the optimized portfolio and that of the stock index formalized as:

$$Min \left| \sum_{i=1}^{n} r_i \, w_i - R \right|$$

Formula 39

3-6-3 Calculation of Returns

Given that the objective of this study is to closely track the index of the Tehran stock exchange (which itself is a measure of the returns of individual listed stocks), the data used to calculate the returns of each stock must also reflect their overall returns in a daily timeframe.

The overall returns of a stock can be calculated by taking into account different factors such as price, earnings per share paid out in cash and stock options and capital increases (based on source of capital increase). In order to gather the necessary data TSEClient 2 software was used to extract all necessary data for each individual stock whilst actual returns were calculated using the following formula:

$$R = ln\frac{I_t}{I_{t-1}}$$

Formula 40

Where I_t is considered to be the value of assets at the end of the time frame and I_{t-1} to be the value at the beginning of the time frame.

3-6-4 Exact Algorithm Implementation

In order to solve the research question, we first attempt to solve the question on a smaller scale using exact algorithms. In order to show the effects of uncertainty in the model parameters (Expected rate of returns of index and tracking portfolio) we then attempt to solve the problem by utilizing the results from the LSTM model as a substitute for the expected rates of return whilst using robust optimisation to show the effectiveness of such approaches in the minimization of tracking error. The final part consists of the comparison of the performance of the optimised portfolio in regards to its ability to track the stock market index using out of sample data.

3-6-4-1 Formalisation

In order to formalize the problem, we must first introduce the notation used:

Input Parameters:

R: Expected returns of the stock market index

r_i: Expected returns of the i^{th} stock from the selection pool of stocks available for the formation of and index tracking portfolio

K: Optimised number of constituents of the index tracking portfolio

Decision Variables:

Zi: Binary variable used to show the inclusion or absence of i^{th} stock in the final index tracking portfolio. 1 is used to signal inclusion of the stock and 0 is used to signal its absence

Wi: Weight of the ith stock in the index tracking portfolio

In order to solve the index tracking portfolio selection problem with regards to integer restraints the problem is formalized as:

$$Min \left| \sum_{i=1}^{n} r_i w_i - R \right|$$

Subject to:

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

$$\sum_{i=1}^{n} z_i = k$$

$$LBz_i < w_i < UBz_i, i = 1,2, \dots, n$$

$$z_i \in \{0,1\}$$

Formula 41

The first part of the above equation is the tracking error function (objective function) to be minimized. The pursuing constraints formalizes the constraints which ensures the assigned weights of all constituent members ($w_0 \ldots w_i$) of the portfolio add up to one hundred percent whilst ensuring that the number of chosen stocks adhere to the prerequisites set by the creator of the portfolio with regards to integer restraints.

3-6-4-2 Robust Counterpart with Adjustable Rate of Conservatism

In order to offer a robust model, the uncertainty of the index returns and those of the chosen portfolio are viewed as a linear range of variations which can be formalized as:

$$Ur=\{\widetilde{ri} \in (E(ri)-\sigma ri, E(ri)+\sigma ri)\}, UR=\{\tilde{R}\epsilon(ER-\sigma RER+\sigma R)\}$$

Formula 42

In the robust counterpart model, we are looking for the optimum solution which presents the best results in the worst possible scenario with respect to the range of deviations of the uncertain parameters. Thus, to transform the research question into its robust counterpart we devise:

$$\min_{w} \quad \max_{r\iota\in \widetilde{Ur}, \tilde{R}\in UR} \quad \sum_{i=1}^{n} \tilde{r}\, w\mathrm{i}\text{-}\tilde{R}$$

Formula 43

We can then use a complimentary variable $y \geq 0$ in order to define the upper boundary of the deviation in order to transform the proposed model into a classical linear format.

$$\left|\sum_{i=1}^{n} \widetilde{ri} - \tilde{R}\right| \leq y$$

$$if\, \tilde{r}w_i < \tilde{R}: \left|\sum_{i=1}^{n} \tilde{r}w_i - \tilde{R}\right| \leq y \approx \sum_{i=1}^{n} E_{r_i} W_i + \sum_{i=1}^{n} \sigma_i wi^{-E} R^{-6} R \geq -y$$

$$if\,\tilde{r}w_i < \tilde{R}: \left|\sum_{i=1}^{n} \tilde{r}w_i - \tilde{R}\right| \leq y \approx \sum_{i=1}^{n} E_{r_i}W_i - \sum_{i=1}^{n} \sigma_i wi^{-E} R^{-6} R \geq -y$$

Formula 44

Furthermore, we can formalize J as a set of uncertain coefficients (expected returns of the selected stock and that of the index) and S as a subset of J with a value of Γ. Thus, buy using the two above equations and also taking into account that a maximum Γ of uncertain coefficients adopt values other than their face values we can devise:

$$Min\, y$$

$$\sum_{i=1}^{n} E_{r_i} w_i - E_R + \max_{\{s \subseteq J, |s| = \Gamma\}} \left\langle \sum_{i \in s} \sigma_{r_i} w_i + \sigma_R \right\rangle \leq y$$

$$\sum_{i=1}^{n} E_{r_i} w_i - E_R - \max_{\{s \subseteq J, |s| = \Gamma\}} \left\langle \sum_{i \in s} \sigma_{r_i} w_i + \sigma_R \right\rangle \geq -y$$

Formula 45

In order to solve the internal maximization problem in the above equations we substitute the linear version as:

$$\max_{\{s \subseteq j, |s| = \Gamma\}} \left\langle \sum_{i} \epsilon_5^6 r\,, wi_R^{+\sigma} \right\rangle \to \max \left\langle \sum_{i=1}^{n} \sigma_{r_i} w_i + \sigma_R \alpha_0 \right\rangle$$

Subject to:

$$\sum_{i=1}^{n} \alpha_i + \alpha_0 \leq \Gamma$$

$$\alpha_i, \alpha_0 \in \{0,1\}$$

Formula 46

The dual minimization equivalent of the maximization equation above can be rewritten as:

$$\min \sum_{l=1}^{n} v_i + v_0 + \Gamma\lambda$$

Subject to:

$$\lambda + v_i \geq \sigma_{r_i} w_i \quad \forall i$$

$$\lambda + v_0 \geq \sigma_R$$

$$v_i \geq 0 \ \ \forall i$$

$$\lambda \geq 0$$

Formula 47

Now we can replace the maximization parts with the minimization equation above. Also, we add the following constraint so that the integer restraints are satisfied:

$$v_i \leq z_i \quad \forall i$$

Formula 48

Finally, the robust equivalent of the research question with and adjustable conservatism ratio can be formalized as:

$$\min y$$

$$subject\ to:$$

$$\sum_{i=1}^{n} E_{r_i} w_i - E_R + \sum_{i=1}^{n} v_i + \Gamma > +v_0 \leq y$$

$$\sum_{i=1}^{n} E_{r_i} w_i - E_R - \sum_{i=1}^{n} v_i - \Gamma\lambda - v_0 \geq -y$$

$$\lambda + v_i \geq \sigma_{r_i} wi, \forall i$$

$$\lambda + v_0 \geq \sigma_R$$

$$v_i \geq 0 \ \ \forall i$$

$$v_i \leq z_i \ \ \forall i$$

$$\lambda \geq 0$$

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

$$\sum_{i=1}^{n} z_i = k$$

$$LBz_i < W_i < UB_{z_i, i=} 1,2,3, \ldots, n$$

$$z_i \in \{0,1\}$$

Formula 49

3-6-4-3 Problem Solving & Comparison of Results

After Solving the both the preliminary equation and its robust counterparts, the resultant portfolios are compared based on their index tracking performance. Factors such as the correlation coefficient of the portfolio and index returns in a time series, mean squared root error and an average of increased returns will be used to evaluate the relevant performance of the resultant portfolios and highlight whether the introduction of robustness (As an adjustable degree of conservatism) is beneficial to the final framework.

3-7 Implementation of LSTM RNN in order to predict future returns of Index

As previously mentioned in this part of the study we will attempt to devise a LSTM recurrent neural network which will attempt to make predictions on future index returns using previous historical data. In order to implement this model, we first utilise python coding language to read a list of daily index values for an extended time period and create a lookup table containing returns values for the index at set intervals within the extended time period.

By setting a look back variable n we are then able to group entries from the look up table into time series arrays which include the returns values for the set intervals, n steps in the past. By formatting the data in such a way, we are able to utilise the “back propagation though time” characteristics of LSTMs to identify hidden relationships between the variables.

In order to prevent bias creation, the data is then normalized using a min-max scaler meaning that the minimum and maximum value of a variable will be within the [0,1] interval accordingly. Since the range of values of raw data varies widely, in some machine learning algorithms, objective functions will not work properly without normalization. Therefore, the range of all features should be normalized so that each feature contributes approximately proportionately to the final distance.

3-7-1 LSTM Model Architecture

The LSTM model devised for this study is comprised of three layers; Two LSTM layers with 64 and 32 neurons respectively and a single neuron output layer. Each layer is designated a dropout ratio to ensure overfitting does not occur and a number of neurons are randomly selected and dropped from the layer. Furthermore, in order to ensure that the model does not get stuck in a local optimal solution an adaptive learning rate methodology is implemented.

Once the model has been initiated, the normalized training data set is input and the model is compiled.

The testing data set is then input into the resultant model and the predictions from the model can be evaluated against real world data.

3-8 Implementation of GA for solving of Research Problem

The search space of k stocks from n number of available stocks in considered to be an extensive search space which logarithmically increases with an increase of k number of stocks in the final portfolio and also the widening of n number of stocks available for selection. Thus, the utilization of an innovative heuristic technique such as genetic algorithms becomes important in completing the search in the shortest amount of time possible whilst ensuring the consistent and efficient search of the available solution space.

The algorithm presented above is utilized as the objective fitness function of the genetic algorithm where the minimization of the aforementioned algorithm will lead to a higher level of fitness between the portfolio returns and those of the index. We will provide an overview of the design and development stages of the genetic algorithm in the next section.

3-8-1 Genes and Chromosome Structure

The most important factor in solving a problem using genetic algorithms is to devise a structure for genes and chromosome in such a way that each chromosome represents a valid potential solution.

In the structure offered by this study, the designed chromosome is constructed of a k member strand where k is smaller than n (number of available equities for selection). Each gene which makes up the chromosome is representative of an asset in the underlying portfolio which has been randomly chosen from n possible assets. For example, if k=6 assets are selected from a possible n=100 assets, the corresponding chromosome could be portrayed as the shape below where assets numbered 8, 12, 20, 85, 55, 71 have been selected.

8	12	20	85	55	71

Figure 10 Suggested chromosome structure in robust GA

The initial population is created by randomly created instances of chromosomes thus, the integer constraints can be satisfied.

3-8-2 Chromosome Fitness Assessment

Once the initialization function has created an initial population and a set of index tracking portfolios as chromosomes within the constraints of the problem, the genetic algorithm begins its search to find the optimal solution. Linear Programming is used to optimize the weights of each constituent asset in the portfolio resulting in a fitness score for each chromosome. In summary the powerful search abilities of genetic algorithms and the high degree of accuracy of linear programming are used in conjunction to attain the optimal solution.

3-8-3 GA Operators

3-8-3-1 Selection Ratios

In every subsequent generation CRxP children are created using the crossover operator and MRxP children are generated using the mutation operator. Where P is the population in each generation and CR and MR are the crossover and mutation ratios. For the creation of each child chromosome parents are chosen using two distinct methods, the first of which is the roulette wheel selection process where a higher probability is given to those parents which are considered closer to the optima (to be used as the crossover parents) and the second method is the equal distribution of probability for all parents (to be mutated).

3-8-3-2 Crossover Operand

The crossover operation utilized in this model is the single point crossover (single splice) method where a singular point in the chromosome is randomly selected in both parent chromosomes. Genes in a parent chromosome up to that of the crossover point will be transferred to the child chromosome whilst the rest of the genes are pulled from the second parent. In other words, a child chromosome will have genes corresponding to two parents. How much of the underlying genes from each parent is transferred however, is controlled by the crossover point. This is shown in the figure below:

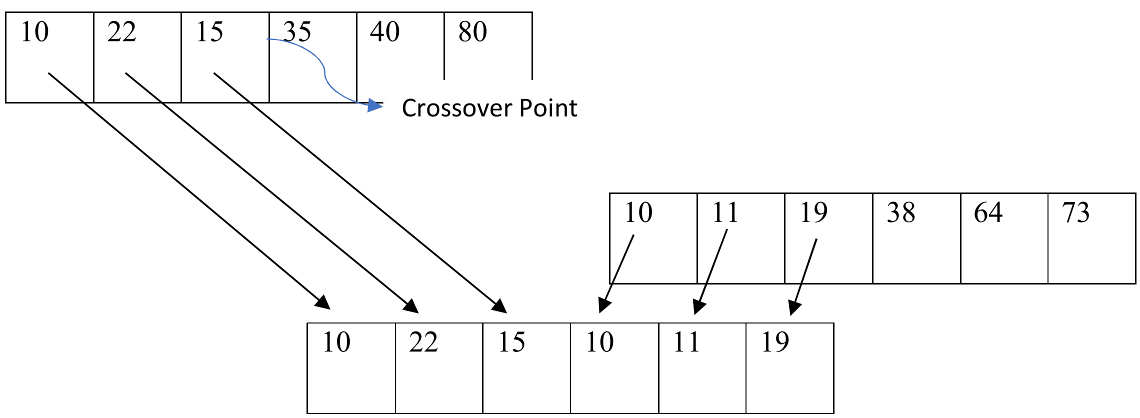

Figure 11 Crossover Operand in GA

In order to find valid solutions in regards to the research question, each chromosome must contain k number of genes corresponding to k distinct assets in an index tracking portfolio. If there exists such a gene that is present in both parents (such as the example above) then the crossover operand is coded so that only distinct genes from the second parent are transferred over to the child chromosome thus ensuring that all genes in the child's chromosomes are distinct. In regards to the research question this ensures that all underlying assets in the index tracking portfolio are distinct.

