An Objective Analysis of Alternative Risk-to-Reward Ratios.

A thesis submitted for the degree of

Doctor of Philosophy

by

Josephine Gerken

18th June 2015

Department of Mathematics

University College London

Declaration

I, Josephine Gerken, confirm the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis is devoted to the task of investigating the merits of employing a generalised reward-to-risk ratio for the dual purpose of ranking assets and forming optimal portfolios. In quantitative portfolio optimisation and risk management, the optimal allocation of assets often relies on traditional and well-established measures including the Sharpe ratio. Despite their frequently-documented drawbacks, these quantitative tools continue to be relied on. This is partly due to familiarity and ease of use. Another reason stems from the empirical evidence that alternative and more complex measures often produce results highly correlated with those based on their standard counterparts. This thesis presents an objective analysis of various risk-to-reward measures, for the purpose of portfolio optimisation, which are flexible enough to represent asymmetry in risk and return preferences. In particular, we regard the one-sided variability ratio (Farinelli and Tibiletti (2002)), $\phi(b, p, q)$, as an intuitive tool to be used in an optimal allocation model on account of its flexibility and ability to account for any distributional model. Focusing on two parameterisations of $\phi(b, p, q)$ (Omega $= \phi(b, 1, 1)$ and Sortino ratio $= \phi(b, 1, 2)$), we show how these ratios can be analytically reduced to functions of the Sharpe ratio based on Student t returns. This sheds unfavourable light on such alternative and supposedly

superior measures and verifies what practitioners commonly observe: that several theoretically advocated ratios replicate the same, or near identical, outcomes as the Sharpe ratio. Analogously, Eling and Schuhmacher (2012) provide further support for this hypothesis, conditional on the location-scale property being satisfied. Relaxing this distributional assumption, however, we cannot claim that ratios including Omega and Sortino do not produce substantially different results to the Sharpe ratio. In an empirical setting, it is questionable to suppose that return distributions satisfy the location-scale property. We consider anisotropic bivariate distributions proposed by Shaw and Lee (2008) to facilitate the simulation of situations where alternative measures produce different results in terms of asset rankings and dissimilar optimal asset weights. In light of this, we argue that it is critical to take an objective approach to constructing a portfolio optimisation model from modelling the underlying data to selecting the relevant risk and performance measures.

Acknowledgements

I would like to thank my PhD supervisor, Professor William Shaw, for supporting me during these past four years and for encouraging me to carry out my PhD studies in the first place. I am grateful to him for passing on his knowledge, expertise and advice from both an academic and practical perspective. Apart from his academic guidance, I would also like to thank Professor Shaw for sourcing my DTA with EPSRC and financial support. I would also like to express my gratitude to other members of the mathematics department at UCL, including Dr. Andrea Macrina, Robb McDonald and Helen Higgins. I will forever be thankful to my sponsors, Zephyrus Partners, who without their financial support I would not have been able to complete this PhD. In particular, I would like to thank Christophe Louis and Julien Le Tyrant whom I worked with during my time there.

Finally, but most importantly, I am eternally grateful to my parents who have always been there for me throughout my academic career. Their support has been invaluable to me and I hope I have made them proud. I would like to dedicate this thesis to them.

For Mum and Dad. x

Contents

Chapter 1

Introduction

The Sharpe ratio is often the first performance measure that comes to mind in the context of measuring the relative risk and performance potential of a collection of assets. In short, it is the expected return in excess of the risk-free rate divided by the standard deviation of returns. An extensive list of alternative performance measures exists, however, the majority of which were devised for the purpose of overcoming the drawbacks associated with the Sharpe ratio. We focus on a particular family of performance measures, the one-sided variability ratio (Farinelli & Tibiletti, 2002), to be explained in due course, due to its ability to account for any distributional model and capture a broad spectrum of investor preferences. We believe this measure is an intuitive tool for the dual purpose of ranking assets and constructing optimal portfolios.

The task of setting up a framework for portfolio optimisation requires making a number of decisions relating to the following questions: Which utility function is most appropriate for our objective?; Which distribution or family of distributions shall we use to model the marginal densities? On the same train of thought, which dependency structure should we employ to characterise the relationship between the assets (returns) in our investment universe? Furthermore, what are the constraints, if any, that must be satisfied?[1] All these questions demand well-chosen answers in the sense that we must strive for accuracy in modelling the 'truth' yet, for practical purposes, be able to produce results in a reasonable and sensible amount of time.

The optimal allocation problem in portfolio optimisation involves combining a number of appropriate optimisation and risk tools. In a risk-return analysis, we generally have an objective function (a function of the returns) that we wish to maximise whilst simultaneously minimising a measure of risk; this relates to the first question above. Commonly used risk and performance functions include the Sharpe ratio, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR) and variance[2]. There are, however, a large number of other measures, some which are more relevant for certain objectives and more effective in turbulent market conditions.

[1]Such constraints can include control of turnover on rebalancing the portfolio, short-selling restrictions, deviation from an index, sector exposure as well as more exotic requirements.

[2]VaR and CVaR are measures of loss and have widespread use in the finance industry. In particular, VaR is industry-standard for calculating regulatory capital. In short, for a given probability level, α, and time horizon, VaR(α) is the portfolio loss that, with probability α should not be exceeded (Rockafellar & Uryasev, 2002). CVaR is then the conditional expectation of losses beyond VaR(α).

Traditionally, the risk attached to a certain allocation is often measured by the standard deviation of the expected returns of the portfolio. This is often associated with the seminal work of Markowitz (1952; 1956; 1959). However, as we shall discuss later on in this thesis, standard deviation was employed as a measure of risk due, predominantly, to its familiarity and convenience. In his earlier paper, Markowitz (1952) proposes a methodology for selecting a diversified portfolio based on what is, and equally important, what isn't desirable. Here is where the 'expected returns - variance of returns' rule was first introduced into the finance literature, the use of which serves 'both as a hypothesis to explain well-established investment behaviour and as a maxim to guide ones own action' (Markowitz, 1952).

Markowitz is often erroneously associated with advocating variance as a measure of risk and the assumption of a quadratic utility function.[3] In a more recent publication, Markowitz (2010) reiterates the statement he made in his 1959 paper where he asserted that if a utility function can be defined as a quadratic for a 'sufficiently wide range of returns', then the expected value of that utility function can be characterised in terms of the expected return and variance. Essentially, what this is saying is that, provided a quadratic utility function suffices for a broad range of returns, then the first two moments (mean and standard deviation) are sufficient for describing the data. There was never the implication that standard deviation is the optimal measure of risk. Rather, we should

[3] As a result of assuming returns should be modelled with the Gaussian distribution.

aim to utilise the most appropriate method which is computationally practical in favour of a precise method which is computationally impossible or unfeasible (Markowitz, 1991). It must be noted that various alternative measures of risk were considered in Markowitz' 1959 paper for the purpose of carrying out a risk-return analysis for selecting a 'good' portfolio. We consider this further in Chapter 2.

Our view is that one should avoid making assumptions concerning the return distribution and which risk measure or portfolio performance measure should be employed. Central to this thesis is the view that, imposing a fixed model for portfolio optimisation from the beginning without fully characterising the objective, inputs and outputs could possibly yield results which hide imperfections due to unmanaged risk or foregone upside potential. This view is common with that of Markowitz (1959) in that an optimal portfolio should represent a 'balanced whole' in the sense that it should represent the needs of the investor, both in terms of 'potential and protection'. There exists a real danger of investors becoming complacent with existing asset allocation methodologies: not knowing what their objectives are and how to define their targets.

As a single investor, we may have a certain belief regarding a particular set of stocks or sectors and wish to construct a portfolio with significant exposure to our preferred assets. Hedge funds typically have the capital and technology to take on more risk and invest in financially complex instruments. Consequently, the distribution of their portfolio return will be significantly different from that of, say, a government portfolio.

For a pension fund, the time horizon over which they typically invest is on the scale of decades rather than a few years or even months. The objective of a pension fund will therefore be very different to that of a hedge fund, the main difference arising from the fact that hedge funds are much less regulated than pension funds who are required to abide by constraints set by an external actuary. The management of defined benefit pension funds, for example, requires satisfying solvency constraints as a primary objective based on, amongst other things, credit quality and the debt level of the underlying company.

Several steps must be taken when solving a generic portfolio optimisation problem:

1. Establishing the random variables: which of our inputs do we need to model?

2. Defining model parameters and estimating of random variables: Unless our approach is entirely empirical, an important stage of the portfolio optimisation process requires estimating the characteristics of our random variables. Parametrically speaking, this means first selecting one distribution or a family of distributions and subsequently estimating the relevant parameters to best fit the data. It is important to mention at this point the difficulty one can face when estimating the parameters of such a problem. This is a challenge in itself.Where necessary, we use a maximum likelihood method in Mathematica.[4]

3. Defining the optimal outcome: one must consider in what sense a portfolio is optimal

[4] In Mathematica, we use the function *EstimatedDistribution[]* which employs maximum likelihood estimation by default.

or not. For this reason, one needs to mathematically characterise the objective, identify what it is we want to maximise and what exactly we want to minimise or preferably eliminate.

4. Defining our risk measure: defining in what sense a portfolio is optimal or not will dictate which risk measure is relevant. The chosen risk measure can either be incorporated within the objective function or applied to the portfolio optimisation problem as a constraint.

5. Modelling and computing the optimal portfolio: the objective function illustrates how we map the random variables, namely returns, into a measure that characterises our optimal outcome. It is then a question of computing asset weights such that our objective function is maximised and our constraints are satisfied.

The subject of this thesis concerns points 3. and 4. above. A thorough mathematical study of all topics above can be found in Meucci (2009). In the following chapter, we discuss two main approaches to mathematically formulating the portfolio objective function. Of course, one cannot review these approaches without mentioning Markowitz who can be credited for bringing portfolio theory into the domain of mathematics. His work at the time fits into the broader category of what is known as modern portfolio theory. In a nutshell, this approach to constructing an optimal portfolio is founded on a risk-return analysis, where risk was traditionally characterised by the variance of returns. Related to this, Sharpe (1966) postulated constructing the optimal portfolio by maximising what has come to be called the Sharpe ratio; a risk-to-reward ratio given by the ratio of expected return, in excess of the risk-free rate, to the standard deviation of returns. For asset i, when expected return and standard deviation are defined as E_i and σ_i, respectively, and p denotes the risk-free rate, the Sharpe ratio was defined in Sharpe (1966) as follows:

$$\frac{E_i - p}{\sigma_i}. \tag{1.1}$$

Several drawbacks associated with modern portfolio theory are well-documented and will be discussed further in Chapter 2. By extending the general framework underlying modern portfolio theory to include alternative risk measures, a new branch of portfolio theory emerged: post modern portfolio theory. This extension to modern portfolio theory focused on analysing returns about a specified target. Motivated by the work of academics including Sortino and Price (1994), Stone (1973) and Fishburn (1977), alternative methods of measuring risk were proposed.

Chapter 3 summarises the key components of a parametric portfolio optimisation model. Here, we consider distribution choices, portfolio performance measures and, briefly, dependency models. On the former of these, we focus our attention on appropriate substitutes for the normal distribution. We review some of the well-known papers analysing financial returns and discuss other distributions capable of capturing the observed properties of returns beyond mean and variance, of which only the Gaussian distribution can be fully described. In particular, we highlight papers advocating the Student t distribution as a model yielding the highest likelihood in terms of fitting the empirical data. In the greater part of Chapter 3, we discuss several performance measures and how these are defined in terms of incorporating risk and return. Here, we explore various measures of risk stemming from the work of Stone (1973). This section forms the basis and motivation for the rest of this thesis.

Numerous researchers have focused their efforts on studying portfolio performance measures. For example, Cogneau and Hubner (2009) review 101 performance measures and categorise them on the basis of how the objective is defined and, consequently, how assets are selected, how flexible they are in terms of being able to represent the general investor and how returns are built in; whether it is absolute vs. relative return which matters or excess return vs. gain measure. Farinelli and Tibiletti and co-authors propose a more flexible formulation for performance measurement than the Sharpe ratio to allow for one-sided performance and risk evaluations (see Farinelli & Tibiletti, 2002; Farinelli et al, 2006; Farinelli et al, 2008). Wiesinger (2010) reviews performance tools in terms of how returns are adjusted for risk. On the subject of risk itself, as far back as 1973, Stone recognised the limited use of standard deviation as a proxy for risk and proposed a broad family of risk measures of which Fishburn's two-parameter family is a subset (Fisburn 1977). We discuss these later on in this thesis. Building on the work of Stone (1973), Pederson and Satchell (1998) offer an even more general class of risk measures in addition to proposing a decision criteria for which of these risk measures are appropriate. More specifically, whether the choice of performance measure really matters is investigated by Eling and Schuhmacher (2005), and Huyen (2007) for the purpose of analysing hedge funds. Zakamouline (2011) asserts that, for the evaluation of hedge funds, the choice of performance measure does matter.

The phenomenon of risk is pertinent in multiple disciplines and it is interesting to read that risk measures that have been used for applications ranging from psychology to management science have also been proposed for their potential to measure risk in financial returns (Pederson & Satchell, 1998). We dedicate a large part of Chapter 3 to the one-sided variability ratio first introduced by Farinelli and Tibiletti (2002). This ratio is known as the Farinelli-Tibiletti ratio in Eling and Schuhmacher (2012). It is the ratio of upper partial moment of order p to the lower partial moment of order q where deviations are quantified about a benchmark, b.

$$\phi(b,p,q) = \frac{\mathbb{E}[|(X-b)^{+}|^{p}]^{1/p}}{\mathbb{E}[|(X-b)^{-}|^{q}]^{1/q}}. \tag{1.2}$$

In (1.2) above, $()^{+}$ and $()^{-}$ refer to maximum and minimum functions, respectively. More specifically, $()^{+}$ is equivalent to $\max(x,0)$. Similarly, $(x)^{-}$ is equivalent to $\min(x,0)^{+}$. The one-sided variability ratio can be tailored to the needs of any investor on account of its three parameters b, p and q. This measure, $\phi(b,p,q)$, possesses many desirable characteristics. Firstly, it is general in three senses: how we define optimal return, how we define risk and about which value or variable we measure return and risk.

Secondly, it can be used for the application in both an empirical and parametric analysis. Furthermore, in a parametric framework, it is capable of capturing all moments of the distribution of the underlying random variables. It is for these reasons why we advocate its use for applications in portfolio optimisation. Further explanation regarding equation (1.2) is provided in Chapter 3. In addition, we discuss the motivations that lead to the creation of a general performance measure.

In Chapter 4, we focus on two parametrisations of (1.2): the Omega Index, $\phi(b, 1, 1)$ and Sortino ratio, $\phi(b, 1, 2)$. We show how to reduce the problem of computing the Omega Index and Sortino Ratio to the evaluation of analytical functions of the Sharpe ratio when the underlying random variable is modelled by a Student t distribution. Based on several statistical studies over many decades, we propose to use the Student t model as a suitable representation of asset returns. We take the view that, since this model is at least less worse than the Gaussian model and exhibits the empirically observed fat tails, we have grounds for employing the Student t distribution for the purposes of risk management or portfolio optimisation, (Breymann, Luthi & Platen, 2009). Prior to selecting values for parameters p and q in $\phi(b, p, q)$, we show how to simplify the performance measure, in its general form, when the underlying random variable is a linear transformation of a random variable with zero mean and unit variance. We then present the reductions of Omega and Sortino in a Gaussian model for comparative purposes. Following this, we give a brief overview of the Student t distribution to facilitate calculations made in the rest of the thesis.

It was our belief that a wider family of risk-to-reward performance measures could be reduced to a function of the Sharpe ratio and that, under certain conditions, these functions may possibly be monotonically increasing functions of the Sharpe ratio. In carrying out research to support this belief, we encountered the work of Eling and Schuhmacher (2007; 2012), who proved that a large class of risk-to-reward ratios are monotonically increasing functions of the Sharpe ratio provided the location-scale property is satisfied and certain properties exist regarding the performance measures. In addition, each reward and risk component must satisfy certain conditions on their functional form. In brief, two random variables satisfy the location-scale property if their de-meaned and de-scaled equivalents are equal in distribution. We review the paper of Eling and Schuhmacher (2012), as their result generalises our analysis provided in Chapter 4, and consider the implications of their findings in the context of asset allocation. The conclusions reached in Eling and Schuhmacher (2012), however, are due to restrictive statistical properties which we claim should not be imposed on general grounds. On the one hand, the results in Eling and Schuhmacher (2012) may lead the reader to believe that these alternative measures are, in fact, no more superior than the Sharpe ratio. Consequently, due to its simplicity and familiarity, one may rely solely on the Sharpe ratio as the most adequate risk-to-reward ratio. We claim that consistency between the Sharpe ratio and other risk-to-reward ratios does not exist in general. To maintain the flow of the thesis, we take the Omega Index and Sortino ratio as special cases of the one-sided variability ratio, $\phi(b, p, q)$, and test whether asset rankings and optimal portfolios based on these measures are consistent with the Sharpe-optimal equivalent portfolios when we first impose the location-scale property and then relax this condition. Our approach in Chapter 6 is to consider anisotropic bivariate distributions, as formulated by Shaw and Lee (2008), and asymmetric dependency structures, two concepts not considered in Eling and Schuhmacher (2012). Our results show that one doesn't need to suppose highly exotic distributional models for asset returns to lead to inconsistencies between performance ratios. The analysis in Chapter 6 thus further emphasises the need to expand our toolbox of risk-to-reward ratios beyond the traditional measures such as the Sharpe ratio.

Everything considered, this thesis studies the merits of a flexible class of performance measures, namely the one-sided variability ratio in Farinelli and Tibiletti (2002), not just in terms of how they value individual assets but also on how they combine these assets via the process of optimising a portfolio. Whilst several authors investigate the strengths and weaknesses of various performance measures and explore whether they provide any advantages relative to more traditional measures such as the Sharpe ratio, the focus is usually aimed at questioning whether they differ in terms of individual asset rankings. Exploring the benefits of various parameterisations of $\phi(b, p, q)$ from both an individual asset and portfolio perspective is, to the best of my knowledge, where this thesis constributes to the existing literature. Our conclusions are of importance for the purpose of understanding why it is important to consider a broad spectrum of performance measures in a general portfolio optimisation setting not just theoretically but also for practical implementation.

1.1 Data description and optimisation process

In order to investigate the various parameterisations of the one-sided variability ratio in terms of how they compare with the Sharpe ratio, we calculate optimal portfolios by maximising four parameterisations of (1.2): $\phi(b,1,1)$ (Omega Index), $\phi(b,1,2)$ (Sortino ratio), $\phi(b,1,3)$ and $\phi(b,1,4)$ where $b \in \{5\%, 10\%, 15\%\}$. The resulting optimal portfolios are compared with the Sharpe-optimal equivalent portfolios. Individual asset rankings are also calculated and compared. Four sets of optimal portfolios are calculated for this set of performance measures. Two sets use empirical return data from two distinct time periods (to be explained in due course). The same optimisations are then carried out using Student t simulated return data where the first two moments of each asset are taken from the empirical asset returns and we assume four degrees of freedom, homogeneously.[5] This approach essentially assumes returns follow a particular continuous statistical distribution to which risk-adjusted performance measures and portfolios optimal in terms of these performance measures are calculated, accordingly. The point here is to simulate

[5]The choice of four degrees of freedom is trivial but a natural choice given the findings in Platen and Sidorowicz (2007) and Breymann, Luthi and Platen (2009).

return data such that the marginal densities satisfy the location-scale property in accordance with the framework underlying the work in Eling and Schuhmacher (2012). What we are interested in is whether the theoretical results presented in Eling and Schuhmacher (2012) agree with our results. As far as this author is aware, there is no obvious relationship between asset rankings and optimal asset weights for any risk-adjusted performance measure. We will find that agreement in asset rankings between a subset of risk-adjusted performance measures does not imply agreement in optimal asset weights based on the same set of measures. It is this distinction we wish to make clear and is the motivation for carrying out these tests.

Initially, we carry out the above analysis on a universe of UK equities. To add another dimension, we also carry out the same analysis where our universe of assets consists of three major currencies and three commodities. UK equities tend to be positively correlated whereas non-equity asset returns do not necessarily tend to move in the same direction. The latter optimisation analysis will therefore provide further clarification of whether the broad range of performance measures in (1.2) offers anything different to the Sharpe ratio under alternative scenarios.

The first dataset to be used consists of daily returns of eighteen stocks selected from the FTSE 100 Index traded on the London Stock Exchange as of September 2014. The dataset covers the period January 2003 to March 2012. The eighteen stocks are chosen due to their return distributions exhibiting various shapes. That is, the stocks chosen offer a broad range of statistical characteristics. The second dataset consists of six asset return series: three currencies (EUR, GBP and JPY) and three commodities (Gold, Oil and Silver) denominated in USD.

The two periods under investigation are January 2003 to July 2007 and August 2007 to March 2012. The former represents the years considered to be, by and large, economically thriving. The latter covers the recent financial crisis. The reason for choosing these time periods is to show that the various parameterisations of the one-sided variability ratio behave as we would expect them to over diverse economic climates.

The results presented in this thesis are all produced using the package, *StochOptia*, in Mathematica. This package is centered on a Monte Carlo framework which carries out optimisations by simulating both random returns and random portfolios. For the benefit of the reader, we provide an overview of the optimisation method carried out when running the StochOptia package.

In general, we wish to use return data of a set of assets to calculate those asset weights which yield the optimal value of a pre-defined objective function, subject to possible constraints. We would like to have the option to use empirical or simulated data where asset returns are drawn from standard distributions. The method proposed is Monte Carlo in nature. We carry out a computational search amongst possible portfolios based on any multivariate return model (including the empirical dataset), inter-asset dependency structure, objective function as well as any investor-specific constraints. Optimum portfolio weights are chosen such that no other set of simulated portfolio weights produce a higher objective value. The extensive computational search involves simulating random portfolios in addition to simulating portfolio return samples. Calculation of the objective function is carried out on each set of feasible portfolio weights. The search then chooses the set of weights yielding the optimal value of the objective function.

Simulation of random portfolios in an un-biased manner is achieved using a 'Face-Edge-Vertex Biased' scheme (Shaw, 2011). This method of generating random portfolios is chosen due to its ability to capture solutions along the faces, edges and vertices as well as the interior of a feasible hyperspace.[6] Essentially, weights are sampled as a function of uncorrelated Uniform random variables on $(0, 1)$ and rescaled to sum to unity:

$$w_i = \frac{f(U_i)}{\sum_i f(U_i)}. \tag{1.3}$$

The finer details of the theory underpinning this method are beyond the scope of this thesis. For the interested reader, however, the author would recommend referring to Shaw (2011) where the sampling methods of Devroye (1986) are summarised. For now, let us consider an investment universe comprising of N assets. We generate M sets of samples of N uncorrelated uniform random variables, $U^p_{n,m}$, taken to the 2^p-th power. For

[6] This compares with even simplicial sampling whose solutions tend to cluster around the equally-weighted point rather than producing possible solutions on or towards the edges of the feasible hyperspace.

some integer p, a full set of simulated portfolios is then generated as follows:

$$w_{n,m} = \frac{U_{n,m}^{2p}}{\sum_{i=1}^{N} U_{i,m}^{2p}}. \tag{1.4}$$

To illustrate the various levels of face-edge-vertex bias that can be achieved with this method of random portfolio sampling, Figure (1.1) exhibits the 3-asset case where p takes the values $p = \{0, 1, 2, 3, 4, 5, 6\}$ and the number of samples is 3000 (that is, $m = 3000$ in (1.4)). One can see that as we increase the value of p, more weights on the boundaries of the hypercube will be selected. When incorporating constraints, both linear and non-linear constraints are dealt with simply by rejecting those set of simulated portfolio weights which do not satisfy the criteria. A visualisation of how the feasible space is shrunk for two different sets of constraints is illustrated in Figure (1.2). In both cases $p = 3$ and $m = 10000$. In the plot on the left-hand side of Figure (1.2), optimal asset weights are bound below at -0.1. Feasible portfolios are constrained more tightly in the plot on the right-hand side where we impose a lower bound of -0.5 on all assets and constrain the total weight of short positions to lie in the range $(-0.5, -0.2)$.