3-8-3-3 Mutation Operand

The mutation operand is inspired by natural genetic mutation and creates a child chromosome through random change in a singular chromosome. Furthermore, mutation actually changes the search space, moving the algorithm away from the local optima and increasing the quality of the final solution.

The mutation operand randomly selects a gene and another randomly chosen gene from the gene seed (available assets) replaces the randomly selected gene in the chromosome. This is shown in the figure below:

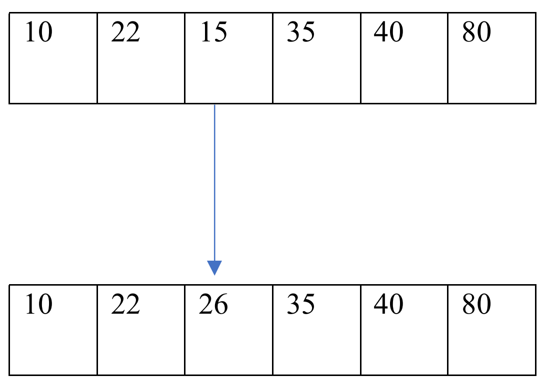

Figure 12 Mutation Operand in GA

3-8-4 Constraints Handling

As highlighted before integer constraints are handled by the population function and genetic operands designed to find an optimal solution to the research question. The proposed models' other constraints however are programmed using a linear programming library of python (PuLP), which requires that all input and decision variables and also constraints regarding the problem are codified for the inbuilt solvers of the library to produce suitable solutions.

3-8-5 Number of Generations

The proposed genetic algorithm has been set to run for 100 generations before offering the optimal calculated solution. In other words, the algorithm will attempt to calculate the optimal solution after having calculated the fitness function for 100xpopulation of chromosomes (possible solutions/portfolios) which will have naturally evolved by the use of GA algorithms through the course of each generation.

3-8-6 Comparison of Results

The genetic algorithm will provide the optimal (or close to optimal) solution which in this case is the constituent assets in an index tracking portfolio and their corresponding weights. In addition to the tracking error a schematic comparison between the resultant portfolio and that of the base index returns is presented.

3-9 Implementation of Core-Satellite Strategy

In the equation below, the objective of the core-satellite approach to investment is defined as the maximization of the expected returns of a portfolio of assets in relation to those of the base index whilst minimizing the differential variance between the returns of the portfolio and those of the index. The proposed portfolio is comprised of an index tracking fund (C) and risky assets (S) such as units of a specific fund.

$$P = wS + (1 - w)C$$

$$MaxU(w) = E(P - B) - \lambda\sigma^2(P - B)$$

Formula 50

The aim of this mathematical model is to find the optimal weight of the allocatable funds to be dispersed between the index tracking fund and more risker assets in such a way that risk is managed to the confines of the fluctuations of the base index whilst attaining higher returns than that of the index.

The first step to solving this model is to construct an index tracking fund using the genetic algorithm model proposed previously. In the next step, the top five publicly traded funds by NAV and market capital which have been traded in the timeframe of our study are chosen as the riskier assets (S). Finally, each constituent's weight and the performance of the resultant portfolio (P) is calculated in Lingo software and comparisons to the base index are made. The steps necessary to solve the model are outlined below.

3-9-1 Formulization of Core-Satellite Problem

In order to formulize the problem we must first introduce and define the notation used:

- Input Variables:

    E(B): Expected Returns of the index

    E(Si): Expected Returns of i^{th} risky asset

    E(C): Expected Returns of the index tracking fund

    $\sigma 2$(B): Index Variance

    $\sigma 2$(Si): i^{th} risky asset variance

    $\sigma 2$(C): Index tracking fund variance

    Cov(Si,B): Covariance between i^{th} risky asset and index

    Cov(Si,Sj): Covariance between risky assets i and j

    Cov(C,B): Covariance between index tracking fund and index

- Decision Variables:

W_0: Weight of the index tracking fund in portfolio P

W_i: Weight of i[th] risky asset in portfolio P

In order to solve the problem with regards to integer constraints first the problem is formalized as:

$$MaxE(P-B)-\lambda\sigma^2(P-B)$$

Subject to

$$\sum w_i + w_0 = 1$$

$$\varepsilon_i \leq w_i \leq \delta_i$$

Formula 51-1

Taking into account that portfolio (P) is made up of an index tracking fund and risky asserts we can expand the objective function using the equations below:

$$P = \sum w_i s_i + w_0 c$$

$$E(P-B) = \sum N_i E(s_i) + w_0 E(C) - E(B)$$

$$6^2(P-B) = \sigma^2(P) + \sigma^2(B) - 2\operatorname{cov}(P,B)$$

$$\sigma^2(P) = \sigma^2(\sum w_1 s_i + w_0 c)$$

$$\sigma^2(\sum \omega_i s_i + w_0 c)$$

$$= \sum w_i^2 \sigma^2(s_i) + w_0^2 \sigma^2(c) + 2\sum w_0 w_l\colon \operatorname{cov}(s_i, c) + 2\sum\sum w_i w_J \operatorname{cov}\left(s_i, s_J\right)$$

$$\operatorname{cov}(P,B) = \operatorname{cov}(\sum W_i S_i + W_0 C, B) = \sum W_i \operatorname{cov}(S_i, B) + W_0 \operatorname{cov}(C, B)$$

Formula 51-2

By substituting the equations in the objective function the final mathematical model of the core-satellite problem is proposed as:

$$Maxz = \sum N_i E(s_i) + W_0 E(C) - E(B) - \lambda_k w_i^2 6^2(s_i) + w_0^2 \sigma^2(C) + 2\sum w_0 w_i \operatorname{cov}(Si_1^C)$$
$$+2\sum\sum W_i \log \operatorname{cov}(s_i, j) + \sigma^2(B) - 2[\sum w_i \operatorname{cov}(s_i, B) + W_0 \operatorname{cov}(C, B)]$$

Subject to

$$\sum w_i + w_0 = 1$$

$$\varepsilon_i \leq w_i \leq \delta_i$$

Formula 52

3-9-2 Implementation of GA for Solving the Core-Satellite Problem

In order to solve the aforementioned problem a different chromosome structure is proposed which is made up of two parts. The first part is a k-length strand with integer members where each gene ($s_1, s_2, \ldots s_k$) is representative of the index of an asset in the fund which has randomly been selected from n available assets. For example, in a case where k=6 we select 6 assets from a possible 100. The corresponding first part of the chromosome in this case can be portrayed as in the figure below:

12	13	16	85	21	23

Figure 13 First part of proposed chromosome in the GA

The second part consists of the same number of genes as the first part where (w1,w2,…wk) represent the weights of each asset in the formation of the fund. The initial population function randomly assigns values in the (0,1) range to each gene in this section in such a way that the sum of all weights equals to one thus ensuring that the weight constraints of the equation are met.

$S_1\ S_2\ S_3 \ldots\ S_K\ W_1\ W_2\ W_3 \ldots W_k$

Figure 14 proposed Chromosome structure in classical GA

3-9-2-1 Chromosome Fitness Assessment

Once the population function has created a number of suitable solutions (series of index tracking funds) which adhere to the problem constraints as the possible search space, The genetic algorithm initially calculates the fitness of each solution and selected those with highest fitness value as potential parents for the creation of children in the next generation. It should be noted that in this problem, solutions with the lowest tracking error will have the highest fitness.

3-9-2-2 GA Operators

Selection

In each generation RCxP and RMxP children will be created through crossover and mutation where p is the population size in each generation and RC and RM correspond to crossover and mutation ratios respectively. Parents are chosen in two mechanisms the first being roulette selection where chromosomes with a higher fitness have a higher chance of being selected as parents. This method is used to select the chromosomes for the crossover operation. The second selection method is the uniform selection method where all chromosomes have the same chance of being selected. This method is chosen for the mutation operation.

Crossover Operand

The crossover operation utilized in this model is the single point crossover (single splice) method where a singular point in the chromosome is randomly selected in both parent chromosomes. Genes in a parent chromosome up to that of the crossover point will be transferred to the child chromosome whilst the rest of the genes are pulled from the second parent. In other words, a child chromosome will have genes corresponding to two parents. How much of the underlying genes from each parent is transferred however, is controlled by the crossover point. This is shown in figure 15:

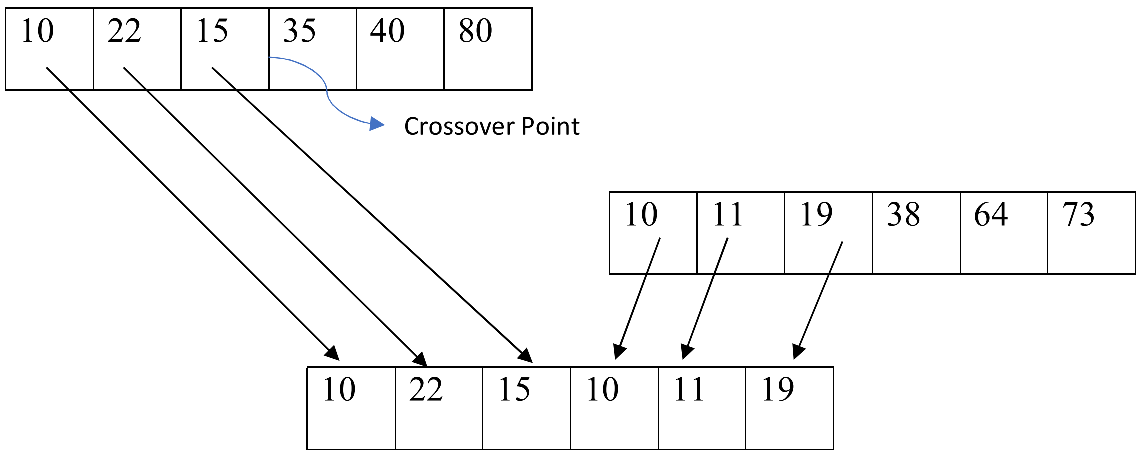

Figure 15: Crossover Operand

In order to find valid solutions in regards to the research question, each chromosome must contain k number of genes corresponding to k distinct assets in an index tracking portfolio. If there exists such a gene that is present in both parents (such as the example above) then the crossover operand is coded so that only distinct genes from the second parent are transferred over to the child chromosome thus ensuring that all genes in the child's chromosomes are distinct. In regards to the research question this ensures that all underlying assets in the index tracking portfolio are distinct.

Mutation Operand

The mutation operand is inspired by natural genetic mutation and creates a child chromosome through random change in a singular chromosome. Furthermore, mutation actually changes the search space, moving the algorithm away from the local optima and increasing the quality of the final solution.

The mutation operand randomly selects a gene and another randomly chosen gene from the gene seed (available assets) replaces the randomly selected gene in the chromosome. This is shown in the figure below:

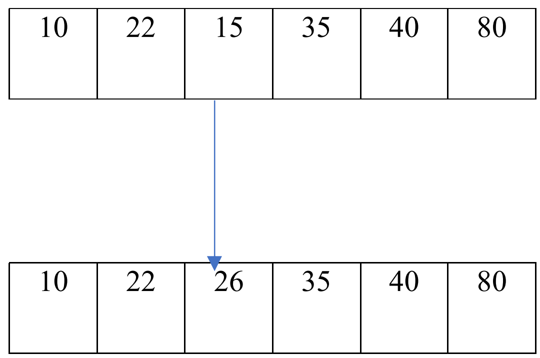

Figure 16: Mutation Operand

3-9-2-3 Constraints Handling

As previously stated, integer constraints are handled by the initial population function and the design of the crossover and mutation operations.

In order to satisfy the weights constraints, weight values are recalculated at every step and after every mutation to ensure that their sum will equal to value 1. In this approach every gene from the second part of the chromosome which contains weight values is divided by the sum of all weights existing.

3-9-2-4 Number of Generations

The proposed genetic algorithm has been set to run for 100 generations before offering the optimal calculated solution. In other words, the algorithm will attempt to calculate the optimal solution after having calculated the fitness function for 100xpopulation of chromosomes (possible portfolios) which will have naturally evolved by the use of GA algorithms through the course of each generation.

3-9-3 Comparison of Results

The resultant GA will result in a portfolio which has both core and satellite components with investment weights assigned to each individual portfolio members. By comparing the returns of the resultant portfolio against the returns of the index in a predetermined timeframe we are able to assess whether the portfolio that has been optimised by the core-satellite strategy is able to achieve higher returns than that of the market.

Chapter 4: Analysis of Results

4-1 Introduction

As markets become more competitive and big data becomes more prevalent and accessible through the advances made in technology, more and more investors are looking to leverage advanced quantitative models based on innovative optimization techniques instead of classical approaches. Taking this into account, this research attempts to solve the asset allocation problem with regards to uncertainty using these advanced techniques.

This chapter of the study introduces the data sets used and offers explanation on how they were set up in order to solve the research question and then highlights the effectiveness of robust optimization in the formation of an index tracking fund by comparison between different methodologies.

In the latter part of the chapter a general framework is proposed in order to implement the robust index tracking approach in the TSE and finally, the core-satellite approach of defining a portfolio as a combination of risky assets (satellite components) and index tracking core is explained and analysed.

The final part of this chapter will attempt to answer the research questions and theories outlined in the previous parts of this study.

4-2 Collection of Data

In this research our aim is to create an optimized investment portfolio which tracks the performance of the main market index. The market index itself is a measure of the returns of the stocks residing in the market. Preparation and adjustment of the timeseries dependent price of stocks within the market, based on capital increases or EPS payments is an inevitable part of investment strategies which we will address shortly.