In concluding this introduction, the objective of this thesis is to advocate employing a wide range of performance measures for portfolio optimisation and quantitative risk management. Given the vast amount of literature disregarding the Gaussian distribution as an appropriate model of financial returns, it is clear that alternative measures beyond those that rely solely on mean and variance should be considered. In the greater part

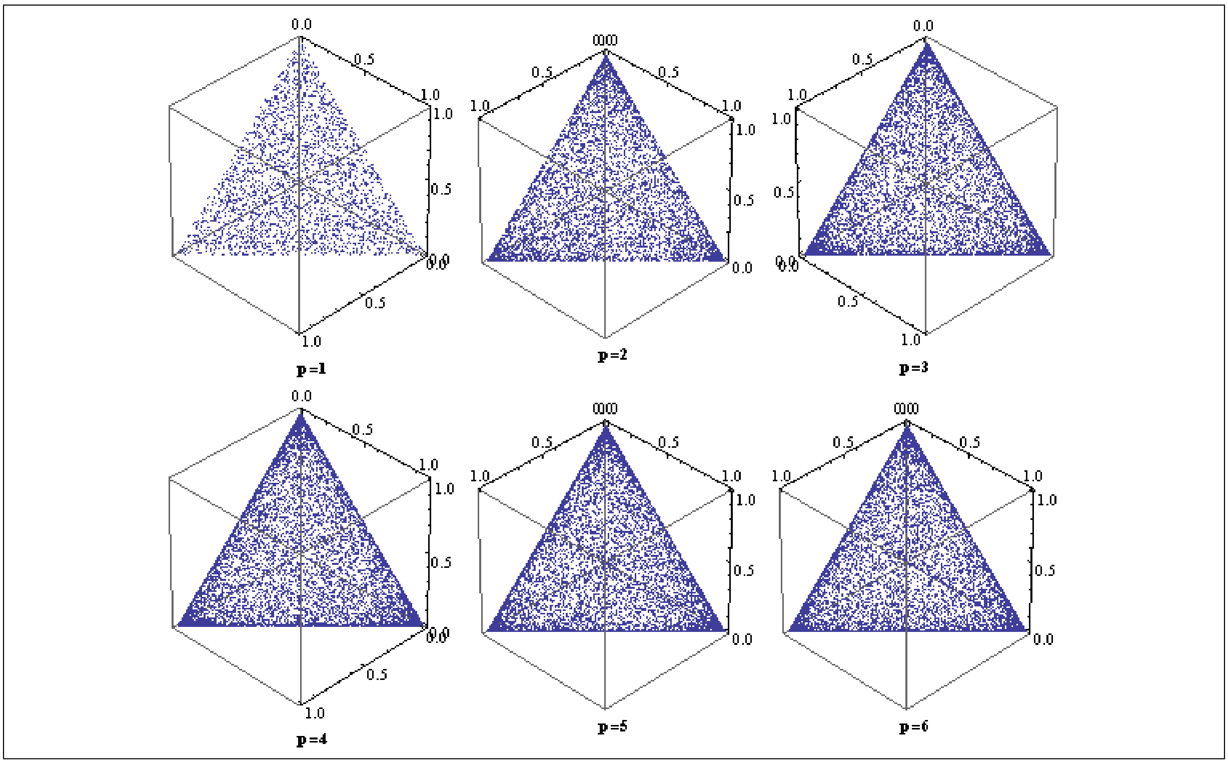

Figure 1.1: *A visual representation of various levels of random sampling that can be achieved using Face-Edge-Vertex Biased for sampling three asset weights.*

of the thesis, we concentrate on two parameterisations of the one-sided variability ratio: The Omega Index and Sortino ratio. In light of the recent financial crisis, the emphasis on risk management has never been higher. We believe that limiting the set of risk management tools to one, or even a few similarly defined risk tools can be risky in itself. In the context of asset allocation, combining the Omega Index and Sortino ratio to the box of tools for optimisation purposes may lead to different optimal portfolios that will force the portfolio manager to question exactly what the investor's objectives are. By no means should the reader interpret this to mean we ought to disregard traditional

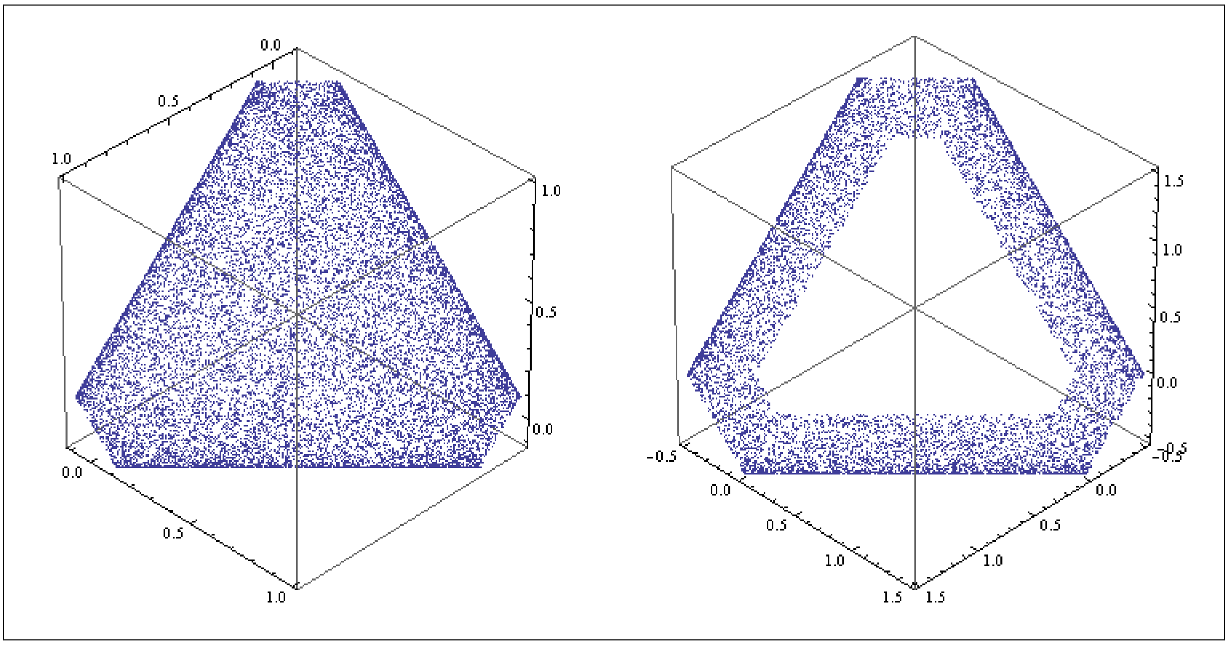

Figure 1.2: *A visualisation of how a feasible portfolio space for three assets is reduced under certain constraints. The figure on the left displays how imposing a lower bound of* -0.1 *chops the feasible portfolio space. By setting a lower bound of* -0.5 *on all assets and constraining the total weight of shore positions to lie in the range* $(-0.5, -0.2)$ *we illustrate the level to which our feasible portfolio space is reduced.*

measures including the Sharpe ratio. In fact, what we wish to emphasise is that certain measures serve their purpose better in different scenarios. What is necessary is to select the appropriate measures based on the nature of the inputs and the objective of the investor. A timely article on this matter by Hoffman and Rollinger (2013) provides evidence that alternative risk-adjusted performance ratios such as the Sortino ratio are being considered in industry.

Whilst the mathematical content in what follows predominantly focuses on two particular performance ratios, broadly speaking, the proceeding analysis and general philosophy of this paper extends to other ratios and components of a portfolio optimisation model.

Chapter 2

Portfolio optimisation

'... uncertainty has to do with the range of returns that could occur, and risk has to do with some of those returns.'

Sortino

The task of allocating an amount of wealth amongst a number of assets is an enduring challenge. In its simplest form, this problem consists of making a choice as to what percentage of our wealth we should allocate to n assets. From a more theoretical perspective, however, we generally want to share our wealth amongst those n assets such that we optimise a well-chosen objective function. Additionally, we may have to reduce our set of feasible allocations in order to satisfy particular constraints or targets. Finally, it is often required that our optimal allocation represents a degree of diversification. Having all your eggs in one basket, as the age-old proverb states, is generally not a safe bet.[7] Note that a portfolio can be diversified in a number of ways. Traditionally, the most diversified portfolio is the equally-weighted or '$1/n$' portfolio.[8] Distributing ones wealth equally amongst a basket of assets is a tried and tested method. However, alternative definitions of a well-diversified portfolio have come to fore. For example, an asset allocation methodology gaining in popularity is the equal risk contribution portfolio, (Maillard, Roncalli & Teiletche, 2010), which is diversified in terms of variance rather than market capitalisation. With this in mind, a general formulation for the portfolio optimisation

[7] Whilst this is a well-known saying and one often reciprocated in the context of portfolio optimisation, it is in fact an interpretation of the original Spanish proverb. Carrying out a thorough search of the original version, one will uncover many different translations where there is no mention of 'eggs' anywhere. Regardless of the translation, the proverb as it is know in modern times appears to make sense in the context of numerous situations. Having all your eggs in one basket, and then dropping that basket, leaves you with nothing. On the contrary, however, Mark Twain (1894) can be quoted as saying: "Put all your eggs in one basket and watch that basket."

[8] The concept of sharing one's wealth equally amongst all assets is certainly not a modern strategy. As quoted by DiMiguel, Garlappi and Uppal (2009) and Benartzi and Thaler (2001) (who both cite 'Babylonian Talmud: Tractate Baba Mezi'a, folio 42a' as the source), as far back as the fourth century, the Rabbi Isaac bar Aha offered the following advice: 'One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand.'

problem is as follows:

$$\text{Aim:} \qquad \max_w(\Psi(w))$$

$$\text{Subject to:} \qquad \text{constraints and targets,}$$

where $\Psi(.)$ represents our objective function.

In this chapter, we discuss the two main branches of portfolio optimisation; modern portfolio theory and post-modern portfolio theory. Firstly, with reference to modern portfolio theory, we introduce the pioneering work of Harry Markowitz whose contribution to portfolio analysis cannot be ignored. On account of the drawbacks associated with the work of Markowitz and modern portfolio theory in general, we discuss the succeeding area of research known as post-modern portfolio theory.

2.1 Modern portfolio theory

The pioneering work of Harry Markowitz during the 1950s (1952; 1956; 1959) initiated the mathematical theory of portfolio selection. These studies laid the groundwork for what is known as 'Modern Portfolio Theory' (Markowitz, 1952; 1959). Modern Portfolio Theory (henceforth MPT) is a method for portfolio selection which promotes diversification. This is achieved by seeking to reduce the variance of a portfolio through the methodology of combining assets which are not perfectly positively correlated. The MPT framework is based on a number of restrictive assumptions, notoriously on the return distribution and on the definition of risk. For instance, it postulates that all investors are risk averse and markets are perfect. More infamously, though, is the perception of normally distributed returns and the use of standard deviation as a proxy for risk.

In the preliminary paper on portfolio selection via mean-variance analysis (Markowitz, 1952), which is a particular application of MPT, risk was characterised by the variance. This is where the mistaken belief that Markowitz solely advocated the standard deviation to represent risk lies. If one refers to his later publication, the reader will verify that Markowitz had it in mind to use semi-variance as a superior and investor-tailored risk measure (Markowitz, 1959). Whereas variance is an average of the squared deviations of a data set either side of the mean or target value, semi-variance solely measures the squared deviations below the mean (or target value when we refer to target semi-deviation). This is defined as $\mathbb{E}[((x-b)^{-})^2]$ or, computationally, in the empirical (discrete) case:

$$\text{Semi-variance } = \tfrac{1}{n}\sum_{x<b}^{n}(b-x)^2. \tag{2.1}$$

For comparative purposes, Markowitz (1959) considered the standard deviation, semi-variance, expected value of loss, expected absolute deviation, probability of loss and maximum loss as potential measures of risk in a risk-return analysis. Efficient portfolios based on the latter five of these measures were compared with the corresponding efficient portfolio based on mean-variance analysis. Associating these measures with their utility function, Markowitz concluded that portfolios optimal in the sense of expected loss, expected absolute deviation, maximum loss and probability of loss should not be trusted on theoretical grounds as they may lead to undesirable portfolios.[9] Markowitz asserted that a portfolio efficient in terms of expected return and semi-variance is always on par with, if not superior than, the equivalent efficient portfolio resulting from mean-variance analysis. Despite this conclusion, the technology available at the time led Markowitz to settle for variance. The problem of calculating an optimal portfolio was therefore characterised mathematically in terms of a function of the covariance matrix. For n assets, optimal weights, $\underline{\omega} = \{\boldsymbol{\omega_1}, \ldots, \boldsymbol{\omega_n}\}$, could then be determined by minimising the following function given, the return vector, $\underline{R}$, covariance matrix, $\underline{C}$, and Lagrange multiplier, λ, representing the investor's level of risk aversion:

$$f(\omega_1, \ldots, \omega_n, \lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j C_{ij} - \lambda \sum_{i=1}^{n} R_i \omega_i. \tag{2.2}$$

In addition, constraining the weights to be positive and imposing a fully-invested strategy provided the following constraints:

$$\begin{aligned} \omega_i &\geq 0 \\ \sum_{i=1}^{N} \omega_i &= 1. \end{aligned} \tag{2.3}$$

There are several issues with this well-known model. Primarily, as we mention above, risk is characterised exclusively by the variance. The notion that the general investor endeavours to reduce variance should be rejected purely on intuitive grounds;

[9] See Markowitz (1959) for a thorough review on the approach to optimising a portfolio via utility theory.

variance in returns contributes to generating profits and is not necessarily an undesirable characteristic. In addition, defining risk in terms of variance alone can often be linked to an assumption that the distribution of returns is Gaussian. This implicit assumption is further illustrated by the dependency model between returns being represented by the covariance.[10]

Whilst the long-only and fully-invested constraints are standard in portfolio optimisation models, in a general framework there may be more complex constraints. As an example, limits on sector exposure and individual stock weights or allowing for shorting might need to be imposed.

Several studies, both theoretical and empirical, assert how the mean-variance approach to portfolio selection can be appropriate only for certain conditions. On the one hand, Samuelson (1970) defends the mean-variance model on the grounds that it remains useful in situations where riskiness is limited. In which case the quadratic solution is approximately true. When these conditions cannot be assumed, however, both Feldstein (1969) and Samuelson (1970) highlight the limitations of the mean-variance model both in terms of restrictions on distribution and quadratic utility approximations. Restricting the model in one direction (that is, exclusively relying on mean and variance in the analysis) leads to further restrictions elsewhere. Feldstein (1969) rightly states that a

[10] A common approach, at least traditionally, to asset allocation was to begin with a utility function whose expected value was to be maximised. This methodology yields equation (2.2) if an exponential utility model is employed and again Gaussian returns are assumed (Shaw, 2011).

preference ordering (when a portfolio contains more than one risky asset) in terms of μ and σ can be justified only if each asset is distributed in such a way that any linear combination of these assets can be represented by a distribution that depends only on two independent parameters.

From a more economical viewpoint, Simaan (1993) considers the opportunity cost of mean-variance strategies relative to the investors optimal strategy on the basis of maximising his utility function. The opportunity cost is measured by the dollar amount the investor would require in order to invest in the mean-variance portfolio instead of his utility-maximising optimal portfolio. This premium depends upon, amongst other things, the joint distribution of asset returns in addition to the investor's utility function. Simaan (1993) concludes that for general distributions, the mean-variance strategy is inferior to the optimal strategy.

In response to the theoretical progress in formulating portfolio selection, Sharpe (1966) defined a technique for measuring the performance of a mutual fund. Originally named the reward-to-variability ratio, it is now referred to as the Sharpe ratio. Considering the choices a mutual fund manager is required to make, the Sharpe ratio was proposed as a means to which a portfolio can be selected amongst all feasible portfolios exhibiting an appropriate degree of risk. Based on the expected rate of return, $\mathbb{E}[r]$, and the standard deviation of returns, $\sigma(r)$, Sharpe proposed selecting the portfolio yielding the highest rate of excess return per unit of volatility (or variability),

$$S = \frac{\mathbb{E}[r] - b}{\sigma(r)}, \tag{2.4}$$

where b is the reference rate about which return is measured and is often the risk-free rate.[11] Equation (2.4) is derived from the fact that all feasible efficient portfolios lie on the straight line expressed as,

$$\mathbb{E}[r] = b + \lambda\sigma(r). \tag{2.5}$$

Here, λ represents the risk premium and is assumed to be positive in accordance with the assumption that investors are risk-averse. As pointed out by Kaplan and Knowles (2004), the Sharpe ratio quantifies the sign and magnitude of average return relative to risk given by the standard deviation. Evidently, the key assumption is that two statistics (mean and variance) are sufficient for evaluating portfolio performance (Avouyi-Dovi et al., 2004). In this respect, we observe how Sharpe's work relates to Markowitz' mean-variance model.

Despite its drawbacks, the Sharpe ratio is still frequently used today both in its own right and as a tool for comparing other performance measures (Avouyi-Dovi et al., 2004). Its use is widespread even amongst the hedge fund community despite that fact that, by and large, they seek negatively skewed returns (Passow, 2005). Stimulated by the need for an improved version of modern portfolio theory, several portfolio measures based on

[11] In Sharpe (1966), the reference rate in (2.4) was given by p. We define it as b here to be consistent with the definitions given in the rest of this thesis.

the formulation of the Sharpe ratio were introduced. We discuss several of these in the subsequent chapter.

2.2 Post-modern portfolio theory

Due to its limitations, the MPT framework has been expanded to what is now known as Post-Modern Portfolio Theory, henceforth PMPT. PMPT was first introduced to the finance literature due to Rom and Ferguson (1994) at Investment Technologies. The work of Sortino and Price (1994) provided a theoretical foundation for this branch of portfolio theory and is a response to the shortcomings of MPT. In contrast to MPT, PMPT is concerned with the return that must be achieved (or exceeded) in order to meet future liabilities and satisfy solvency constraints. Hence, PMPT lends itself well to pension fund optimisation where both assets and liabilities need to be managed.[12] The origins of this framework lie in the theory presented in Fishburn's 1977 paper (see Chapter 3 for

[12]When managing pension portfolios, it is likely that pension funds will be in a state of deficit due to liabilities (current and future) being greater than assets (returns). In the context of pension fund optimisation, in particular for the Defined Benefit setting, we have the following situation: A pension fund will carry a deficit taking into account past, current and future liabilities. Given a specified time horizon, the objective is to manage the fund such that solvency, as defined by an actuary, is achieved by the end of this time period. The length of the time horizon is due to factors including credit quality and asset performance. Every three years, under pension regulations, an actuary assesses the performance of the pension fund and ensures the solvency plan is performing.

further comments on this paper). Thanks to increased computational power and further progress in portfolio analysis, several of the limitations embedded in Markowitz' standard mean variance portfolio theory had, by the early 90s, been overcome. As an alternative to variance, downside deviation (equivalently, target semi-deviation) was introduced as a measure of risk. Rather than treating volatility equally over the entire distribution, only variation in returns below a specified target were captured; this is more consistent with the way investors perceive risk. The reasons for this are two-fold. Firstly, variation in returns above the benchmark should not necessarily be penalised. In doing so, one is limiting the potential upside to an optimal investment. Secondly, when the distributional properties of our returns exhibit skewness, using a symmetrical measure to construct an optimal portfolio can lead to underestimating (or overestimating) potential risk. Even when returns are symmetrically distributed, downside deviation can provide different information than standard deviation. As a simple illustration, let us consider an asset whose returns are normally distributed with mean, μ and standard deviation, σ. We compare the squared deviation and downside squared deviation of our returns relative to a benchmark b. For the variance,

$$\begin{aligned} \mathbb{E}[(X-b)^2] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-b)^2 e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}dx \\ &= (\mu-b)^2+\sigma^2 = V_{(b)}. \end{aligned} \tag{2.6}$$

When we measure the downside deviation, we obtain

$$\begin{aligned}
\mathbb{E}[((X-b)^-)^2] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{b}(b-x)^2 e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}dx \\
&= \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}b^2}(-b\sigma-2(\mu-b))+(\sigma^2+(\mu-b)^2)\Phi(b) \\
&= DV_{(b)}, \qquad (2.7)
\end{aligned}$$

where $\Phi(.)$ is the cumulative distribution function for the standard normal distribution.

We can observe immediately that $DV_{(b)}$ is not linearly related to $V_{(b)}$,

$$DV_{(b)} = \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}b^2}(-b\sigma-2(\mu-b))+V_{(b)}\Phi(b). \qquad (2.8)$$

Based on the concept of measuring performance and risk about a benchmark, Sortino and Price (1994) sought to define a measure in the same spirit as the Sharpe ratio which measured performance on a risk-adjusted basis. The motivation for seeking a new measure stemmed from the concept that risk should only be associated with a subset of uncertain returns. Only returns that occur below a pre-specified benchmark, not necessarily the mean, are considered unfavourable and should be minimised accordingly. With this in mind, the reference level should act as a separation between good and bad outcomes. Therefore, it does not make sense for 'good' outcomes to enter the risk calculation. Taking everything into consideration, the Sortino ratio is defined as follows:

$$SR_X(b) = \frac{\mathbb{E}[(X-b)^+]}{\sqrt{\mathbb{E}[((X-b)^-)^2]}}. \qquad (2.9)$$

We define the reference, b, here to denote the benchmark about which the investor perceives good and bad outcomes. Originally, the benchmark b stood to represent the minimum accepted return (Hoffman & Rollinger, 2013). In this context, the Sortino ratio is an appropriate tool for measuring performance and,in particular, constructing an optimal portfolio for a pension fund. We find, as will be illustrated in Chapter 4, that the Sortino ratio turns out to be a particular parameterisation of the one-sided performance ratio in (1.2).

In Chapter 4, we propose and analytically show that the Omega index and Sortino ratio are functions of the Sharpe ratio for Student t distributed returns. In Chapter 5, we discuss the recent work of Eling and Schuhmacher (2012) who prove that several post-modern risk-to-reward ratios are monotonically increasing functions of the Sharpe ratio.[13] Based on the conclusions in Eling and Schuhmacher (2012), one can justify the consistent use of the Sharpe ratio in industry. Nevertheless, towards the end of this thesis, we provide situations where agreement between asset rankings based on the Sharpe ratio, Omega and Sortino does not exist in the general case.

The key components to any portfolio optimisation problem in the PMPT framework are the distribution models and which performance measure to employ. We reiterate our view that reducing our tools to one distribution and contending ourselves with one risk performance measure is risky in itself. We discuss, in finer detail, the choice of distribution and risk-adjusted performance measure for portfolio optimisation in the next

[13]This conclusion, however, is founded on certain conditions on the distribution of the underlying variables and the performance ratio being satisfied.

chapter. There, we discuss the consequences this choice can have in the context of portfolio optimisation and risk management.

Chapter 3

Parametric portfolio optimisation

This chapter discusses the components that contribute to constructing a parametric portfolio optimisation model. In a parametric setting, we have three key decisions to make:

1. Which distribution to use for modelling the random variables, primarily the returns;
2. How we define our objective function;
3. How to characterise the relationship between the underlying random variables.

In this thesis we are, primarily, interested in investigating various alternative risk-adjusted performance ratios which can be used for the purpose of constructing a flexible portfolio optimisation model. In this chapter, we draw attention to a number of those that have been discussed in the literature. This concerns the second decision above.

For completeness, however, we discuss the nature of dependencies and marginals with familiar distributions. This is what we mean by a 'parametric' setting. When it comes to simulation, a copula and marginal quantiles suffice for simulating a parametric return dataset.

3.1 Modelling the return distributions

Accurately modelling asset returns in a parametric setting remains a challenge today, despite huge advancements in both computer technology and mathematical data analysis. Since Markowitz' seminal 1952 paper, several alternative statistical distributions have been proposed for modelling financial return data. Even as far back as 1963, Mandelbrot recognised that returns were far from Gaussian.

The task of modelling financial returns is concerned with analytically characterising the relevant features and patterns of the data in order to capture as best as possible its 'true' behaviour.[14] In the context of asset allocation and portfolio management, our choice of distribution plays a critical role in measuring risk and performance. The calculation of risk, whether on a single asset basis or at the portfolio level, is highly dependent on the assumptions and choices we make in the modelling process. Beyond the existence of higher moments, by ignoring or misrepresenting certain observable facts, risk measures

[14]What we mean here is that one endeavours to capture enough of the features of the data to illustrate the point under investigation.

including VaR become practically useless for quantifying the risk of extreme price swings (Rachev, Iotova & Stoyanov, 2010). A vast amount of literature spanning several decades on the subject of estimating return distributions exists. Yet, there is still no agreement as to which distribution is considered universal nor is there a concise methodology for selecting the most appropriate distribution under any scenario. Furthermore, there is yet to be a widespread abandonment of the normality assumption.