EPS after ratification are funds which are released by the company and belong to shareholders. This causes the daily value of the company to fall by the same value. The theoretic re-opening price of the stock after the ratification of EPS is calculated by taking the pre-closure price of stock and subtracting the EPS from that price thus getting the theoretical opening price of the stock after the ratification. This causes some problems however as there may be cases where the subtraction of the EPS from pre-closure stock prices could result in a negative stock price. In order to solve this issue CRSP research institution has proposed a new approach which has become the standard adopted by a most of the price announcing services and websites such as Yahoo and StockCharts. In this approach prices are multiplied by a constant adjustment factor. In this study the adjustment factor is calculated in a similar way and is used to adjust the time series dependent price of stocks being studied.

$$Adjustment\ Factor = 1 - \frac{d}{P}$$

Where:

d: EPS

p: Stock price at closure before ratification of EPS

In order to remove the effects of capital increases a similar adjustment factor is used:

$$Adjustment\ Factor = 1 - \frac{\Delta}{p}$$

Where:

Δ: Difference in price before and after capital increase

P: Stock price before Capital Increase

4-3- Data Collection

The main bulk of data collection in this study is based upon daily index values and stock prices, however in order to maintain a standard within the model and also factor in consistency, a series of filters were used to lower the original number of stocks available for selection in the proposed optimised portfolio.

Each stock in the market was considered a viable choice for the portfolio if they had 100 trading days in the year or 400 for a 4-year time period. It should be noted that the 100 top stocks by market capital were exempt from this prerequisite. Stocks which did not have any price variation in a day or those that did not meet their minimum trade volume quotas were marked as not being tradeable on that day as in fact they couldn’t have any effect on the index.

Other Factors were also used to filter out the number of stocks viable for the portfolio:

- Stocks which had trading dates outside of the study period were omitted as they could have no effect on the market index.
- Stocks suffering from thin trading were also omitted as the very limited number of trades in the study period could not affect the index in a meaningful way.
- Stocks which had not gone through their initial public offering during the time period of the study.
- Stocks whose prices could not be simulated due to erratic trading and limited information. These included stocks which were suspended for a myriad of reasons or had been re-opened for trading after a long suspension during the time period of the study and would be considered as high-risk assets.

After the initial filtering, the number of stocks viable for selection in the proposed portfolio reached 232 stocks. The daily returns of the stocks during the first 3 year interval became the training set and the last years (4th year) daily returns were utilised as the testing set for the model, which were used to evaluate the performance of the optimised index tracking fund. In order to calculate the expected returns and the standard deviation of the index and the filtered 232 stocks, average daily returns and their corresponding time series were used.

Finally, the algorithm developed to create the optimized index tracking portfolio within the core-satellite approach required all viable stocks to have complete returns data for the time period of the study. 16 of the 232 stocks identified had missing daily returns values due to different issues such as trading suspensions or IPO statuses in the period of study and econometric methodologies such as EGARCH simulation had to be used to eliminate missing values and ensure that a complete range of daily returns were available for analysis by the models.

4-4 Utilised Software & Parameters

The initial data collection, sorting and filtering was done using Microsoft Excel. The modelling of the genetic algorithm and LSTM neural networks were coded in python to allow for cross compatibility with a myriad of different linear and quadratic solvers (such as PuLP and Lingo). The use of python and its many data science driven libraries such as Scipy and Numpy also allowed the direct manipulation of data as n-dimensional arrays whilst allowing for future expansion and efficiency through parallel processing. The test system which was used for the implementation and solving of each model was a windows-based system running an AMD Ryzen 3700x processor and 16GB of Ram.

Taking into account that genetic algorithms are random in nature, each problem set was solved by the models five times and the best results from the objective function and other indicative variables outlined as performance measures, were selected as results.

4-4-1 GA Parameters utilised

Keeping in mind that the genetic algorithms used to solve the index tracking portfolio problem and the core-satellite approach to investment are different in nature, different optimised parameters are identified for each one separately.

Each genetic algorithm needs to have optimised parameters regarding initial population and genetic operands and initial experiments are conducted to find the optimal values for each parameter. Crossover and mutation ratios are bound to the (0,1) range so in the case of these two parameters we can start each ratio at 0 and incrementally increase their values by 0.05 at each iteration of the experiment. The initial population number can also be experimented on by having an initial value of 5 and then increasing that to a maximum of 200.

Utilising a sequential approach to these experiments and assigning other parameters randomly at each stage will allow us to find optimal points for all three parameters. In each stage 10 possible iterations of other parameters are used to run the model and the best outcomes are noted.

After running the robust genetic algorithm using the optimised parameters for different number of generations it becomes apparent that after 100 generations the tracking error becomes stable and further generations are not effective in providing a higher quality solution. Therefore, the number of generations parameter of the robust genetic algorithm is set to 100.

Through the design of experiments outlined above, the optimal parameters for the robust genetic algorithm in the index tracking problem are highlighted below:

Number of solutions in initial population	100
Crossover ratio	0.75
Mutation ratio	0.15
Number of generations	100

Figure 17: Parameters for index tracking portfolio problem GA

The optimal parameters for the core-satellite approach to investment were identified as:

Number of solutions in initial population	150
Crossover ratio	0.8
Mutation ratio	0.2
Number of generations	100

Figure 18: Parameters for core-satellite problem GA

After the initialization of the optimised parameters in each appropriate model, the index tracking portfolio problem is solved followed by the core-satellite investment problem.

4-5 Selection of a Series of Index Tracking stocks and Optimisation Using Exact Algorithm

One of the main objectives of this research is the study of the effectiveness of robust optimisation in the formation of an index tracking portfolio with integer constraints where the performance of the portfolio mimics the performance of the index in the proposed time frame of the study. The median absolute deviation (MAD) of the expected returns of the portfolio against the index is a measure used to gauge the effectiveness of the robust optimisation. In order to assess the effectiveness of robustness on the portfolio optimisation problem, the problem is first solved with the assumption that random variables are of a certain nature. The results of the resultant portfolio are then compared with a solution to the problem that has been derived from a model which has implemented uncertainty as a factor. To this end, we utilised real world data from the TSE to compare portfolios based on factors such as root mean squared error (RMSE), Correlation ratio and the average of excess returns. Furthermore, the model offered in the third chapter of this study is solved using Lingo software in different scenarios with varying conservatism values and finally a comparison is made based on these results.

By utilising exact algorithms, we are unable to form a portfolio based on the initial selection pool of 232 stocks. By keeping in mind that the aim of solving the problem on a smaller scale is to measure the effectiveness of implementation of uncertainty in the model and the performance of its robust counterpart, a common filtering factor is used to limit the initial stock selection pool to the top 20 stocks in terms of market capital.

There are two main strategies in regards to the maintenance of the resultant portfolio and the assessment of its performance in the tracking of the index using the test range identified at the beginning of this chapter:

1. Readjustment of stock weights in the portfolio with the aim of conserving initial weight ratios.
2. Re-adjustment of stock weights based on their price fluctuations (Torrubiano & Suarez).

The first approach requires active management of the portfolio with a subsequent higher trading cost corresponding to a higher number of trades in order to maintain the initial weight ratios. The second approach is more attractive however as stocks weights are automatically adjusted based on their price fluctuations and the change in their relative value in the portfolio. In this study we utilise the second methodology of portfolio maintenance in order to assess the performance of the resultant portfolio in the tracking of the index using test data sets. Furthermore, we also set the upper and lower boundaries of investment in each stock (weights assigned to each asset) to the [0.01-0.2] range to ensure variety in the portfolio. In order to assess the relationship between the number of constituents of the index tracking portfolio and its performance criterion the research problem was solved using different k values (integer constraint) and results are highlighted below:

Integer Constraint (k)	**Correlation ratio**	**RMSE**	**Average of Excess Returns**
5	%41.37	0.0041	%0.32
10	%54.21	0.002	%0.11
15	%73.57	0.0012	%0.1
20	%77.88	0.00086	%0.02

Figure 19: Analysis of sensitivity of performance measures against different scenarios of integer constraints

As it becomes apparent in the table above, the performance criterion of the resultant portfolio improves with the increase in the number of portfolio constituents. This is due to the fact that an increase in constituents results in more similarity between the index and the portfolio which results in a 88 percent increase in the correlation ratio and nearly five times more tracking precision. An interesting point to note, is the relative reduction of the correlation gains at each step of increasing portfolio members. For example, there is only a 4 percent correlation ratio gain when increasing the number of portfolio members from 15 to 20. This becomes important to portfolio managers when trying to balance the performance gains of a portfolio by increasing constituents, against the inherent transactional costs that arise from the same increase in members.

In order to statistically assess the relationship between number of constituents of a portfolio and tracking error we have calculated Pearson's correlation coefficient to be %-93.4 rejecting the null assumption and confirming the existence of a relationship between number of stocks in a portfolio and tracking error.

According to modern portfolio theory 20 is the optimum number of constituents of a portfolio with regards to diversification. Research conducted by Elton & Gruber (1977) focused on the benefits of diversification when forming a portfolio from a possible 3290 stocks. Results show that a single member portfolio had average standard deviation (risk) of %49.2 and a 1000-member portfolio a value of %19.2. The study also showed that a portfolio with 20 randomly chosen stock of equal weights had an average risk of %20 with only a %0.8 decrease in risk resulting from the addition of all remaining stocks to the portfolio. Thus, this research utilises a maximum of 20 stocks as the components of the portfolio to be assessed and analysed.

It should also be noted that a correlation coefficient of around 80%, an average excess return close to zero and tracking error to the degree of 0.001 highlights the fact that the optimum number of constituents in an index tracking portfolio based on the TSE index is also 20 and that is well within the parameters identified by previous literature and research. Furthermore, by taking into account the direct positive relationship between number of constituents and the quality of index tracking, our analysis on the assessment of effectiveness of other parameters will be confined to portfolios having 20 members who have the best performance. In order to assess the effectiveness of robust approach in optimisation we continue to analyse the sensitivity of the value of the objective function in regards to changes in the degree of conservatism (Γ) by utilizing data from the test range. The degree of conservatism is a function of the number of uncertain variables of the model (expected returns of the stock and index) which in a 20-member portfolio will be between 0 and 21. In other words a degree of uncertainty value of 0 signifies certainty in all model variables whilst a degree of 21 highlights complete uncertainty in regards to the returns of the 20 stocks forming the portfolio and the index itself.

Degree of Conservatism	Objective Function Value	Change
0	0	0
5	0.0128	0.0128
10	0.0175	0.0047
16	0.0229	0.0046
21	0.0244	0.0015

Figure 20: Sensitivity of robust problem objective function in regard to changes in Conservatism

By reviewing the result above it becomes clear that with an increase in the degree of conservatism the value of the objective function also moves away from the optimal solution. It should be also be noted that there is a smaller deviation from the optimal solution as Γ is increased in each step which is in line with previous studies and literature. Although it could be argued that by increasing the degree of conservatism and reducing the risk tolerance of the model, will lead to deviation from the optimal solution, results and the performance indexes of the portfolio in the testing data series, actually highlight the benefits of the implementation of uncertainty constraints in the model and thus the effectiveness of robust optimization:

Degree of Conservatism (Γ)	Correlation Coefficient	RMSE	Average of Excess Returns
0	%42.37	0.00408	%0.34
5	%72.21	0.00135	%0.11
10	%75.81	0.00119	%0.10
16	%77.88	0.00086	%0.02
21	%44.19	0.00402	%0.34

Figure 21: Analysis of sensitivity of performance measures against different scenarios of Conservatism

As outlined earlier it becomes apparent that the implementation of uncertainty in the model leads to a substantial increase in the performance of the portfolio and its tracking ability. This is apparent in the decrease of RMSE and increase in the correlation coefficient as Γ is increased. Generally, it could be argued that by foregoing an implementation of uncertainty in the model (best case scenario outlook to the problem), unrealistic results stemming from deviation of expected returns of the portfolio against actual returns in the testing data range are created. On the other hand, a pessimistic outlook on the problem (complete uncertainty) leads to a loss of many potential solutions. Results from this part of the study (solving the robust counterpart of the research problem) with an adjustable degree of conservatism are in line with previous studies conducted such as Chen & Kwon (2012) and Bertsimas & sim (1999). As outlined in the previous table an optimistic outlook to the problem (fulfilment of all expected rates of return of index and portfolio members) reduces the quality of tracking in the testing period resulting in a correlation coefficient of only %42. Whereas, by taking a realistic outlook on the problem and the implementation of a certain level of uncertainty leads to significant improvement in the tracking performance criterion. Taking into account uncertainty for 16 (out of possible 21) variables causes a five-fold improvement in the quality of tracking and increases the correlation coefficient to nearly %78. Finally, the results show that a pessimistic approach to the model (uncertainty of expected returns of the index and all portfolio members) causes a significant drop in tracking quality (correlation of only %44) during the testing period.

By taking three portfolios resultant of the solving of the research question and its robust counterpart and applying a paired sample t test we get the following results:

			95% confidence interval of the difference				
P value	**df**	**T**	**Lower**	**Upper**	**Standard Error Mean**	**Standard deviation**	**Mean**
0.006	2	12.89	0.23176	0.463906	0.0269771	0.0467257	0.3478333

Figure 22: Comparison analysis of index tracking portfolios and their robust counterparts based on correlation

Taking into account the P value above, the null assumption is negated, proving that the implementation of robustness has a meaningful effect on the model. Furthermore, by highlighting the positive nature of the mean we can ascertain that robust index tracking portfolios have better performance. Finally, by reviewing the performance parameters outlined thus far we can ascertain those portfolios that have degrees of uncertainty modelled on 16 of their members to be the best index tracking portfolios.