Constructing a model for the behaviour of price returns entails forming a hypothesis on the distribution capable of exploiting the relevant features, commonly referred to in the literature as stylised facts, (Breymann, Luthi & Platen, 2009; Jondeau, Poon & Rockinger, 2007; Cont, 2001). It is well-known that variations in asset prices are influenced by numerous events, predominantly of an economical or political nature, and that not all asset price movements will be affected in the same way. Nevertheless, based on an empirical and somewhat qualitative approach, Cont (2001) presents a set of properties common to a wide range of assets. In this paper, Cont (2001) lists a set of stylised statistical properties shared amongst a broad class of financial assets. These include absence of autocorrelations, heavy tails, asymmetry between perceptions of gain and loss, aggregational Gaussianity, intermittency, volatility clustering, conditional heavy tails, slow decay of autocorrelation in absolute returns, the effect of leverage, volume and volatility correlation and asymmetry in time scales. Faced with the challenge of choosing the best-fitting parametric model for characterising asset returns, one should therefore aim to replicate as many of these empirically observed properties as practically possible. On its own, the existence of heavy tails mentioned in Cont (2001) confirms the inadequacy of the Gaussian distribution for modelling financial data.

In light of these stylised facts, numerous phenomena must be taken into account. Depending on the nature of our model, the properties listed in Cont (2001) will have varying degrees of applicability and importance. One must consider the type of asset, or assets, in our investment universe as returns associated with all asset classes do not necessarily behave in the same way. It should not be assumed that the returns of a stock, commodity or hedge fund, for example, will behave according to the same distribution. Exogenous factors including information arrival, political occurrences and other external stochastic conditions can play a huge part in the distributional properties of an asset. In addition, the distribution of an asset can differ across varying time horizons. One often assumes that a return distribution converges to a Gaussian model as the time horizon increases. However, as shown in Breymann, Luthi and Platen (2009), this is not necessarily true. For a more detailed and theoretical analysis on the estimation of 'market invariants', see Meucci (2009).

It is common practise during the process of constructing a parametric portfolio model to make certain assumptions relating to the underlying random variables. The notorious assumption of normality in returns has frequently been discredited in the finance literature since Mandelbrot (1963) first highlighted the non-Gaussian properties of historical returns. Several academic papers based on the analysis of financial returns highlight the non-normal properties of the data, (Shaw, 2008; Mandelbrot, 1963; Fama, 1965). Yet, due to the analytical properties and readily available (off-the-shelf) numerical code for the functions associated with its distribution, the Gaussian model is still frequently used today despite the fact that this model, based solely on two moments, leads to a poor representation of the characteristics of returns. One procedure that allows the practitioner to remain within the Gaussian picture involves explaining away, and consequently eliminating, large price changes to selected causes (Mandelbrot, 1963). Removing such outliers should be avoided, particularly in the field of risk measurement where capturing extreme returns is of great importance (Platen & Sidorowicz, 2007). The two-parameter Gaussian story is a far too simplistic explanation of reality.

The main criticism of the normal distribution is that it fails to capture the existence of heavy tails on account of its zero excess kurtosis, (Platen & Sidorowicz, 2007; Rachev, Iotova & Stoyanov, 2010; Shaw, 2008; Jondeau, Poon & Rockinger, 2007). Thus, it renders the model ineffective for calculating risk which will arise due to data occurring in the tails of the distribution. As mentioned in Shaw (2011) and Fergusson and Platen (2006), in going from a Gaussian distribution to a Student t distribution with 4 degrees of freedom, we are approximately 10^{130} times more likely to observe a 25σ event. Another example is provided in Shaw (2011) where, for a Student t model with 2.25 degrees of freedom, a 30σ event is possible within roughly 100 years. A 30σ event is expected to occur approximately once in more than a few lifetimes of the universe, however, in a Gaussian model with the same mean and variance. In the context of estimating portfolio risk, assuming normally distributed returns for calculating VaR, for example, could result in a drastic underestimation of potential portfolio loss.

In light of the above, it is clear that several cases call for the need to deviate from the Gaussian world. Several alternative distributions for modelling returns have been proposed in the literature. For capturing the fat-tailed phenomenon, a Student t density is the most frequently documented alternative (Shaw, 2011; Blattberg & Gonedes, 1974; Fergusson & Platen, 2006; Breymann, Luthi & Platen, 2009; Platen & Sidorowicz, 2007; Rachev, Iotova & Stoyanov, 2010). From a practical perspective, one advantage of the Student t density is the ease with which it can be employed. Its statistical relationship with the Gaussian density means that many of the built-in Gaussian functions can be recycled in the Student t model. Furthermore, in terms of parameters, it is as well understood as the Gaussian model. A theoretical discussion of this distribution can be found in Chapter 4.

Platen and Sidorowicz (2007) conduct an empirical investigation into the distributions of log-returns of diversified world stock indices. They first hypothesise that each random variable, X, representing the return is governed by a normal mean-variance mixture distribution as follows,

$$X = m(W) + \sqrt{W}\sigma Z. \tag{3.1}$$

Here, Z is standard normal ($Z \sim N(0,1)$), $\sigma \in \mathbb{R}$ is a constant, $m : [0,\infty) \to \mathbb{R}$ is a measurable function and W≥ 0 is a non-negative random variable, independent of Z. More specifically, they apply the maximum likelihood methodology to the symmetric generalised hyperbolic density function. Based on an equally-weighted index[15] and a 99.9% level of significance, the Student t distribution with approximately 4 degrees of freedom best fitted the empirical data. The results showed a significant superiority of the Student t density compared with the normal inverse Gaussian, hyperbolic and variance gamma densities.

Employing the same methodology, Breymann, Luthi and Platen (2009) also estimate the parameters of the generalised hyperbolic distribution to fit the observed log-returns of a well-diversified stock index. For short time horizons, the Student t density with $4-6$ degrees of freedom yielded the highest likelihood. More interestingly, though, was that as the time horizon was lengthened, they did not observe an often assumed transition to the normal density. Rather, the Student t density with 4 and 6 degrees of freedom remained superior.

Since the Student t model belongs to the family of elliptical distributions, an obvious drawback is its inability to capture skewness.[16] Limiting our distribution choice to the

[15]Platen and Sidorowicz (2007) advocate the equally-weighted index on the grounds that it 'forms the best diversified portfolio' and does not require using the data to calculate the portfolio weights.

[16]These distributions are symmetrical and fully characterised by the location parameter, dispersion parameter and the generator of the probability density function (Meucci (2009)).

elliptical family may have negative consequences further on in the optimisation process. The obvious solution is to consider one of the numerous distributions capable of modelling asymmetry.

Several distributions exist that are capable of modelling asymmetry. Our purpose here is not to discuss all possible distributions, merely we wish to highlight their existence and their applicability to financial data modelling. A few of the most well-known distributions able to represent skewness include the general hyperbolic distributions, exponential distribution and lognormal distribution. One distribution in particular that is gaining in popularity in the quantitative finance literature is the Johnson family (Passow, 2005; Choi, 2001; Kaplan and Knowles, 2004 and references therein). In brief, the four-parameter Johnson family, introduced in 1949 (see Johnson (1949) for the original paper), represents a system of distributions capable of modelling a wide spectrum of shapes. It includes the normal and lognormal distributions as special cases. A random variable, X, follows a Johnson distribution if it can be expressed as,

$$X = \zeta + \lambda f^{-1}\Big(\frac{Z-\gamma}{\delta}\Big), \tag{3.2}$$

where ζ is the location parameter, λ represents the scale and γ and δ are shape parameters. The random variable Z is standard normal ($Z \sim N(0,1)$) and f can take on one of four particular functions[17]. The reader will observe that a Johnson distributed random

[17]Please refer to the appendix in Kaplan and Knowles (2004) for the four functional forms of the Johnson distribution family.

variable is simply a transformation of a standard normal variable. Its probability density function is defined as:

$$p(x) = \frac{\delta}{\lambda\sqrt{2\pi}} f'\Big(\frac{x-\zeta}{\lambda}\Big) \exp\Big(-\frac{1}{2}\Big(\gamma + \delta f\Big(\frac{x-\zeta}{\lambda}\Big)\Big)^2\Big). \tag{3.3}$$

For brevity, we mention just one member of the Johnson family: the Johnson *SU* unbounded distribution, obtained by defining f as follows,

$$f(z) = ln(z + \sqrt{z^2+1}). \tag{3.4}$$

Substituting (3.4) into (3.3), we obtain,

$$\begin{aligned} p_{S_U}(x) = \quad & \frac{\delta}{\lambda\sqrt{2\pi}} \left(\frac{(x-\zeta)^2+\lambda^2}{\lambda^2}\right)^{-1/2} \\ & \times \ \exp\left(-\frac{1}{2}\Big(\gamma + \delta \ln\Big(\frac{x-\zeta}{\lambda} + \frac{(x-\zeta)^2+\lambda^2}{\lambda^2}\Big)\Big)^2\right). \end{aligned} \tag{3.5}$$

A Johnson *SU* distributed random variable is then expressed as

$$X = \zeta + \frac{\lambda}{2}\Big(\exp\Big(\tfrac{Z-\gamma}{\delta}\Big) - \exp\Big(\tfrac{\gamma-Z}{\delta}\Big)\Big). \tag{3.6}$$

The ease with which this flexible distribution can be simulated is obvious. In Figure (3.1), we plot three Johnson *SU* probability distribution functions with equal mean and standard deviation but varying degrees of skewness. These plots illustrate the flexibility of this distributional family for the application of modelling financial data.

Several applications of the Johnson *SU* distribution in mathematical finance can be found. In the context of portfolio optimisation, Passow (2005) analytically derives

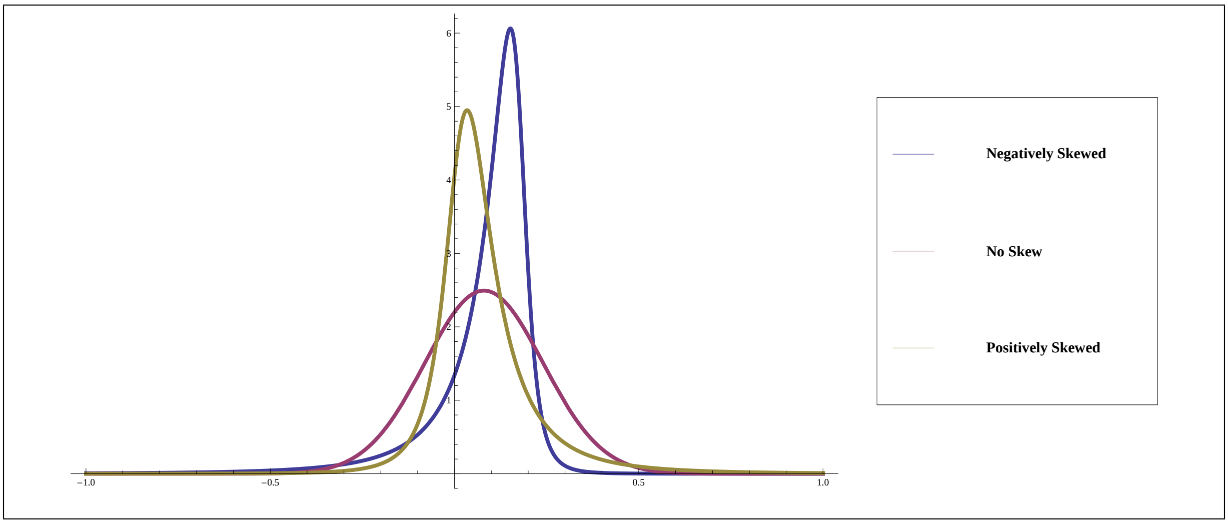

Figure 3.1: *Johnson (SU) distribution function plots. All distributions have mean = 0.08 (8%) and standard deviation = 0.16 (16%)*

the Omega ratio where returns are assumed to follow the Johnson distribution. Choi (2001), on the other hand, presents a GARCH-*SU* model for the purpose of estimating VaR[18]. Kaplan and Knowles (2004) consider assuming returns are governed by the Johnson SU distribution for the purpose of estimating Kappa, a generalised risk-adjusted performance measure which we explain later in this chapter. The aim here was to assume a particular continuous distribution in order to calculate Kappa using return distribution characteristics rather than empirical return data. In particular, Kaplan and Knowles (2004) calculated Kappa values of eleven hedge fund indices using both raw return data and estimated Johnson SU return distributions. Rankings of these hedge fund indices in

[18]In brief, a GARCH (Generalised Autoregressive Conditional Heteroskedasticity) process is a time series estimation methodology where the variances of the error term are not assumed to be constant.

terms of the resulting Kappa values based on both approaches were calculated for the purpose of quantifying whether assuming the Johnson SU distribution as a model of returns can be justified. The results showed that, for certain parameterisations of Kappa, rank correlations between empirical and Johnson SU estimated Kappa values were at or very close to 100%. What is of interest here, is the use of the Johnson distribution for the application to portfolio analysis. Kaplan and Knowles (1994) conclude that, for the purpose of 'evaluating competing investment alternatives', the Johnson SU estimated Kappa calculations were robust.

For illustrative purposes, Figure (3.2) exhibits four distributions with equal mean and standard deviation: a negatively skewed Johnson *SU* distribution, a normal distribution and two Student t distributions, one with 8 degrees of freedom, another with 3 degrees of freedom. Visually, these plots demonstrate the inadequacy of basing risk and performance measures solely on the first two distributional moments.

The subject of the next section concerns the interpretation of a risk measure, the definition of which is of great importance when formulating a risk-adjusted performance measure.

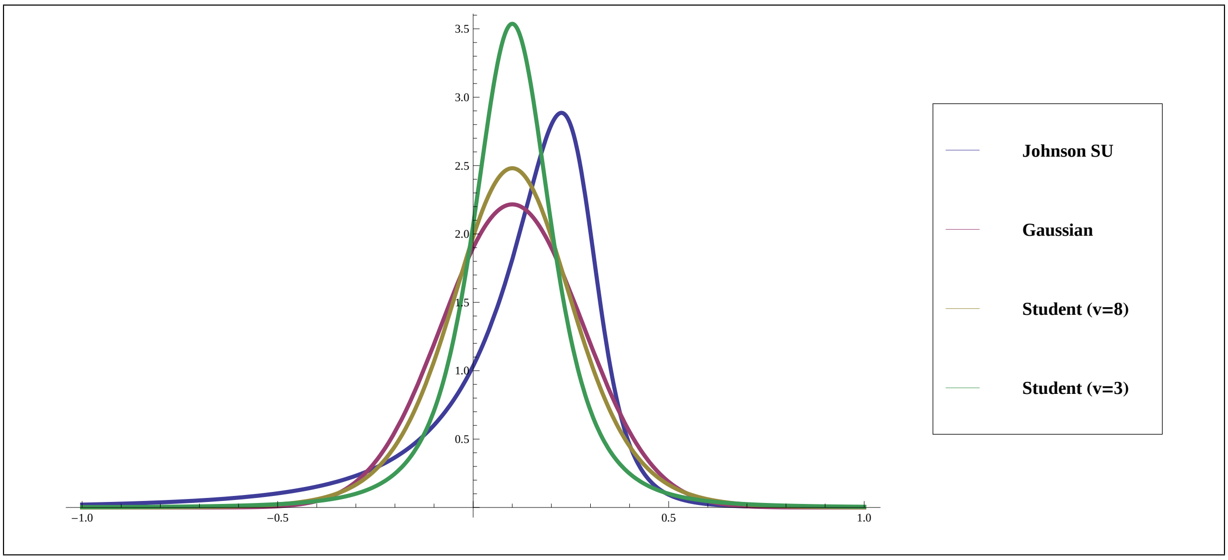

Figure 3.2: *Distribution plots with equal means and standard deviations.*

3.2 Risk measures

Before proceeding to the topic of risk-adjusted performance measures, we focus our attention on how one defines risk. Incorporating risk into a portfolio optimisation model can be done in one of two ways. Firstly, it can be integrated into the constraint set. This way, specific bounds on losses, for example, can be maintained. Second, and of particular relevance to the subsequent work in this thesis, a risk measure can be included in the objective function to facilitate maximising some measure of return and managing risk simultaneously.

An open challenge in optimal asset allocation is choosing which risk measure is most appropriate. Several factors play a role in making this decision. First and foremost, how risk will be incorporated into any analysis needs to be considered. For use in finance, investor preferences play a fundamental role in determining how risk is characterised. For the purpose of calculating an optimal portfolio, the nature of the underlying data will help to identify a suitable risk measure.

Pederson and Satchell (1998) propose a set of axioms for choosing a desirable measure of risk. These 'Basic Properties' include non-negativity, homogeneity, sub-additivity, shift-invariance and, for assigning zero weight to returns not considered to lie in the risky part of the distribution, a downside-only axiom. On page 15 of Pederson and Satchell (1998), a risk measure is defined as follows:

DEFINITION 3.1. *A risk measure in the financial sense will be a measure which, given a random variable y, returns a non-negative number on the real line:*

$$R[y] : \Omega \to \mathbb{R}^{+}.$$

Everything considered, the motivation of this section is to call attention to some of the risk measures which have been suggested in the finance literature. Since the list of risk measures that have been proposed to date is vast, we are not able to discuss the entire set without detracting from the flow of this thesis. Rather, our aim is to bring to light the wide range of tools at our disposal beyond conventional measures such as standard deviation. We concern ourselves, in what follows, with the risk measures which allow for limits on the range of returns to be included in the calculation of risk.

As far back as 1973, Stone defined a general class of risk measures encompassing several well-known dispersion measures as special cases. Two forms of this class, $L(W_0, k, A)$ and $R1(W_0, k, A)$, were presented as we show below:

$$R1(W_0, k, A) = \Big(\int_{-\infty}^{A} |W - W_0|^k dF(W)\Big)^{\frac{1}{k}} = L(W_0, k, A)^{\frac{1}{k}}, \quad k \geq 0. \tag{3.7}$$

Evidently, this family of risk measures is governed by three parameters: k, A and the target or benchmark level W_0. Depending on the choice of parameterisation, Stone's class of risk measures demonstrate various ways of calculating deviations of wealth from a specified reference level. The parameter k represents the sensitivity of deviations from W_0. When $k \geq 1$, larger values of k imply greater degrees of sensitivity to large deviations from the reference level. For k in the range $(0, 1)$, smaller deviations are weightier than larger deviations from W_0. When $k = 1$, all levels of deviation are given equal weight. Lastly, the parameter A states the range of returns accountable for risk. When $W_0 = A$, $R1(W_0, k, A)$ is equivalent to the denominator of the one-sided variability ratio in (1.2).

In the thread of research that stems from Stone's paper, Fishburn (1977) defined a two-parameter risk function represented by a probability-weighted dispersion measure. For $\alpha > 0$, Fishburns' $(\alpha - t)$ model, also referred to as a mean-risk dominance model (Fishburn, 1977), is given by:

$$F_\alpha(t) := \int_{-\infty}^{t} (t - x)^\alpha dF(x) \quad \alpha > 0, \tag{3.8}$$

where t determines the level of returns below which risk is calculated and α signifies the degree to which deviations from t are weighted. One will observe that when $W_0 = t$, $k = \alpha$ and $A = t$ in $L(W_0, k, A)$ in (3.7) above, we obtain $F_\alpha(t)$. For $\alpha = 2$, (3.8) leads to the target semi-variance model that Markowitz originally intended to use to characterise risk (Markowitz, 1959; Fishburn, 1977; Harlow, 1991).

On the basis that investors perceive gains and losses disproportionately, Harlow (1991) utilises lower partial moments for characterising a downside risk measure. For returns, R_p, and when τ represents the reference return level, Harlow (1991) defines the lower partial moments as:

$$LPM_n(\tau) = \sum_{R_p=-\infty}^{\tau} p_p(\tau - R_p)^n. \tag{3.9}$$

In (3.9) above, n represents the degree to which deviations from the reference level, τ, should be weighed and is equivalent to k in Stone's risk measure (3.7). The probability of R_p occurring is denoted by p_p. Smaller values of n indicate that one is more concerned with returns not falling below τ. On the other end of the spectrum, large values of n imply a greater sensitivity to returns falling far below τ in the lower tail of the distribution.

Various parameterisations of $LPM_n(\tau)$ lead to measures frequently mentioned in the risk literature. For any τ, $n = 0$ gives the probability of returns falling below τ, $n = 1$ yields the target shortfall and $n = 2$ leads to the target semi-variance. When τ is equal to the mean return and $n = 2$, we obtain the standard semi-variance which, for Gaussian returns, is exactly half the variance (Harlow, 1991).

Rather than depending on specific moments of the distribution, $LPM_n(\tau)$ takes the entire distribution into account (Harlow, 1991). For a given τ, the resulting value of $LPM_n(\tau)$ is determined by how much weight is given to certain parts of the return distribution about τ. For this reason, optimal asset allocations calculated via the minimisation of $LPM_n(\tau)$ for different n can differ dramatically. As will be revealed in the next section, $LPM_n(\tau)$ features in the denominator of Kappa (Kaplan & Knowles, 1994) and the one-sided variability ratio (Farinelli & Tibiletti, 2002).

Following Harlow's article on the topic of downside risk (Harlow, 1991), Pederson and Satchell (1998) survey a broad range of risk measures employed in the field of psychology, operations research, management science, economics and finance in order to construct a general class of risk measures. Based on Stone's three-parameter risk measure (3.7), for some bounded function $W(.)$ and parameters A,b,α,$\theta > 0$ ($A,\ b,\ \alpha,\ \theta \in \mathbb{R}^+$) this family of risk measures is given by:

$$R2[A, b, \alpha, \theta, W(.)] = \left(\int_{-\infty}^{A} |y - b|^{\alpha} W(F(y)) f(y) dy \right)^{\theta}. \tag{3.10}$$

Letting $W(x) = x$, $b = W_0$, $\alpha = k$ and $\theta = 1/k$ and interchanging y for W yields Stone's measure:

$$R2(A, W_0, k, \tfrac{1}{k}, W(x) = x) = \Big(\int_{-\infty}^{A} |W - W_0|^k dF(W) \Big)^{\dfrac{1}{k}} = R(W_0, k, A). \tag{3.11}$$

As with other risk measures discussed in this section, many parameterisations of the general family in (3.10) yield common risk measures including, but not limited to, standard deviation, semi-standard deviation, semi-variance and lower partial moments.

The breadth of research focusing on incorporating downside or one-sided risk into a general family of risk measures highlights the intuitive appeal for optimising a function of the returns where only those returns that negatively impact upside performance are minimised. The concept of splitting the return distribution and calculating risk based, solely, on those considered as 'bad' returns motivated the risk measures introduced above.

In the subsequent section, we introduce risk-adjusted performance measures. Of relevance to this thesis, we study the one-sided variability ratio proposed by Farinelli and Tibiletti (2002) which incorporates both upside potential and a one-sided risk measure. We use this ratio in an optimal asset allocation framework for the purpose of highlighting its differences from conventional measures, most notably the Sharpe ratio.

3.3 Risk-adjusted performance measures

Evaluating portfolio performance solely on the basis of total return ignores the notion that investors regard return in terms of the risk required to achieve it. This trade-off between risk and return is most simply characterised by the Sharpe ratio (Sharpe, 1966), the drawbacks of which have already been discussed.

Whilst Sharpe (1966) first proposed a way of quantifying risk-adjusted performance, Modigliani and Modigliani (1997) realised that the values it produced lacked any clear economic meaning. Consequently, despite its simplicity, it is not necessarily easily understood amongst investors or fund managers who are less familiar with financial theory (Modigliani & Modigliani, 1997).

To combat this shortcoming, the Modigliani risk-adjusted performance measure ($MRAP$) was proposed as an alternative to the Sharpe ratio. $MRAP$ is risk-adjusted in the sense that the relationship between a porfolios' risk and return is adjusted by the risk of the benchmark portfolio. Letting r_f represent the risk-free interest rate, r_p the portfolio return and σ_p and B representing portfolio risk and benchmark risk, respectively, $MRAP$ is defined as follows:

$$MRAP = \frac{(r_p - r_f)\sigma_B}{\sigma_p} + r_f. \tag{3.12}$$

What $MRAP$ tells us is what the return of the evaluated portfolio is when its risk is matched to that of the benchmark portfolio. By construction, it incorporates the Sharpe ratio. Multiplying by the benchmark portfolio risk, however, yields a value given in basis points which should be more economically meaningful. One can then consider it as a more economically intuitive and easier to interpret version of the Sharpe ratio.