There is another important factor of the portfolio optimisation problem which needs to be addressed; In order to efficiently maintain the portfolio, the timeframe in which the resultant portfolio is considered have an acceptable correlation rate with the index needs to be identified, so that the constituting members of the portfolio and their respective weights can be reassessed and the portfolio restructured, so that the efficiency of performance tracking is preserved. Ascertaining the optimal interval is uncertain and will depend on various market factors. Beasley et al (2003) identify a time frame of less than six months, but it should be considered that this preference is suitable for mature efficient markets which are considered less volatile, suggesting that the TSE market should have a smaller interval between reviews and readjustment. According to the study index tracking is of a suitable quality up to around 4 months (with a correlation coefficient of %77.88). In order to ascertain the length of the interval more precisely changes in the correlation coefficient of the proposed index tracking portfolio during the testing period are evaluated:

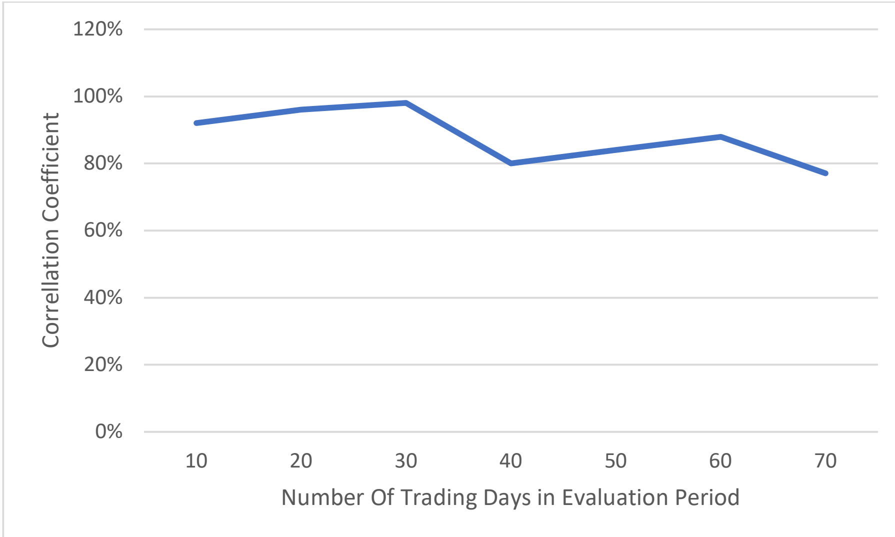

Figure 23: Changes in correlation between portfolio and index returns

Taking into account the above chart, it is argued that with a satisfactory correlation coefficient level set at %80, the resultant portfolio will need to be reviewed and adjusted at around 70-day intervals.

4-6 Formation of index tracking portfolio utilising genetic algorithms and comparison of performance against index tracking portfolios utilising robust optimisation

In order to solve the robust counterpart of the problem on a larger scale we will utilise a genetic algorithm. The benefits of the proposed genetic algorithm become apparent when the resulting portfolios are compared statistically with those of the portfolio formed using the exact optimisation algorithm of the previous section of this chapter. The performance of the portfolio formed using the proposed genetic algorithm model can also be used to gauge the efficiency of the designed model.

Furthermore, we solve the model proposed in the third chapter of this study by implementing varying degrees of conservatism and compare the resulting solutions. The implementation of

GA on all 232 stocks in the initial pool can be considered as the initial population overcoming the limitations in the exact algorithm implementation of the previous attempt which could only search for optimal solutions in a 20-stock search space.

In the initialization of the model, we set the minimum weights of each constituent in the portfolio within the [0.01-0.2] range. As observed in the previous section, the performance factors of the portfolio increased as the number of constituents was also increased, highlighting a direct relationship between the tracking quality of the portfolio and the number of its members. Thus, we will base our analysis of other influential factors on 20 member portfolios which have the best performance. Therefore, the genetic algorithm proposed is looking for the best 20-member selection of stocks out of a possible 232 which not only minimizes the expected returns of the resultant portfolio and the index (tracking error), but also attains reasonable performance metrics in the testing process.

We will also be looking to measure the effectiveness of a robust approach in optimisation by assessing the sensitivity of the objective function and performance metrics under different scenarios of conservatism using the data from the test ranges.

Degree of Conservatism	**Objective Function Value**	**Change**
0	0.000449	0
5	0.010886	0.0113
10	0.014927	0.004
15	0.01921	0.0043
21	0.020876	0.0018

Figure 24: Sensitivity of robust problem objective function towards changes in Conservatism (Γ)

The results from the above table show that with an increase in Γ, the resultant solution is moving away from the optimal. The results also show that there is less of a deviation from the optimal at higher degrees of conservatism (higher Γ lead to smaller movements away from the optimal answer). The performance metrics of the portfolio using the testing time-series of data however highlights the apparent benefits from the implementation of uncertainty (robust optimisation) in the model.

Degree of Conservatism (Γ)	**Correlation Coefficient**	**RMSE**	**Average of Excess Returns**
0	%38.28	0.00099	%-2.4
5	%52.30	0.00091	%-0.03
10	%.56.39	0.00090	%-0.2
15	%68.07	0.00086	%-0.21
21	%32.50	0.00109	%-0.09

Figure 25: Analysis of sensitivity of performance measures against different scenarios of Conservatism

As previously mentioned, the implementation of uncertainty in the model increases the portfolios' ability to track the index; With the increase of Γ there is a visible decrease in the RMSE (deviation of the expected returns of the portfolio from those of the index) and an increase in the correlation coefficient.

Generally, it could be argued that by foregoing an implementation of uncertainty in the model (best case scenario outlook to the problem) unrealistic results are created stemming from deviation of expected returns of the portfolio against actual returns in the testing data range. On the other hand, a pessimistic outlook on the problem (complete uncertainty) leads to a loss of many potential solutions. By looking at the results in the above table it becomes apparent that an optimistic outlook on the problem in the testing data range based on the realisation of expected returns of the index and portfolios stocks leads to a drop in the quality of tracking (correlation coefficient of 38%) whereas a realistic problem outlook and the implementation of uncertainty will lead to an increase in tracking performance. An implementation of uncertainty for 15 of the 21 uncertain variables in the model affects the tracking quality positively whilst increasing the correlation coefficient to 68%. On the other end of the scale a pessimistic outlook on the implementation of uncertainty for the expected returns within the testing period (where all uncertain variables including all constituents and the index are deemed to be uncertain) will lead to a drastic drop in tracking quality (correlation coefficient drops to 32%).

An important factor to note is the evident weaker performance of the index tracking portfolios formulated as a result of the genetic algorithm processing 232 stocks when compared to those formulated using the top 20 stocks by capital as the search space. The clear cause of this is that although statistical analysis in the previous section of this chapter validates the use of genetic algorithms in attaining the close to optimal solution, the widening of the search space to include stocks which are naturally of a lower market cap will lead to relative deviation of the returns of the resultant portfolios to those of the index. In other words, the optimal solution provided is a portfolio whose members have diversity and irregularity within their market caps leading to tracking error and deviation from the index within the testing range.

In practical terms the unique phenomena of buying and selling queues which occur in the TSE negate the possibility of only selecting high market cap stocks for the index tracking portfolio. Also, it is argued that a retail investor is unable to regularly create and maintain a diverse index tracking portfolio whilst only utilising high value stocks. These facts further highlight the need to select stocks from a larger solution search space through the implementation of meta-heuristic algorithms.

As previously mentioned, market capital is one of the popular metrics utilized in the selection of stocks influential to the index, when implementing an index-based investment approach. Therefore, we look to improve the performance metrics of the proposed algorithm in a way which also allows the model to select stocks from a wider search space.

An in depth look at the output of the genetic algorithm, leads to the logical detail that a number of stocks keep recurring in the final resultant portfolios. Using this logic as a base it is argued that a new multi-step algorithm could find better solutions with lower tracking errors. In other words, the search space of the genetic algorithm could now be reduced from 232 stocks to 147. In reality this approach acts as a new filtering metric based on the previous output of the genetic algorithm. By taking into account that the stocks selected by the genetic algorithm in the previous pass are those which are deemed to be influential to the index, it is expected that this new filtering metric will lead to better solutions in terms of index tracking and higher values of the correlation coefficient whilst reducing the search space complexity.

Due to the multistep nature of this new approach, which utilizes the output from one stage as the input for the next stage, it can be named the Multistage Genetic Algorithm. In order to further optimize this new algorithm against the testing range, the remaining 147 stocks are ranked by market cap value and then an exact algorithm is utilized to assign weights to the top 20 ranked stocks. The sensitivity of the multi stage genetic algorithm's objective function towards differing levels of conservatism is outlined in the table below:

Degree of Conservatism	Objective Function Value	Change
0	0.000459	0
5	0.010655	0.01111
10	0.014894	0.00424
15	0.018913	0.000462
21	0.020860	0.00134

Figure 26: Sensitivity of objective function towards changes in conservatism (multi stage GA)

Increasing the Γ value and risk tolerance of the model results in deviation from the optimal solution but the performance metrics of the resultant portfolio highlight the benefits of the implementation of uncertainty (effectiveness of robust optimization) in the model:

Degree of Conservatism (Γ)	**Correlation Coefficient**	**RMSE**	**Average of Excess Returns**
0	%42.32	0.004122	%0.33
5	%81.39	0.001043	%0.07
10	%81.39	0.001043	%0.07
16	%81.37	0.00078	%0.02
21	%44.14	0.004056	%0.33

Figure 27 Analysis of sensitivity of performance measures against different scenarios of Conservatism (multi stage GA)

Changes in the Γ value do not necessarily result in a change in the members of the portfolio and their weights, thus the tracking performance metrics may remain unchanged. This is apparent from the above results where the implementation of uncertainty in 5 and 10 variables of the model does not affect the composition of the portfolio or the members weights.

By analysing the metrics introduced, it is concluded that that the implementation of uncertainty on 16 model variables will result in the best performing index tracking portfolio which benefits from a high correlation with the index (above 80%), whilst not requiring constant maintenance and reshuffling.

In the next section of this chapter a composition of the best index tracking portfolio and the weights of its members is offered.

4-7 Proposed index tracking portfolio

The final robust index tracking portfolio consists of 20 members which have been selected using a mixture of multi stage genetic and exact optimization algorithms. Uncertainty has been implemented on 16 of the model variables.

Stock	Industry Sector	Weight
Iranian Telecommunications	Telecoms	20%
National Copper Industries	Metals & Minerals	3.9%
Esfahan Mobarakeh Steel	Metals & Minerals	3%
Omid Investment Co.	Multi Industrial	4.7%
Ghadir Investment Co.	Multi Industrial	3.2%
Pasargad Bank	Banking	1.3%
GolGohar Mining	Mining	3.5%
Chador Malo Mining	Mining	3.7%
Khuzestan Steel	Metals & Minerals	2.6%
Pardis Petrochem	Chemical	4.1%
Mellat Bank	Banking	4.1%
Saderat Bank	Banking	20%
Parsian bank	banking	4%
Tejarat Bank	Banking	4.5%
Eqtesad Novin Bank	Banking	3.8%
Esfehan Oil Refinery	Oil	4.3%
Mines and Industries Development Co	Mining	3.2%
Amir Kabir Petrochem	Chemical	1%
Khark Petrochem	Chemical	3.2%
Total		100%

Figure 28: composition of optimized robust index tracking portfolio

Generally, the use of exact optimization algorithms in the solving of the problem ensures that assets are allocated within the lower and upper boundaries defined in the model, whilst the implementation of uncertainty within the model results in the dispersion of weights throughout the constituent stocks, leading to the better performance of the robust index tracking portfolios when compared to other portfolios.

We can display the classification of constituent stocks based on their industries using the graph below:

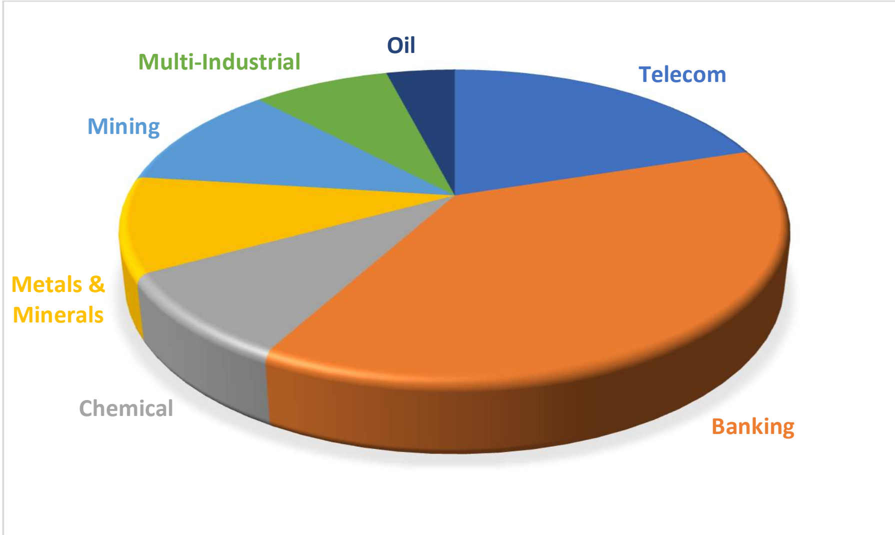

Figure 29: Composition of Robust Index Tracking Portfolio by Industry

The figure above clearly shows that the banking and telecommunication companies make up the bulk of the portfolio with the rest of the portfolio members mainly from the mining, metals and minerals and chemical sectors. Oil and multi-industrial companies round up the selected stock industry sectors. Thus, it is argued that the aforementioned sectors have the most impact on the index within the timeframe of this study.

4-8 Operational framework for the formation of a robust index tracking portfolio

Based on the results of this research we outline the proposed framework for the formation of a robust index tracking portfolio:

1. Required Data: The data required for the proposed model are the expected returns and standard deviation of the returns of stocks and the index within the studied timeframe. Taking into account that the overall objective is the tracking of the performance of the index, the returns of each individual stock needs to be adjusted to portray the inherent effects of changes in price and also other factors such as capital increases or EPS

payments. In order to calculate the expected returns, the average daily returns of stocks and index are utilised.