For the purpose of evaluating mutual funds, Simons (1998) surveys Morningstar's rating methodology, the results of which are determined by use of their risk-adjusted performance measure. In this fashion, funds are rated depending on how they rank in terms of Morningstar's risk-adjusted performance measure. Where Morningstar's measure differs from the Sharpe ratio is due to how return and risk are defined. Morningstar's return for any fund is the ratio of its sales-loaded excess return (relative to the 90-day Treasury bill rate) to the average excess return for the asset class the fund under analysis belongs to. In replacement for standard deviation, risk is defined as the ratio of the funds' downside risk (based on monthly negative returns) to the average underperformance of the funds' asset class. The difference between both components above (Morningstar return - Morningstar risk) leads to Morningstar's risk-adjusted performance measure.

Thus far, the above performance measures discussed above lack the level of generality we uncovered in the previous section on the theme of risk. To accommodate the spectrum of investor preferences, one calls for a flexible class of risk-adjusted performance measures capable of representing the general investor. The Kappa ratio (Kaplan & Knowles, 1994) offers generality through the use of lower partial moments. Essentially, for any threshold return τ, Kappa is the ratio of excess return $(\mu - \tau)$ to the normalised lower partial moments:

$$K_n(\tau) = \frac{\mu - \tau}{\sqrt[n]{LPM_n(\tau)}}, \quad n \geq 0, \tag{3.13}$$

where $LPM_n(\tau)$ is given in (3.9) and was used in Harlow (1991) as a downside risk measure. When plotting $K_n(\tau)$ against τ, the steepness of the curve is governed, inversely, by the parameter n. Kaplan and Knowles (1994) evaluate the sensitivity of $K_n(\tau)$ to skewness and kurtosis for values of the threshold, τ, near to and further away from the mean return. They conclude that Kappa is relatively insensitive to skewness and kurtosis when τ lies near to or is just above the mean return but is sensitive to skewness and kurtosis for values of τ much further below the mean return. Therefore, one can think of n as representing the risk appetite for skewness and kurtosis risk aversion for threshold returns below the mean, (Kaplan & Knowles, 1994).

A similar, yet more general measure was proposed by Farinelli and Tibiletti (2002) and is the subject of the next section. Special cases of Farinelli and Tibiletti's family of performance measures, including the Omega ratio and Sortino ratio, are also special cases of Kappa.

3.3.1 A general class of performance ratios

The one-sided performance ratio (Farinelli & Tibiletti 2002), shown in equation (1.2) is flexible enough for representing a continuum of investor objectives and suitable for separating upside and downside risk, a valuable property which mean-variance analysis fails to achieve. In order to validate the method of mean-variance optimisation, one has to assume the underlying returns are normally (and hence symmetrically) distributed or that investors have quadratic utility functions (Harlow, 1991)[19]. In addition, investors are assumed to be equally averse to variation across the full distribution of returns.

[19] In defence of mean-variance analysis, if returns are approximately Gaussian, the standard deviation, being less computationally complex and more familiar, may suffice (Harlow, 1991).

Broadly speaking, minimising risk over the entire distribution will not represent the general investor's objective. The possibility that one may have a certain inclination towards upside risk and hence wish to focus on penalising downside risk means that some form of risk separation is required. Mean-variance optimisation fails to capture this asymmetry in perception of risk. Therefore, alternative tactics need to be employed.

Taking into account the spectrum of risk aversion preferences and avoiding the need to impose strong assumptions on the return distribution, Farinelli and Tibiletti (2002) propose the one-sided variability ratio, henceforth OVR, as a suitable measure when the underlying random variable potentially possesses asymmetry. Based on the investor's choice of the benchmark parameter, b, and parameteres p and q representing the orders of the right-sided and left-sided moments, respectively, OVR captures 'good' volatility above the benchmark and 'bad' volatility below the benchmark. For certain choices of p and q, the performance ratio can prove to be a more appropriate measure than quantile-based risk measures such as VaR and other risk-adjusted performance measures including Sharpe ratio. The class of performance ratios is given below:

$$\phi(b,p,q) = \frac{\mathbb{E}[|(X-b)^{+}|^{p}]^{1/p}}{\mathbb{E}[|(X-b)^{-}|^{q}]^{1/q}}. \tag{3.14}$$

For any parametrisation, OVR is a ratio of the upper partial moment and lower partial moment with respect to the boundary level, b.[20] The reader may observe that

[20] Although not directly mentioned in Farinelli and Tibiletti (2002), we take the curly brackets in their definition of the performance ratio to be modulus signs.

$\phi(b, p, q)$ in (3.14) above incorporates the n-th square root of (3.9). The denominator in (3.14) is also equivalent to the q-th root of (3.8) when $\alpha = q$ and $t = b$ which, in itself, is a particular case of (3.7).

Appropriate choices for parameters p and q in (3.14) depends on the investor's aversion to uncertainty in returns about the benchmark, b. As highlighted in Stone (1973), one must consider: 'what value of the benchmark, about which deviations we are measuring, should we choose; how we perceive small deviations versus large deviations relative to the benchmark; and which outcomes should be included in the calculations of the risk measures'. By specifying values for p and q, we are able to capture all levels of risk aversion preferences. The generality of this class of measures make them an appealing alternative with respect to traditional mean-variance theory since it possesses the ability to provide more information regarding the risk-adjusted performance of a single asset or portfolio. As pointed out by Harlow (1991), when returns are not Gaussian, this approach of allowing for asymmetry in perceptions of risk will result in portfolio decisions which disagree with those made in a mean-variance framework on account of the added information garnered from past asset returns.

The choice of orders p and q depend on our preference, or dislike for, extremal events. High values for p and q indicate greater emphasis to variation in returns far from b in the tails of the distribution. That is, further weight is given to volatile outcomes. To put it another way, high values for p (q) imply a preference (aversion) for a risky trade whereas low values for p and q indicate a low concentration on data occurring deep in the tails. For empirical data, the performance ratio can be highly beneficial in terms of capturing skewness and other distributional characteristics beyond the first two moments.

With regard to asset allocation and risk management, $\phi(b, p, q)$ provides an attractive alternative to Markowitz' mean-variance analysis on the basis of their consistency with the way investors generally perceive risk, (Harlow & Rao, 1989). In concept, $\phi(b, p, q)$ is also consistent with common market practise of calculating performance statistics relative to a benchmark (Mausser, Saunders & Seco, 2006,Kaplan & Knowles, 1994). Furthermore, one does not need to make the same, unrealistic assumptions, particularly on the family of distributions the underlying random variable should follow or the utility function it should satisfy. The benefit of utilising a downside risk measure in a performance tool is that one may employ arbitrary distributions providing a richer approach to reducing risk and, at the same time, augmenting return potential.

To illustrate the differences between various parameterisations of Farinelli and Tibiletti's one-sided performance measure (3.14), we take a universe of 18 stocks from the FTSE 100 Index and calculate optimal portfolio weights based on maximising $\phi(b, 1, 1)$ (Omega), $\phi(b, 1, 2)$ (Sortino), $\phi(b, 1, 3)$ and $\phi(b, 1, 4)$ and, for comparison, the Sharpe ratio. We vary two components of this analysis; time period and return threshold. Two distinct time periods are considered; a 'boom' period (January 2003 - July 2007) and a 'bust' period (August 2007 - March 2012) to cover the recent financial crisis. Three return thresholds are tested, $b \in \{5\%, 10\%, 15\%\}$. We use the empirical return dataset from each period for calculating the one-period optimal portfolio weights. The first four moments of each stock's return data over the boom and bust periods are given in tables (3.1) and (3.2), respectively. Optimal portfolio weights with respect to the boom period are given in table (3.3) and for the bust period, optimal portfolio weights are presented in table (3.4).

Optimal portfolio weights with respect to the boom period are given in table (3.3) and for the bust period, optimal portfolio weights are presented in tables (3.4).

One will immediately notice that portfolio weights resulting from optimising different parameterisations of $\phi(b, p, q)$ differ substantially from the equivalent Sharpe-optimal portfolios. For any return threshold, the correlation between $\phi(b, 1, q)$-optimal weights and Sharpe-optimal weights decreases considerably as we increase the parameter q. This is significantly more apparent for the 'boom' period as exhibited in Table (3.5). We leave the discussion regarding performance ratio based asset rankings to Chapter 5. In the interim, the numbers presented thus far illustrate the benefits of utilising a richer family of performance measures such as OVR for calculating optimal portfolios.

In summary, the set of risk and performance measures available to us is vast. Whilst we advocate employing one or a number of the one-sided risk-adjusted performance measures defined in (3.14), it is critical to bear in mind that traditional tools including the Sharpe ratio are not redundant. For certain distributional properties of the underlying random variables, the ranking and allocation of assets based on the Sharpe ratio are

Table 3.1: Boom period (January 2003 - July 2007): Stock Return Statistics.

Asset:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Mean Return (% p.a.)	9.50%	23.81%	5.22%	26.71%	8.56%	13.68%	23.97%	19.10%	18.23%	-8.56%	11.37%	37.04%	45.92%	3.03%	18.83%	3.95%	18.78%	34.49%
Standard Deviation (% p.a.)	15.97%	18.07%	19.98%	22.02%	19.38%	18.83%	21.81%	20.37%	23.41%	36.47%	20.85%	59.72%	30.81%	17.28%	16.23%	23.20%	18.33%	36.66%
Skewness	-0.13	1.59	0.39	0.51	0.08	0.33	0.97	0.65	1.63	-20.58	0.18	28.95	1.12	-1.62	-0.14	-1.53	1.10	-0.19
Kurtosis	3.70	14.28	4.57	5.40	1.17	2.49	7.96	9.44	22.99	599.23	14.09	944.90	7.78	17.44	2.14	24.65	8.49	5.46

Table 3.2: Bust period (August 2007 - March 2012): Stock Return Statistics.

Asset:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Mean Return (% p.a.)	11.48%	20.28%	2.53%	-10.34%	2.46%	12.19%	17.40%	-10.71%	-3.88%	-23.50%	6.27%	9.66%	17.91%	1.28%	1.31%	5.01%	7.03%	47.03%
Standard Deviation (% p.a.)	24.00%	26.35%	34.49%	41.73%	33.18%	26.00%	38.70%	38.11%	37.35%	80.17%	31.14%	26.30%	35.97%	24.29%	25.83%	30.85%	26.63%	44.25%
Skewness	0.01	0.39	-0.44	-0.14	0.17	-0.12	0.33	0.03	-1.07	-1.61	-0.31	0.16	0.20	0.19	0.02	0.73	-0.39	0.41
Kurtosis	3.68	8.47	5.93	2.98	4.87	4.18	3.09	3.18	13.95	33.15	5.70	3.75	3.32	5.68	9.93	9.71	4.65	6.32

Table 3.3: Boom period optimal portfolios.

Performance Measure	Benchmark	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Sharpe Ratio	5%	0.12%	20.56%	2.01%	18.90%	0.00%	4.84%	14.92%	0.00%	0.00%	0.00%	0.02%	5.37%	19.81%	0.07%	7.00%	0.01%	6.37%	0.00%
$\phi(b,1,1)$ - Omega	5%	0.05%	12.39%	0.02%	17.71%	0.81%	0.49%	3.35%	0.04%	0.00%	0.01%	0.64%	36.01%	19.12%	0.11%	6.84%	0.12%	1.45%	0.84%
$\phi(b,1,2)$ - Sortino	5%	0.00%	54.48%	0.24%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.02%	9.95%	22.80%	0.00%	0.27%	0.00%	0.07%	12.15%
$\phi(b,1,3)$	5%	0.00%	75.03%	0.02%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	5.60%	0.00%	0.01%	0.00%	0.00%	0.00%	19.30%
$\phi(b,1,4)$	5%	0.00%	78.54%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%	21.45%
Sharpe Ratio	10%	0.01%	39.74%	0.00%	10.58%	0.07%	0.03%	3.99%	0.01%	0.00%	0.03%	0.02%	8.33%	28.10%	0.00%	8.67%	0.01%	0.36%	0.06%
$\phi(b,1,1)$ - Omega	10%	0.00%	17.12%	0.01%	0.11%	0.00%	0.03%	0.00%	0.00%	0.38%	0.00%	0.00%	51.41%	30.88%	0.00%	0.04%	0.00%	0.03%	0.00%
$\phi(b,1,2)$ - Sortino	10%	0.00%	63.50%	0.00%	0.76%	0.01%	0.00%	0.00%	0.22%	0.00%	0.00%	0.00%	3.61%	16.48%	0.00%	0.00%	0.00%	0.06%	15.36%
$\phi(b,1,3)$	10%	0.00%	78.86%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	21.14%
$\phi(b,1,4)$	10%	0.00%	79.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	20.91%
Sharpe Ratio	15%	0.00%	17.18%	0.01%	21.68%	0.00%	0.00%	0.03%	0.03%	0.48%	0.02%	0.00%	9.63%	50.45%	0.00%	0.06%	0.00%	0.44%	0.00%
$\phi(b,1,1)$ - Omega	15%	0.01%	8.72%	0.01%	6.57%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.36%	47.53%	36.80%	0.00%	0.00%	0.00%	0.01%	0.00%
$\phi(b,1,2)$ - Sortino	15%	0.00%	55.35%	0.01%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	28.34%	0.00%	0.00%	0.01%	0.00%	16.27%
$\phi(b,1,3)$	15%	0.00%	78.49%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	21.44%
$\phi(b,1,4)$	15%	0.03%	77.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	22.09%

Table 3.4: Bust period optimal portfolios.