2. Filtering metrics in the selection process: The aim of developing suitable filtering metrics is to reduce search space complexity whilst increasing the correlation of the proposed algorithms, in order to find stocks which have the highest influence on the index. Some of the initial filtering metrics proposed include the selection of stocks with highest market cap, suitable liquidity and an absence of thin trading. Also, only stocks with at least 100 days trading days annually were deemed suitable for the initial population. Finally, the search space of the optimal solution was limited to stocks which appeared at least once in the results of the genetic algorithm.
3. Proposed Mathematical model: The proposed mathematical models in the literature depend on the absolute data (historical returns) for the formation of an index tracking portfolio through theorizing that the resultant solution of the minimisation problem will hold true for an extended period of time in the future. However, it has been shown that robust optimisation is applicable to uncertain data sets allowing us to develop a model which minimises the tracking error through the use of expected returns instead of the absolute returns. In order to assess the effectiveness of robust optimisation in the index tracking problem and also maintain the linear nature of the model the objective function is summarised as the minimisation of the deviation between the expected returns of the portfolio versus that of the index:

$$\min \left| \sum_{i=1}^{n} r_i w_i - R \right|$$

Formula 53

4. Optimum number of portfolio members: One of the main factors of a successful index tracking portfolio is its ability to balance costs and returns. In other words, one of the main objectives of the portfolio is the tracking of the index with the least amount of portfolio members possible. Based on the results of this research (which is in line with previous literature) the optimum number of constituents in an index tracking portfolio should be limited to 20 members.
5. Timeframe for re-evaluation: Due to market changes, the composition of the index tracking portfolio will need to be adjusted at specific intervals. Based on the analysis of the results of this study the optimal time frame proposed for the re-evaluation of the portfolio should be between 3 to 4 months (70 days).

Based on the framework offered in this study and the results analysed, a step-by-step schematic of the implementation of an index-based investment approach is now offered.

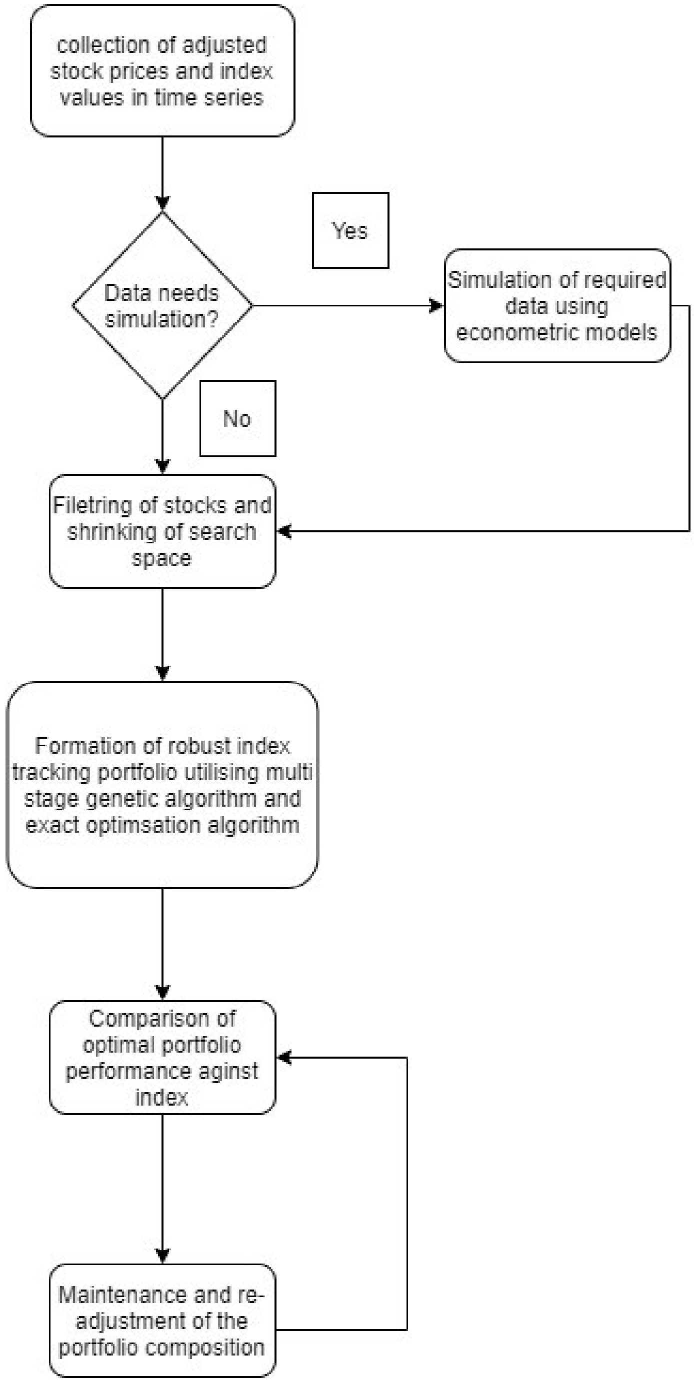

Figure 28: Data flow diagram for proposed framework

4-9 Reducing uncertainty with LSTM predictions

In this study we are looking to harness and utilise the memory state of a LSTM recursive neural network in order to lower the uncertainty factor for model variables, with the aim of minimizing the loss function value. As previously stated, all models display an increase in the loss function value when Γ (uncertainty in each model variable) is increased which makes it a viable question to see what changes the model will experience if we are able to model future expected rates of return and use them to further mitigate risk. For the purpose of this study, the index returns were selected as the model variable which the LSTM neural network would attempt to predict.

Also, it is important to highlight that the index returns were specific to 7-day intervals, so a look-back variable of 7 and look-forward variable of 1 were determined. In other words, the model is trained to look back at the 7 previous weekly returns of the index and make predictions for 1 week into the future.

As previously stated, all LSTM models include a minimum of 3 layers (input layer, hidden layer & output layer). All layers except for the output layer should include a dropout ratio to ensure that overfitting does not occur and that the model is flexible enough to react to unseen (test) data. As the individual testing of suitable dropout ratios falls outside of the scope of this study, a standard dropout ratio of 0.2 is used for all layers. The objective function is set to calculate the mean squared error and an ADAM variable learning algorithm is utilised to prevent the model from becoming stuck in local minima. The LSTM model is programmed to run for 100 epochs with results shown below:

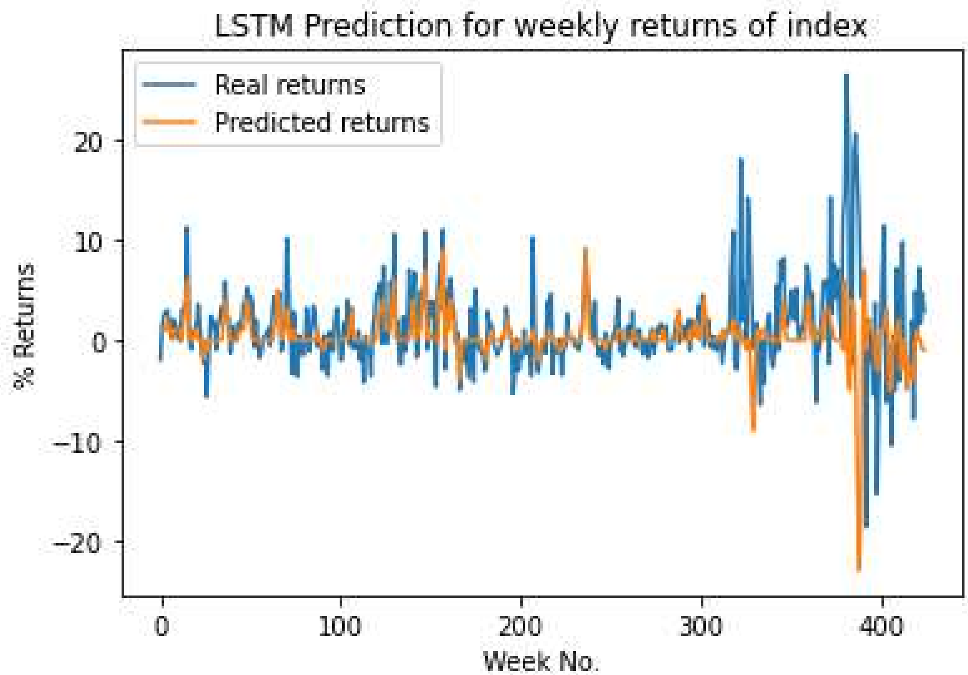

Figure 29: LSTM Predictions for Weekly Returns of Index

At first glance, the graph above looks promising because the prediction seems close to the last seen level. However, the prediction can only be logically performed one timestep in the future — the series is constructed by adding the predicted value to the end of the series, and is then used in the calculation of the next intervals prediction, effectively creating a downward bias which would negatively affect any decisions made based on the model results.

This is further confirmed in the study done by Sirignano and Cont (2018), who argue that the timeseries cannot be predicted because the data used for prediction is often limited to recent intervals, whereas financial data can be fluid and non-stationary, prone to regime changes which may render older data less relevant for future prediction.

Thus, it could be argued that there is far more going on in the stock market data than can be captured simply by looking at a univariate series of historical values. The stock prices are not the result of a couple of underlying causal factors, but a rather a multitude of contributions as well as a good dose of human irrationality (Kahneman et al, 1991). Indeed, it has been claimed that stock data is almost random (Malkiel, 1973).

4-10 Implementation of a core-satellite approach

In this study the core component of the investment approach is in fact the optimal index tracking portfolio which was formed using the meta-heuristic genetic algorithm highlighted in the previous section. The next step will be to calculate the optimal allocation of funds between risky assets (satellite component) and the core. Finally, we will assess the performance of the resultant portfolio which consists of the two core and satellite components, using testing range data based on risk and performance metrics.

4-10-1 Selection of the optimum index tracking portfolio (core) using historical data

In this stage of the core-satellite problem we look to select an optimized index tracking portfolio based on historical data (training data). The objective is to replicate the risk and returns of the index within the resultant portfolio, which is measured using the tracking error, which in this case is the RMSE value. The heuristic algorithm whose resultant portfolio mimics the returns of the market index and has the lowest tracking error will be considered as the better and more effective approach in finding the optimal solution in this problem. The classical genetic algorithm and multi-stage algorithm mentioned in the previous part of this study were compared and the results are shown below:

	Algorithm Type	**Integer Constraint**	**RMSE**	**Correlation coefficient**	**Average of excess returns**
1	Classical GA	20	0.000158	84.55%	-0.03%
2	Compound GA	20	0.000142	88%	-0.03%

Figure 30: Comparison of classical GA vs multi-stage GA in the formation of core component

It becomes apparent from the analysis of the above results that the compound GA which consists of a classical GA working in tandem with quadratic programming, performs better than a singular classical GA in terms of metrics such as tracking error and correlation coefficient. This in line with previous research conducted by Torrubiano & Suarez (2008). This higher performance can mainly be attributed to the quadratic programming component of the algorithm which is able to calculate the optimal weights assigned to each portfolio member, whereas the classical GA is only able to find solutions which are close to the optimal solution. In the next step we apply the multi-stage GA filtering process highlighted in the previous section of this chapter to shrink the search space where all stocks which have been chosen at least once in the optimal portfolio by the GA, are considered viable for the initial population. This reduces the population size from a possible 232 stocks to only 94. The compound GA is then used to find optimal solutions within this search space and finally the remaining stocks are also ranked by their market cap values and the top 20 stocks with the highest market cap are used as the input for the quadratic programming exact algorithm to allow for a comparison between these approaches. The results of the above are highlighted below:

	Algorithm Type	**Integer constraints**	**RMSE**	**Correlation coefficient**	**Average of excess returns**
1	Multi-stage GA	20	0.00018	91.32%	-0.01%
2	Adjusted multi-stage GA	20	0.000114	94.47%	0.01%

Figure 31: Comparison of multi-stage GA vs adjusted multi-stage GA

Based on the above results the simultaneous utilization of the multi-stage genetic algorithm filtering process and the subsequent ranking of search space, results in an improvement of performance metrics. The correlation coefficient now equals nearly 95% whilst the average of excess returns of the resultant portfolio now becomes positive at 0.01%. Finally, it is argued that according to the results of above analysis the adjusted multi-stage genetic algorithm should be used in the formation of the core component of the core-satellite approach.

4-10-2 Selection of the risky assets for the satellite component

In this stage data from www.fipiran.com is used to identify the top five investment funds based on NAV and average annual rates of returns for the previous year. Afterwards Lingo is used to solve the core-satellite problem (assignment of weights) for varying degrees of risk tolerance(λ). The output of this stage will be the optimal weights assigned to the index tracking portfolio (core) and risky assets.

λ	Correlation Coefficient	RMSE	Average of excess returns
0	83.79%	0.000695	14.31%
20	94.71%	0.00042	13.62%
30	95.04%	0.000417	13.66%
40	95.28%	0.00041	13.07%
50	95.38%	0.000407	12.67%
70	95.42%	0.000405	12.21%
100	95.44%	0.000404	11.83%
200	95.62%	0.000382	9.98%

Figure 32: Analysis of sensitivity of performance metrics to changes in risk tolerance for core-satellite model

It becomes apparent that an increase in the risk tolerance rating leads to a deviation from the optimal value in the objective function (lower returns than that of the index). This is due to the fact that as risk tolerance is increased the model allocates a higher weight to the index tracking (core) component of the model which in turn will lead to lower tracking errors and higher correlation coefficients whilst aligning fluctuation of the portfolio to those of the index. As can be seen from the result above a risk tolerance of 0 ($\lambda = 0$) leads the model to allocate 100% of the funds available to a singular fund gaining the highest returns (in relation to the index) at 14%. On the other end of the scale with a risk tolerance value of 200 a substantial ratio of funds available (53.33%) is allocated to the index tracking portfolio (core component) which due to the higher correlation coefficient with the index returns only yields an excess return of 9.98%.

The main objective of the core-satellite approach to investment is to form a portfolio consisting of a core index tracking component coupled with risky satellite assets. Weights should be assigned to each of these components in such a way that it yields returns higher than those of the index whilst controlling the systematic risk inherent in the market. Taking into account that there is a need to balance excess yields and risk we argue that the portfolio with risk tolerance factor of 20 ($\lambda = 20$) is superior to others due to the fact that it manages an excess return of 13.63% whilst achieving high correlation with a minimum tracking error value.

It should also be noted that there is a need to set lower and upper ranges for the weights of the core and satellite components. According to the results of this study setting a minimum of 1% and a maximum of 50% for the weights assigned to each asset results in a correlation coefficient of 94.85%, a RMSE value of 0.000411 and excess returns of 13.19% in relation to the index.

Chapter 5: Conclusions & Recommendations

5-1 Introduction

The final chapter of this research focuses on the fundamental theories utilised in the study and offers an overview of the steps undertaken to complete the research. Later, a discussion on the results and conclusions of the research is offered and the final part of the chapter aims to provide a series of recommendations to market participants (investors and financial institutions) whilst introducing new themes for future research.

5-2 Research Summary

By comparing active and passive investment strategies it becomes clear that each investment paradigm is best suited for different market scenarios. The existence of some level of inefficiency in the market validates the use of active management. However, investors need a sufficient level of experience and expertise to fully utilise this inefficiency and if found lacking, high account management fees and transactional costs will result in the investment becoming a zero-sum game. In other words, it could be argued that investors that are unable to identify and select high performing assets need to adopt a passive investment strategy that can compound returns and exploit the market inefficiency by generating a portfolio of assets that is able to sufficiently manage risk and also uncertainty.

In order to provide a suitable investment framework (which is one of the main objectives of this study) it becomes apparent that a suitable investment strategy must take into account risk reduction factors and minimization of trading frequency and trading costs. When identifying that these factors are the underlying principles of index-based investment approaches it soon becomes clear that these passive strategies are the most efficient approach in attaining the returns of the market.

The resultant framework not only tracks the index by minimising the deviation of expected returns of the fund and the index, but also utilises robust optimisation in an attempt to implement uncertainty within the model. The proposed model has a series of unique advantages, some of which include a robust counterpart linear formulisation, reduced complexity through the utilisation of duality principles and finally the fact that standard solver software can be used to solve the optimisation problem. Specific metrics such as correlation coefficient, RMSE and the average excess of returns were calculated using real world data from the TSE to assess the tracking performance of the model.

The importance of risk management becomes apparent in the development of the Modern Portfolio theory providing mathematical proof that appropriate diversification can minimise portfolio variation at the given return level, whilst quantifying the trade-off between risks and returns. Thus, creating a set of efficient portfolios that maximise returns at the given risk level.

Developments in the fields of quantitative optimisation techniques and mathematical modelling are some of the most invaluable tools in an investor's arsenal. Ever since their introduction in 1952, they have had a profound impact on the financial sector as a whole, with investors looking to identify the most effective and efficient models whilst reducing prediction errors. The underlying principles of portfolio optimisation and mathematical modelling have been a constant driving force in market development, adoption, risk management and financial decision support processes.

By reviewing the relevant literature, it becomes clear that these mathematical models suffer from two main weaknesses. The first is their inability in complex constraint handling which usually requires a conversion of the base model into a non-linear format. Another of the main challenges in utilising these quantitative models is the uncertainty of the input variables of the model. Numerous studies have been conducted on finding an optimal solution to the portfolio optimisation problem which focus on uncertainty, resulting in the introduction of phased and robust models. The current research has attempted to discuss and further develop these models with the specific aim of implementing robustness into its innovative portfolio optimisation model. By implementing robustness as a degree of uncertainty in the objective function of the proposed mathematical model, market participants are able to manage uncertainty whilst maximizing validity of the end results.

The introduction of integer constraints within the portfolio optimisation problem significantly increases the complexity of the model making it very hard to solve. The complex nature of the model prompts the use of heuristic modelling software (in this case a genetic algorithm) in order to solve the problem, however the very stochastic nature of these approaches means that there is no guarantee in finding the optimal solution. With this reasoning we first solve the problem on a smaller scale using exact and genetic optimisation algorithms separately and after having statistically tested the convergence of the proposed heuristic model, utilise it to solve the research problem on a larger scale. In the next stage, the genetic algorithm is further developed to improve the performance of the robust index tracking portfolio resulting in the introduction of the proposed multi-stage genetic algorithm.

The final heuristic model utilises a genetic algorithm to search the solution space whilst quadratic programming is used to calculate the optimum weights assigned to each member of the portfolio. Furthermore, by programming the genetic algorithm to employ a multi-stage search of the solution space and defining a new initial solution population based on its previous output, the model was able to significantly reduce the search space of the optimal solution. In addition, by ranking the remaining securities in the resultant search space based on market capitalization and utilising an exact optimisation algorithm bound to the integer constraints of the portfolio optimisation problem, there was a substantial increase in convergence towards the optimal solution.

Finally, in order to maximize returns and allow investors to take advantage of the inefficiencies of the market, the portfolio optimisation problem was expanded upon to accommodate the implementation of a core-satellite investment approach. Modelling of the core-satellite problem was achieved through the use of a multi-objective function, aiming to maximise the average excess returns of the portfolio whilst minimizing its variance in relation to the returns of the index. In order to determine the composition of the risky assets in the satellite component of the portfolio, a fund within a fund concept was used. In other words, units of trade-able existing managed funds were utilised.

5-3 Research Results

The results of this study research showed that an index tracking portfolio with a moderate level of risk tolerance outperforms others and that there is a direct relationship between the number of constituents of an index tracking portfolio and its ability in tracking the index. The results also showed the optimal number of portfolio constituents to be 20 and that the optimal timeframe for the re-evaluation of the portfolio to be between 3 and 4 months (70 days).

By solving the problem under varying scenarios of risk tolerance values and making statistical analysis we were able to confirm the effectiveness of robust optimisation in the context of an index tracking portfolio selection problem. It should be noted that the resultant portfolio scored highly in all performance metrics for data outside the training range (testing range data used) with a correlation coefficient of over 80% and tracking accuracy of 0.0001 which further highlights the unique benefits of utilising this heuristic approach.

By studying the implementation of robustness in the proposed framework it becomes clear that robust models with higher degrees of conservatism perform better than their non-robust counterparts. It should be noted however that assigning uncertainty to all variables will lead to a drastic drop in performance. The performance metrics show that a realistic outlook on the value of the degree of uncertainty needs to be employed. The results of the research highlight that out of a possible 21 variables (20 portfolio members and the returns of the index) the model which attributed uncertainty to 15 variables had the highest correlation coefficient.

Analysis of the composition of the resultant robust index tracking portfolios showed that the banking, telecommunications, chemical, mining, metals and multi-industrial sectors had the most funds allocated in the portfolio making them the most influential on the index within the timeframe of the study.

The implementation of the core-satellite strategy significantly improved the performance of the resultant portfolio attaining excessive returns of 13.19% whilst previous research limited returns to 2-3% (Bonucci & Queesenberry, 2006).

5-4 Research Recommendations

5-4-1 Recommendations based on Research Results

The framework developed in this research is aimed towards both novice and professional investors who are looking to develop passive index tracking products, giving them access to a powerful decision support toolset which benefits from low transactional cost and risk. The inherent properties of the implemented model make it well suited to the needs of novice investors looking to enter the market with limited experience.

The following recommendations are given to both sets of users which can benefit from the proposed framework:

1. Taking into account the progresses made in the field of information technology and the huge mass of data available in financial markets, it is advised to both retail and institutional investors to utilize the multi-stage compound genetic algorithm outlined in this study as a machine learning tool to lower transactional costs and minimize risk whilst attaining favorable returns.

2. Taking into account the high correlation between number of constituent stocks in an index tracking portfolio and tracking error, it is recommended that diverse portfolios containing 20 stocks be selected. The resultant portfolios should be reassessed every 4 months for optimal performance.

3. Novice investors should look to minimize tracking error which is a low-risk investment strategy. This is one of the inherent benefits of an index tracking approach. Institutional investors should however should look to use the above as a platform to build upon by implementing a core-satellite portfolio to attain higher returns.

4. As the proposed algorithms in this study are fully autonomous, they could be used by governing and regulatory bodies as tools to inform and guide new investors in the market, increasing liquidity in the market whilst ensuring market penetration within new sectors.

5. Investment companies and hedge funds are advised to allocate a portion of their assets to the index-based approach proposed in this study in order to lower their transactional costs and attain returns higher than that of the index.

5-4-2 Recommendations for Further Research

By reviewing the results and limitations of the present research the following recommendations are made for future research into the field of heuristic portfolio optimisation:

1. The minimization of transactional costs is an important factor in the portfolio optimisation problem and although the inherent nature of index tracking funds tackles this issue to some extent, further research could be conducted in how to adjust the existing model so that transactional costs and tax are also mathematically modeled as constraints.
2. Further research should be conducted on the feasibility of using other tradeable financial assets to track the market index.
3. Model could be expanded to contain cash reserves as a tertiary segment of the portfolio with constraints set up to aid cash flow management for fund managers.
4. This research focuses on the use of and adjusted genetic compound genetic algorithm in order to solve the optimization problem. New innovative techniques such as particle swarm analysis should be implemented to measure and compare the efficiency of the resultant model.
5. The modelling the core-satellite approach in conditions of uncertainty and the implementation of robust optimization should also be studied further.
6. This research determined units of tradeable funds to be the risky component of the core-satellite approach. Further research could focus on the substitution of other risky assets such as gold, foreign currency etc.

7. The modelling of the core-satellite problem as a dynamic model with the ability to adjust the permittable deviation from the index as a way to control risk and returns is also an attractive prospect for future research.
8. It should be noted that the significant increase in performance of the core-satellite framework could be linked to the inefficiency of the Iranian market. Further studies could be conducted to assess the performance of the proposed framework in mature and efficient markets.

References

1. *1.8 Million Shareholder Discrepancy Between National Privatisation Organisation & Justice Share Participants.* BoursePress.ir. https://boursepress.ir/m.aspx?newsid=154125

2. *A Statistical Insight Into the Populations Rush Towards The Stock Market.* (2020, May). Bourse24, https://www.bourse24.ir/news/197248

3. Adcock, C., & Meade, N. (1994). *A simple algorithm to incorporate transactions costs in quadratic optimisation*. European Journal of Operational Research, 79(1), 85-94.

4. Ahmadzadeh, A., Yavari, K., Issaii-Tafreshi, M., SalehAbadi, A., (2014), *Analysis of Assessing the Efficiency of the Iranian Capital Market*, Journal of Economic Modelling, Vol 17

5. Akbarifard, H., Alayi, R., Anaraki-Mohammadi, A., (2017). *Portfolio Optimsation Utilsiing Metaheuristic Algorithms.* Tehran Stock Exchange, Volume 10, No. 38, 87-110

6. Allahyari, A., (2009), *Analysis of the Weak Form Efficiency in the Iranian Stock Market*, Tehran Stock Exchange Quarterly, Vol 4, 75-108

7. Amenc, N., Goltz, F., & Grigoriu, A. (2010). *Risk Control through Dynamic Core-Satellite Portfolios of ETFs:Applications to Absolute Return Funds and Tactical Asset Allocation*. The Journal of Alternative Investments.

8. Amenc, N., Malaise, P., & Martellini, L. (2004). Revisiting Core-Satellite Investing. The Journal of Portfolio Management, 31(1), 64-75.

9. Amini, A. (2018). *MIT 6.S191: Introduction to Deep Learning* [Video file]. Retrieved from https://www.youtube.com/playlist?list=PLtBw6njQRU-rwp5__7C0oIVt26ZgjG9NI

10. Asoroosh, A., Atrchi, R., Ramtinnia, S., (2017). *Portfolio Optimization Using Teaching-Learning Based Optimization (TLBO) Algorithm in Tehran Stock Exchange (TSE).* Journal of Financial Research, Tehran University Press, Vol 19, No. 2, 263-280

11. Aven, T., & Hiriart, Y. (2013). *Robust optimization in relation to a basic safety investment model with imprecise probabilities*. Safety Science, 55, 188-194.

12. Bachelier, L., (1900), *Speculation Theory,* Scientific Annals of the École Normale Supérieure

13. Bahr-Oloom, M., Tehrani, R., & Hanifi, F. (2012). *Design of a Heuristic Algorithm in the Portfolio Optimisation Problem*. Journal of Accounting IRI, 13, 20-43.

14. Bamberg, G., & Wagner, N. (2000). *Equity index replication with standard and robust regression estimators*. OR Spectrum, 22(4), 525-543.

15. Banz, R.W., (1981), *The Relationship Between Return and Market Value of Common Stocks,* Journal of Financial Economics, Vol 9, No 1, 3-18

16. Barasod, M., Zomorodian, G., (2019), *Investigation into the Behaviour of Investors When Selecting Stock Portfolios,* Financial Engineering Group, Journal of Investing Knowledge, Tehran Azad University, Vol 8, No 29, 101-114

17. Basu, S., (1977), *Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis,* The Journal of Finance, Vol 32, No 3, 663-682

18. Beasley, J., Meade, N., & Chang, T. (2003). *An evolutionary heuristic for the index tracking problem*. European Journal of Operational Research, 148(3), 621-643.

19. Beightler, C., Phillips, D., & Wilde, D. J. (1979). *Foundations of Optimization* (2nd ed.). Prentice Hall.

20. Benartzi, S., Thaler, R., (1995), *Myopic Loss Aversion and the Equity Premium Puzzle*, The Quarterly Journal of Economics, Vol 110, No 1, 73-92

21. Ben-Tal, A., & Nemirovski, A. (1998). *Robust Convex Optimization*. Mathematics of Operations Research, 23(4), 769-805.

22. Ben-Tal, A., & Nemirovski, A. (1999). *Robust solutions of uncertain linear programs*. Operations Research Letters, 25(1), 1-13.

23. Ben-Tal, A., & Nemirovski, A. (2000). *Robust solutions of Linear Programming problems contaminated with uncertain data*. Mathematical Programming, 88(3), 411-424.

24. Ben-Tal, A., Bertsimas, D., & Brown, D. B. (2010). *A Soft Robust Model for Optimization Under Ambiguity*. Operations Research, 58(4-part-2), 1220-1234.

25. Ben-Tal, A., Ghaoui, L. E., & Nemirovski, A. (2009). *Robust Optimization*. Princeton, NJ: Princeton University Press.

26. Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). *Adjustable robust solutions of uncertain linear programs*. Mathematical Programming, 99(2), 351-376.

27. Ben-Tal, A., Margalit, T., & Nemirovski, A. (2000). *Robust Modeling of Multi-Stage Portfolio Problems*. Applied Optimization, 303-328.

28. Bertsimas, D., & Pachamanova, D. (2008). *Robust multiperiod portfolio management in the presence of transaction costs*. Computers & Operations Research, 35(1), 3-17.

29. Bertsimas, D., & Sim, M. (2004). *The Price of Robustness*. Operations Research, 52(1), 35-53.

30. Besley, S., & Brigham, E. (2008). *Principles of Finance*. Boston, MA: Cengage Learning.

31. Block, H. (1970). *A review of perceptrons: An introduction to computational geometry*. Information and Control, 17(5), 501-522.

32. Bouchey, P. (2010). *Tax Efficient Investing in Theory and Practice*. Seattle: Parametric Ins.

33. Brindle, A., (1981). *Genetic Algorithms for function optimisation,* PHD thesis, University of Alberta

34. Brinson, G.P, Hood, L.R., Beebower, G.L., (1986), *Determinants of Portfolio Performance*, Financial Analysts Journal, Vol 42, No 4, 39-44

35. Cadarso, L., & Marín, Á. (2012). *Recoverable Robustness in Rapid Transit Network Design*. Procedia - Social and Behavioral Sciences, 54, 1288-1297.

36. Cai, H., Schmidt, A.B., (2020). *Comparing Mean Variance and Equal Weights Portfolios for Major US Equity Indexes,* Journal of Asset Management, 21, 326-332

37. Canakgoz, N., & Beasley, J. (2009). *Mixed-integer programming approaches for index tracking and enhanced indexation*. European Journal of Operational Research, 196(1), 384-399.

38. Carrizosa, E., Goerigk, M., & Schöbel, A. (2017). *A bi-objective approach to recoverable robustness based on location planning*. European Journal of Operational Research, 261(2), 421-435.

39. Central Bank of Iran. (2018). *CBI Annual Report of Loans to Deposits Ratio*. Retrieved from CBI website: https://www.cbi.ir/category/6068.aspx

40. Ceria, S., & Stubbs, R. A. (2006). *Incorporating estimation errors into portfolio selection: Robust portfolio construction*. Journal of Asset Management, 7(2), 109-127.

41. Chang, T., Yang, S., Chang, K. (2009). *Portfolio optimization problems in different risk measures using genetic algorithm*. Expert Systems with Applications, 36(7), 10529-10537.

42. Chen, C., Kwon, R. H. (2012). *Robust portfolio selection for index tracking*. Computers & Operations Research, 39(4), 829-837.

43. Cheref, A., Artigues, C., & Billaut, J. (2016). *A new robust approach for a production scheduling and delivery routing problem*. IFAC-PapersOnLine, 49(12), 886-891.

44. Cicerone, S., D'Angelo, G., Di Stefano, G., Frigioni, D., & Navarra, A. (2007). *Robust algorithms and price of robustness in shunting problems*. Proceedings of the 7th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, ATOMS (2007).

45. Coleman, T., Li, Y., & Henniger, J. (2006). *Minimizing tracking error while restricting the number of assets*. The Journal of Risk, 8(4), 33-55.

46. *Comprehensive Iran Sanctions, Accountability, and Divestment Act (CISADA)*, H.R 2194, 110th Cong. (2010), https://www.govinfo.gov/content/pkg/PLAW-111publ195/html/PLAW-111publ195.htm

47. *Consumer Price Index in the Month of Mordad of the year 1400.* (2021, August), Statistical Centre of Iran. https://www.amar.org.ir/english/SCI-News-Archive/ID/15776/Consumer-Price-Index-in-the-Month-of-Mordad-of-the-Year-1400

48. Corielli, F., & Marcellino, M. (2006). *Factor based index tracking*. Journal of Banking & Finance, 30(8), 2215-2233.

49. Crevier, D. (1993). *AI: the tumultuous history of the search for artificial intelligence*. New York, NY: Basic Books (AZ).

50. Davis, L., (1991). *Handbook of Genetic Algorithms*, Van Nostrand Reinhold

51. De Bondt, W.F.M., Thaler, R., (1985), *Does the Stock Market Overreact?* The Journal of Finance, Vol 40, No 3, 793-805

52. Dehdar, F., AliAghayi, M., (2009), *A Study on the Weak Form of Efficiency in Emerging Capital Markets: Evidences from Tehran Stock Exchange (TSE),* Journal of Business Strategy, Shahed University Press, Vol 7, No 38, 59-70

53. DeMiguel, V., & Nogales, F. J. (2009). *Portfolio Selection with Robust Estimation*. Operations Research, 57(3), 560-577.

54. Dose, C., & Cincotti, S. (2005). *Clustering of financial time series with application to index and enhanced index tracking portfolio*. Physica A: Statistical Mechanics and its Applications, 355(1), 145-151. doi:10.1016/j.physa.2005.02.078

55. Du, K.L., Suramy, M.N., (2014). *Neural Networks and Statistical Learning,* Springer Text

56. Duchi, J., Hazan, E., & Singer, Y. (2011). *Adaptive Sub-gradient Methods for Online Learning and Stochastic Optimization*. Journal of Machine Learning Research, 12, 2121-2159.

57. Ehrhardt, M., & Brigham, E. (2008). *Corporate Finance: A Focused Approach*. Boston, MA: Cengage Learning.

58. El Ghaoui, L., & Lebret, H. (1997). *Robust Solutions to Least-Squares Problems with Uncertain Data*. SIAM Journal on Matrix Analysis and Applications, 18(4), 1035-1064.

59. El Ghaoui, L., Oustry, F., & Lebret, H. (1998). *Robust Solutions to Uncertain Semidefinite Programs*. SIAM Journal on Optimization, 9(1), 33-52.

60. Ellis, C. D. (1975). *The Loser's Game*. Financial Analysts Journal, 31(4), 19-26.

61. Elton, E. J., & Gruber, M. J. (1977). *Risk Reduction and Portfolio Size: An Analytical Solution*. The Journal of Business, 50(4), 415.

62. Elton, E. J., Gruber, M. J., & Brown, S. J. (2013). *Modern Portfolio Theory and Investment Analysis, 9th Edition: Ninth Edition*. Hoboken, NJ: Wiley Global Education.

63. Elton, E., Gruber, M., (1997). *Modern Portfolio Theory, 1950 to Date,* Journal of Banking & Finance, 21, 1743-1759

64. Fallahpour, S., Asefi, S., Fallah Tafti, S., BagheriKazemabad, M.R., (2018). *Portfolio Optimization Using the Whale Algorithm with Expected Shortfall as the Measure of Risk.* Financial Engineering and Portfolio Management, Islamic Azad University, Vol 9, No. 37, 110-132

65. Fama, E. F., & French, K. R. (2015). *A five-factor asset pricing model*. Journal of Financial Economics, 116(1), 1-22.

66. Fama, E.F, (1991), *Efficient Capital Markets: II*, The Journal of Finance, Vol 46, No 5, 1575-1617

67. Fama, E.F., (1965), *The Behaviour of Stock-Market Prices,* The Journal of Business, Vol 38, No 1, 34-105

68. Fama, E.F., (1970), *Efficient Capital Markets: A Review of Theory and Empirical Work*, The Journal of Finance, Vol 25, No 2, 383-417

69. Fastrich, B., & Winker, P. (2010). *Robust portfolio optimization with a hybrid heuristic algorithm*. Computational Management Science, 9(1), 63-88.

70. Feghhi Kashani, M., Seyyed Shokri, K., Khoshnood, Z., Akbari Alashti, T., (2013). *Development of Financial Institutions in Iran: Emphasis on the Guidelines Needed for the Formation and Operation of New Financial Institutions,* Monetary & Banking Research Institute

71. Fernandez, A., Gomez, S., (2007). *Portfolio Selection Using Neural Networks.* Computers and Operations Research, Vol 34, 1177-1191

72. *Financial Institutions Act.* Monetary and Credit Congress. (2009). https://www.cbi.ir/page/12777.aspx

73. Fischer, D. E., & Jordan, R. J. (1999). *Security Analysis and Portfolio Management* (5th ed.).

74. Fischetti, M., & Monaci, M. (2009). *Light Robustness*. Robust and Online Large-Scale Optimization, 61-84.

75. Fliedner, T., Liesiö, J., (2016). *Adjustable robustness for multi-attribute project portfolio selection*, European Journal of Operational Research, Vol 252, No 3, 931-946

76. Floudas, C. A., & Visweswaran, V. (1995). *Quadratic Optimization*. Nonconvex Optimization and Its Applications, 217-269.

77. Focardi, S. M., & Fabozzi 3, F. J. (2004). *A methodology for index tracking based on time-series clustering*. Quantitative Finance, 4(4), 417-425.

78. Framework for the Distribution of Justice Share Act. 226037, 9th Majlis. (2006). https://rc.majlis.ir/fa/law/show/135201

79. Frank, M., & Wolfe, P. (1956). *An algorithm for quadratic programming*. Naval Research Logistics Quarterly, 3(1-2), 95-110.

80. Frankfurter, G. M., & McGoun, E. G. (1996). *Toward Finance of Meaning*. The Journal of Investing, 5(3), 7-13.

81. Froot, K.A., Thaler, R., (1990), *Anomalies: Foreign Exchange*, The Journal of Economic Perspectives, Vol 4, No 3, 179-192

82. Fukushima, K. (1980). Neocognitron: *A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position*. Biological Cybernetics, 36(4), 193-202.

83. Gagniuc, P. A. (2017). *Markov Chains: From Theory to Implementation and Experimentation*. Hoboken, NJ: John Wiley & Sons.

84. Gaivoronski, A. A., Krylov, S., & Van der Wijst, N. (2005). *Optimal portfolio selection and dynamic benchmark tracking*. European Journal of Operational Research, 163(1), 115-131.

85. García, J., & Peña, A. (2018). *Robust Optimization: Concepts and Applications. Nature-inspired Methods for Stochastic, Robust and Dynamic Optimization*. InTechOpen

86. *GDP (current US$) - Iran, Islamic Rep.* The World Bank. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?end=2020&locations=IR&start=2000

87. Gen, M., & Cheng, R. (1999). *Genetic Algorithms and Engineering Optimization*. Hoboken, NJ: John Wiley & Sons.

88. Gharekhani, M. (2012). *A Framework for Portfolio Management Utilising Robust Optimisation* (Unpublished doctoral dissertation). Science and Industry University, Tehran, Iran.

89. Gilli, M., & Kellezi, E. (2001). *The Threshold Accepting Heuristic for Index Tracking*. SSRN Electronic Journal. Available at SRN: https://ssrn.com/abstract=1317358

90. Goerigk, M., & Schöbel, A. (2014). *Recovery-to-optimality: A new two-stage approach to robustness with an application to aperiodic timetabling*. Computers & Operations Research, Vol 52, Part A, 1-15.

91. Goerigk, M., Hamacher, H. W., & Kinscherff, A. (2018). *Ranking robustness and its application to evacuation planning*. European Journal of Operational Research, 264(3), 837-846.

92. Goetzmann, K., Stiller, S., & Telha, C. (2012). *Optimization over Integers with Robustness in Cost and Few Constraints*. Approximation and Online Algorithms, 89-101.

93. Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.

94. Goldfarb, D., & Iyengar, G. (2003). *Robust Portfolio Selection Problems*. Mathematics of Operations Research, 28(1), 1-38.

95. Graham, B., Dodd, D. (1962). *Security Analysis: Principles and Technique*. McGraw-Hill

96. Gregory, C., Darby-Dowman, K., & Mitra, G. (2011). *Robust optimization and portfolio selection: The cost of robustness*. European Journal of Operational Research, 212(2), 417-428.

97. Guastaroba, G., Mansini, R., Ogryczak, W., & Speranza, M. (2016). *Linear programming models based on Omega ratio for the Enhanced Index Tracking Problem*. European Journal of Operational Research, 251(3), 938-956.

98. Guastaroba, G., Mitra, G., & Speranza, M. G. (2011). *Investigating the effectiveness of robust portfolio optimization techniques*. Journal of Asset Management, 12(4), 260-280.

99. Halldórsson, B., & Tütüncü, R. (2003). *An Interior-Point Method for a Class of Saddle-Point Problems*. Journal of Optimization Theory and Applications, 116(3), 559-590.

100. Hanifi, F., Bahr-Oloom, M. M., & Javadi, B. (2009). *Comparison of Heuristic Algorithms for the Formulation of Index Tracking funds in TSE*. Management Perspective IRI, 32, 89-108.

101. Harlow, W. (1991). *Asset Allocation in a Downside-Risk Framework*. Financial Analysts Journal, 47(5), 28-40.

102. Haugen, R. A., & Baker, N. L. (1990). *Dedicated stock portfolios*. The Journal of Portfolio Management, 16(4), 17-22

103. Haupt, R. L., & Haupt, S. E. (2004). *Practical Genetic Algorithms* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

104. Hejazi, R., Jaffari Sarasht, D., & Delshadi, M. (2012). *Optimised asset allocation utilising genetic algorithms*. Bourse Stock Exchange Journal, 4(14), 135-157.

105. Hicks, J.R., (1935), *A Suggestion for the Simplifying the Theory of Money,* Economica (February), 1-19

106. Holland, J. H. (1992*). Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge, MA: MIT Press.

107. Huang, D., Zhu, S., Fabozzi, F. J., & Fukushima, M. (2008). *Portfolio selection with uncertain exit time: A robust CVaR approach*. Journal of Economic Dynamics and Control, 32(2), 594-623.

108. Huang, D., Zhu, S., Fabozzi, F. J., & Fukushima, M. (2010). *Portfolio selection under distributional uncertainty: A relative robust CVaR approach*. European Journal of Operational Research, 203(1), 185-194.

109. Huang, X. (2008). *Portfolio selection with a new definition of risk*. European Journal of Operational Research, 186(1), 351-357.

110. Hunton, J.E. (2005). *Toward an Understanding of the Risky Choice Behaviour of Professional Financial Analysts*. Journal of Psychology and Financial Markets, 2 (4), 182–189

111. *Iran Libya Sanctions Act*, H.R. 3107, 104th Cong. (1996), https://www.congress.gov/bill/104th-congress/house-bill/3107/text

112. Izenman, A.J., (2009). *Modern Multivariate statistical Techniques: Regression, Classification and Manifold Learning,* Springer text

113. Jansen, R., Van Dijk, R. (2002). *Optimal Benchmark Tracking with Small Portfolios*. The Journal of Portfolio Management, 28(2), 33-39.

114. Jenkins, L., (2000). *Selecting Scenarios for Environmental Disaster Planning*. European Journal of Operational Research, Vol 121, No 2, 275-286

115. Jensen, P. A., Bard, J. F. (2015). *NLP Method*. Retrieved from www.me.utexas.edu/~jensen/ormm/supplements/methods/nlpmethod/s2_quadratic.pdf

116. Jeurissen, R. (2005). *A Hybrid Genetic Algorithm to Track the Dutch AEX-index* (Unpublished master's thesis). University of Rotterdam.

117. Jouzbarkand, M., Panahian, H., (2020). *Evaluation of Market Efficiency Using Advanced Econometric Models in Tehran Stock Exchange*. Financial Engineering & Portfolio Management, Department of Accounting, Kashan Branch, Islamic Azad University, Kashan, Iran, Vol 11, No 43, 76-105

118. Kahneman, D., Knetsch, J.L., Thaler, R.H., (1991). *Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias,* Journal of Economic Perspectives, 5 (1): 193-206.

119. Katzman, K. (2011). *Iran Sanctions (RS20871)*, Congressional Research Service, https://crsreports.congress.gov/product/pdf/RS/RS20871/301

120. Kingma, D., & Ba, J. (2014*). Adam: A Method for Stochastic Optimization*. ArXiv.

121. Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). *60 Years of portfolio optimization: Practical challenges and current trends*. European Journal of Operational Research, 234(2), 356-371.

122. Kouvelis, P., & Yu, G. (1997). *A Robust Discrete Optimization Framework*. Nonconvex Optimization and Its Applications, 26-73.

123. Kothari, C. R. (2008). *Research Methodology, Methods and Techniques* (2nd ed). New Age International

124. Koza, J. R., Bennett, F. H., Andre, D., & Keane, M. A. (1996). *Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming*. Artificial Intelligence in Design, 151-170.

125. Kutschka, M. (2016). *Robustness Concepts for Knapsack and Network Design Problems Under Data Uncertainty*. Operations Research Proceedings, 341-347.

126. Larsen, G. A., & Resnick, B. G. (1998). *Empirical Insights on Indexing*. The Journal of Portfolio Management, 25(1), 51-60.

127. Le Roux, N., Bengio, Y., & Fitzgibbon, A. (2012). *Improving First and Second-Order: Methods by Modelling Uncertainty*. Optimization for Machine Learning.

128. Leavens, D.H., (1945), *Diversification of Investments*, Trusts & Estates, 80, 469-473

129. Lekovic, M., (2021), *Historical Development of Portfolio Theory*, Tehnika, 76, 220-227

130. Liebchen, C., Lübbecke, M., Möhring, R., & Stiller, S. (2009). *The Concept of Recoverable Robustness, Linear Programming Recovery, and Railway Applications*. Robust and Online Large-Scale Optimization, 1-27.

131. Lintner, J, (1965), *The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,* Review of Economics and Statistics, February

132. Lo, A.W., MacKinlay, C., (1988), *Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test*, Review of Financial Studies, 1(1), 41-66

133. *Loan-to-Deposit Ratio of Iranian Banks & Credit Institutions, Grouped by Province.* (2018). Central Bank of Iran. https://cbi.ir/category/6068.aspx

134. Luss, H., & Rosenwein, M. B. (1997*). Operations Research applications: Opportunities and accomplishments*. European Journal of Operational Research, 97(2), 220-244.

135. Malkiel*, B.G.,* (1973), *A Random Walk Down Wall Street,* W. W. Norton and & Company

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

137. Markowitz, H.M., (1991), *Foundations of Portfolio Theory,* The Journal of Finance, Vol 46, No 2, 469-477

138. Markowitz, H.M., (1999), *The Early History of Portfolio Theory,* The Journal of Finance, Vol.55, No 4, 5-16

139. Meade, N., & Salkin, G. R. (1989). *Index Funds—Construction and Performance Measurement*. Journal of the Operational Research Society, 40(10), 871-879.

140. Meade, N., & Salkin, G. R. (1990). *Developing and Maintaining an Equity Index Fund*. Journal of the Operational Research Society, 41(7), 599-607.

141. Mejia-Giraldo, D., & McCalley, J. (2014). *Adjustable Decisions for Reducing the Price of Robustness of Capacity Expansion Planning*. IEEE Transactions on Power Systems, 29(4), 1573-1582.

142. Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of Machine Learning*. Cambridge, MA: MIT Press.

143. Michalewicz, Z. (1995). *A Survey of Constraint Handling Techniques in Evolutionary Computation Methods,* In: McDonnel et al eds, Evolutionary Programming IV, MIT Press, 135-138

144. Moon, Y., & Yao, T. (2011). *A robust mean absolute deviation model for portfolio optimization*. Computers & Operations Research, 38(9), 1251-1258.

145. *Mutual Fund Comparison.* (2022 June). Financial Information Processing of Iran (FIPiran), Fund Comparison, , http://fund.fipiran.ir/mf/list

146. NadAli, M., Rasooli, M., SoltanolOlamayi, M., HajiDoolabi, H., Nasiri, S., (2017). I*nteraction of Capital and Money Markets in the Iranian Economy*, Ravand Journal, 77, 65-114

147. Najafi, A.A, Mooshkhian, S., (2015). *Modelling an Optimal Solution to the Multi-Period Portfolio Optimisation Problem Utilising Genetic Algorithms.* Portfolio Management, Khajje Nasir-Uddin Toosi University, Vol 5, No 21, 13-31

148. Oh, K. J., Kim, T. Y., & Min, S. (2005). *Using genetic algorithm to support portfolio optimization for index fund management*. Expert Systems with Applications, 28(2), 371-379.

149. Pearl, J. (1984). *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Boston, MA: Addison-Wesley Longman.

150. Qian, N. (1999). *On the momentum term in gradient descent learning algorithms*. Neural Networks, 12(1), 145-151.

151. Quaranta, A. G., & Zaffaroni, A. (2008). *Robust optimization of conditional value at risk and portfolio selection*. Journal of Banking & Finance, 32(10), 2046-2056.

152. Quisenberry, C. H. (2006). *Core/Satellite Strategies for the High-Net-Worth Investor*. CFA Institute Conference Proceedings Quarterly, 23(4), 38-45.

153. Radivojević, N., Lazić J., Cvijanovic, J.M., (2009), *Limitations of Application of Modern Portfolio Theory in Emerging Markets*, Industrija, Vol. 37, No. 4, pp. 1-13, 2009

154. Rafaely, B., & Bennell, J. (2006). *Optimisation of FTSE 100 tracker funds*. Managerial Finance, 32(6), 477-492.

155. Razeghi, M. (2012). *Using Neural Networks for the Development of Index Tracking Funds (Unpublished master's thesis)*. Science and Culture University, Tehran, Iran.

156. Rohweder, H. C. (1998). *Implementing Stock Selection Ideas*. The Journal of Portfolio Management, 24(3), 49-59.

157. Roll, R. (1992*). A Mean/Variance Analysis of Tracking Error*. The Journal of Portfolio Management, 18(4), 13-22.

158. Rom, B.M., Ferguson, K.W., (1993), *Post-modern Portfolio Theory Comes of Age*. The Journal of Investing, Vol. 2, No. 4, 27-33, 1993.

159. *Rouhani's Three Episodes in the Stock Market: An Invitation That Caused Millions to Lose Their Savings.* (2021, August). Dana.ir. https://www.dana.ir/news/1750158.html

160. Roy, A.D., (1952), *Safety First and the Holding of Assets, Econometrica, Vol 20, No 3, 431-449*

161. Rubenstein, M., (2002), *Markowitz's "Portfolio Selection": A Fifty-Year Retrospective*, The Journal of Finance, Vol 57, No 3, 1041-1045

162. Rudd, A. (1980). *Optimal Selection of Passive Portfolios*. Financial Management, 9(1), 57.

163. Rudd, A., Clasing, H. K. (1988). *Modern Portfolio Theory: The Principles of Investment Management*. Probus Publishing Company.

164. Rudolf, M., Wolter, H., & Zimmermann, H. (1999). *A linear model for tracking error minimization*. Journal of Banking & Finance, 23(1), 85-103.

165. Ruiz-Torrubiano, R., & Suárez, A. (2008*). A hybrid optimization approach to index tracking*. Annals of Operations Research, 166(1), 57-71.

166. Samuelson, P. A. (1974). *Challenge to judgment*. The Journal of Portfolio Management, 1(1), 17-19.

167. Samuelson, P.A., (1965), *Rational Theory of Warrant Pricing*, Industrial Management Review Vol 6, No 2, 13-39

168. Scherer, B. (2010). *Portfolio Construction and Risk Budgeting*. Risk Books

169. Scherer, B., (2002). *Portfolio Resampling: Review and Critique.* Financial Analysts Journal, Vol 58, No 6, 98-109

170. Schöbel, A. (2014). *Generalized light robustness and the trade-off between robustness and nominal quality*. Mathematical Methods of Operations Research, 80(2), 161-191.

171. Schoenfeld, S. A. (2004). *Active Index Investing: Maximizing Portfolio Performance and Minimizing Risk Through Global Index Strategies*. Wiley Finance Editions

172. *Securities Market Act.* 7th Majlis. (2005). https://rc.majlis.ir/fa/law/show/97786

173. Sharpe, W. F. (1991). *The Arithmetic of Active Management*. Financial Analysts Journal, 47(1), 7-9.

174. Sharpe, W. F., Alexander, G. J., & Bailey, J. V. (1995). *Investments* (6th ed.).

175. Sharpe, W.F., (1964), *Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk,* The Journal of Finance, Vol 19, 425-442

176. Sharpe, W.F., (1965), *Mutual Fund Performance,* The Journal of Business, Vol 39, 119

177. Shiller, R. (1980*). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?* The American Economic Review, Vol. 71, No. 3, 421-436

178. Shooshtar, Z., Namazi, M., (1995), *Absence of Weak Form Efficiency: Proof*, Journal of Financial Research, Management School of University of Tehran, Vol 2, No 7

179. Singleton, J. C. (2004). *Core-Satellite Portfolio Management*. New York, NY: McGraw Hill Professional.

180. Sirignano, J., Cont, R., 2018). *Universal Features of Price Formation in Financial Markets: Perspectives from Deep Learning,* arXiv Prints, Available at: https://arxiv.org/abs/1803.06917?context=q-fin

181. Soyster, A. L. (1973). *Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming*. Operations Research, 21(5), 1154-1157.

182. Strong, R. A. (2008). *Portfolio Construction, Management, and Protection*. South Western Educational Publishing.

183. Surz, R. (2012). *Correcting Core*. The Journal of Index Investing, 2(4), 67-73.

184. Tabata, Y., & Takeda, E. (1995). *Bicriteria Optimization Problem of Designing an Index Fund*. Journal of the Operational Research Society, 46(8), 1023-1032.

185. Tehrani, R., Fallah Tafti, S., Asefi, S., (2018). *Portfolio Optimization Using Krill Herd Metaheuristic Algorithm Considering Different Measures of Risk in Tehran Stock Exchange*. Journal of Financial Research, Tehran University Press, Vol 20, No. 4, 409-426

186. Till, H. (2004). *Risk Measurement of Investments in the Satellite Ring of a Core-Satellite Portfolio: Traditional Vs Alternative Approaches*. The Singapore Economic Review, 49(01), 105-130.

187. Tobin, J. (1989). *MONEY, CAPITAL, AND OTHER STORES OF VALUE*. General Equilibrium Models of Monetary Economies, 25-37.

188. Tobin, J., (1958), *Liquidity Preference as Behaviour Towards Risk,* Review of Economic Studies, Vol 24, No 1, 65-86

189. TURING, A. M. (1950). *Computing Machinery and Intelligence*. Mind, LIX(236), 433-460.

190. Tütüncü, R., & Koenig, M. (2004). *Robust Asset Allocation*. Annals of Operations Research, 132(1-4), 157-187.

191. Vyas H. N, (2014), *Portfolio Management*, Amazon International

192. Walsh, A. (2013*). Core–Satellite Strategies: Combining Stability and Opportunity in an ETF Portfolio*. The Journal of Index Investing, 4(3), 50-53.

193. Warner, M. (2002). *International Encyclopedia of Business and Management*. Routledge.

194. Watkins, C. (1989). *Learning from Delayed Rewards* (Doctoral dissertation, Kings College, London, UK). Retrieved from https://www.researchgate.net/publication/33784417_Learning_From_Delayed_Rewards

195. Williams, J.B*.,* (1938), *The Theory of Investment Value*, Harvard University Press

196. Xidonas, P., Mavrotas, G., Hassapis, C., & Zopounidis, C. (2017). *Robust Multi-Objective portfolio optimization: A minimax regret approach*. European Journal of Operational Research, 262(1), 299-305.

197. Yu, L., Zhang, S., & Zhou, X. Y. (2006*). A Downside Risk Analysis based on Financial Index Tracking Models*. Stochastic Finance, 213-236.

198. Zeiller, M. (2012). *ADADELTA: An Adaptive Learning Rate Method*. ArXiv. Retrieved From https://arxiv.org/abs/1212.5701

199. Zymler, S., Rustem, B., & Kuhn, D. (2011). *Robust portfolio optimization with derivative insurance guarantees*. European Journal of Operational Research, 210(2), 410-424.