Perfomance Measure	Benchmark	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Sharpe Ratio	5%	0.00%	43.40%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	56.55%
Omega	5%	0.02%	44.83%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	55.15%
Sortino	5%	0.00%	18.59%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	81.38%
$\phi(1,3)$	5%	0.00%	40.01%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	11.84%	0.00%	0.03%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	48.07%
$\phi(1,4)$	5%	0.00%	0.00%	0.00%	0.00%	44.42%	0.02%	0.00%	0.00%	13.57%	0.00%	0.05%	0.19%	0.00%	0.05%	0.00%	0.00%	0.00%	41.68%
Sharpe Ratio	10%	0.01%	28.80%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	71.13%
Omega	10%	0.00%	30.45%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	69.52%
Sortino	10%	0.00%	12.16%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	87.83%
$\phi(1,3)$	10%	0.08%	37.73%	0.00%	0.00%	0.00%	0.01%	0.02%	0.00%	11.26%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.83%
$\phi(1,4)$	10%	0.00%	0.02%	0.00%	0.00%	49.84%	0.00%	0.40%	0.00%	4.34%	0.00%	0.01%	0.00%	0.00%	0.32%	0.02%	0.01%	0.00%	45.04%
Sharpe Ratio	15%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Omega	15%	0.00%	4.52%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	95.48%
Sortino	15%	0.00%	6.65%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	93.34%
$\phi(1,3)$	15%	0.00%	0.09%	0.00%	0.01%	0.01%	0.00%	0.06%	0.00%	11.19%	0.00%	20.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	68.53%
$\phi(1,4)$	15%	0.00%	0.00%	0.00%	0.00%	49.29%	0.00%	0.00%	0.03%	7.03%	0.00%	0.00%	0.00%	0.35%	0.00%	0.16%	0.02%	0.00%	43.13%

Table 3.5: Correlations between $\phi(b,1,q)$-optimal and Sharpe-optimal portfolios for $q \in \{1,2,3,4\}$

Perfomance Measure	Benchmark	Boom period	Bust Period
$\phi(b,1,1)$ - Omega	5%	56.4%	100.0%
$\phi(b,1,2)$ - Sortino	5%	60.6%	90.1%
$\phi(1,3)$	5%	43.8%	98.0%
$\phi(1,4)$	5%	43.3%	46.3%
$\phi(b,1,1)$ -Omega	10%	53.9%	100.0%
$\phi(b,1,2)$ - Sortino	10%	85.5%	96.8%
$\phi(1,3)$	10%	72.3%	94.9%
$\phi(1,4)$	10%	72.4%	57.5%
$\phi(b,1,1)$ -Omega	15%	68.2%	99.9%
$\phi(b,1,2)$ - Sortino	15%	57.2%	99.7%
$\phi(1,3)$	15%	19.1%	94.8%
$\phi(1,4)$	15%	18.9%	62.9%

similar, if not the same, to the equivalent analysis based on more complicated ratios. A thread of research in accord with this statement is given in Eling and Shuhmacher (2005) and Eling and Schuhmacher (2012) which is the subject of Chapter 5.

3.4 Dependency

As a final note in this chapter, we briefly discuss a few methods for modelling dependence between asset returns. For a number of applications in finance, modelling the dependence structure between multiple random variables is a necessary procedure. Solely relying on the correlations[21] is not always adequate for capturing the relationship between assets in a general framework. This can only be justified when the underlying distributions are

[21] More specifically, Pearson's correlation coefficient.

elliptical, for example, all Gaussian or Student t. When the underlying distributions do not belong to the elliptical family, Pearson's correlation coefficient can be a misleading metric. Of several potential undesirable properties, one may calculate a positive correlation value between two random variables when, in fact, they are not positively related to each other. See, for example, Jondeau, Poon and Rockinger (2007).

Modelling the returns of a number of assets with a multivariate Gaussian distribution is, generally, over simplistic. When such a multivariate model cannot be assumed, an analytical expression or numerical model for the multivariate distribution may not exist. This is particularly true in situations where a set of returns do not exhibit equal marginal distributions. Copula functions provide the answer to such situations and offer a generalisation to the linear dependence structure, provided by the correlation, to include non-linear dependence.

One cannot discuss the concept and use of copulas without mentioning Sklar's theorem as much of the theory associated with copulas stems from this result. For concreteness, we present Sklar's theorem below.

Sklar's Theorem.. *Let* $\mathbf{H}$ *be a two-dimensional distribution function with marginal distribution functions* $\mathbf{F}$ *and* $\mathbf{G}$*. Then there exists a copula* $\mathbf{C}$ *such that*

$$\mathbf{H}(x, y) = \mathbf{C}(\mathbf{F}(x), \mathbf{G}(y)). \tag{3.15}$$

Conversely, for any univariate distribution functions $\mathbf{F}$ *and* $\mathbf{G}$ *and any copula* $\mathbf{C}$*, the* *function* **H** *is a two-dimensional distribution function with marginals* **F** *and* **G**. *Furthermore, if* **F** *and* **G** *are continuous, then* **C** *is unique.*

In summary, what Sklar's theorem tells us is how any joint distribution can be decomposed into separate marginals and a copula. In other words, a copula provides the missing information needed to combine the marginal distributions in order to to fully describe the joint probability distribution. The theory of copulas is based on the following key results:

Result 1. Let X be a random variable. Regardless of the distribution of X, the random variable obtained from applying its cumulative distribution function to X is uniformly distributed,

$$F_X(X) \equiv U \sim U_{[0,1]}. \tag{3.16}$$

Meucci (2009) refers to these uniform variables as 'grades'.

Result 2. Given a specified target distribution, one can transform a uniform random variable into another random variable with any distribution by feeding it through the inverse cumulative distribution function of the desired distribution. This method is known as inverse transform sampling.

The uniform random variables, or grades, mentioned in Result 1 are generally not independent and it is their joint distribution which is the subject of copulas. One of the most powerful properties of copulas is that we can model any marginal distribution based on the dependency structure of the original dataset. This splitting up of the joint distribution into its marginals and copula, often referred to as the copula-marginal decomposition, is a two-stage process. The first stage requires separating the distribution into its marginals and its copula. The second process involves gluing the copula to arbitrary marginals, creating a new joint distribution.

Copulas can, broadly, be categorised into two types: parametric and nonparametric. In a nonparametric framework, empirical distribution functions can be used to construct an empirical copula, the details of which can be found in Jondeau, Poon and Rockinger (2007) (see also Meucci, 2009 and Nelsen, 2006). In a parametric setting, several families of copulas exist. Within the elliptical family, the Gaussian and Student t are familiar. The advantage of these copulas is that analytical expressions can be used to derive them. The obvious disadvantage is the symmetry of dependence they impose due to equality of dependence in both tails. These copulas share the property that they are derived from multivariate distribution functions. An alternative copula formulation exists in the Archimedean family, many of which have closed form expressions yet do not result from multivariate distribution functions. Archimedean copulas are defined by the following theorem,

Archimedean Copula:. Let φ be a continuous, strictly decreasing function from $[0, 1]$ to $[0, \infty)$, $\varphi : [0, 1] \to [0, \infty)$, such that $\varphi(1) = 0$ and $\varphi^{-}1$ be the inverse of φ. Then, the function from $[0,1]^2$ to $[0,1]$ given by

$$C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \tag{3.17}$$

is a copula if and only if φ is convex and non-increasing. The function φ is known as the generator.

Examples of Archimedean copulas include the Frank family, Gumbel and Clayton copulas. In Figure (3.3) we illustrate four different parameterisations of a Clayton copula.

We leave the discussion of copulas here. For the interested reader, an in-depth text on the subject of copulas is given in Nelsen (2006).

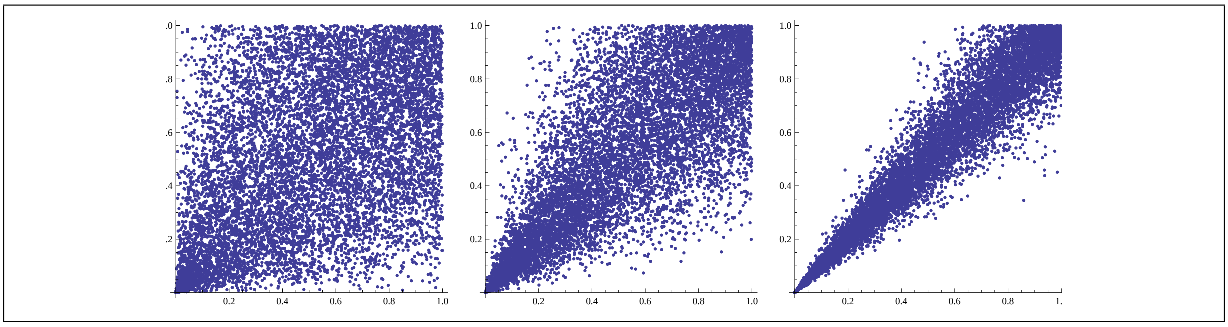

Figure 3.3: *Three parametrisations of the Clayton copula illustrating the various degrees of dependency that can be represented by the archimedean family of copulas.*

Chapter 4

Omega & Sortino Ratio analytics

4.1 The Omega Index & Sortino ratio

In this section, we introduce the Omega Index and Sortino Ratio and study their properties in the context of asset allocation and portfolio risk management. In Chapter 3, we discussed the one-sided variability ratio introduced by Farinelli and Tibiletti (2002). This measure is the ratio of normalised upper and lower partial moments and is defined as follows:

$$\phi(b, p, q) = \frac{\mathbb{E}[|(X-b)^+|^p]^{1/p}}{\mathbb{E}[|(X-b)^-|^q]^{1/q}}. \tag{4.1}$$

Unlike other commonly used measures in industry, this functional form is capable of representing a diverse range of investor preferences on account of the choice of parameters

b, p and q. The values we assign to p and q are indicative of the risk aversion preferences of the investor.

We consider two particular parameterisations of (4.1): the Omega Index ($p = q = 1$) and the Sortino ratio ($p = 1, q = 2$). These special cases have been recognised in the existing literature for applications in portfolio optimisation (Avouyi-Dovi, Morin & Neto, 2004; Chaudry & Johnson, 2008; Hoffman & Rollinger, 2013; Kane et al., 2007; Mausser, Saunders & Seco, 2006; Passow, 2005; Shaw, 2011). More often than not, Omega and Sortino are utilised for comparitive purposes with more traditional measures (Chaudry & Johnson, 2008, Eling & Schuhmacher, 2005; Eling & Schuhmacher, 2012; Nguyen-Thi-Thanh, 2007; Zakamouline, 2011). The reasons for this are twofold. Firstly, as measures for portfolio selection, they lack the same familiarity as, for example, the widely-used Sharpe ratio. Secondly, portfolios optimal with respect to Omega or Sortino can differ dramatically to optimal portfolios calculated via more traditional methods such as mean-variance analysis (Harlow, 1991). For this reason, certain financial institutions are cautious about employing new methods unless competitors simultaneously build them into their methodologies.

In this chapter, we provide an overview of the Omega and Sortino ratios and illustrate how these functions can be reduced to analytical expressions when the underlying return is distributed according to the Student t model. Familiar measures including standard deviation, VaR and the Sharpe ratio continue to be used in a risk and portfolio management setting based on the assumption of Gaussian returns. Maintaining this assumption yields neat, analytical functions expressed in terms of well-understood parameters and built-in functions. We show that Omega and Sortino can also be reduced to neat expressions without the need to assume Gaussian returns. The following analysis demonstrates the ease with which one can deviate from the restrictive Gaussian framework. Furthermore, we are able to express the Omega Index and Sortino ratio in terms of the Sharpe ratio. As we discuss in Chapter 5, Eling and Schuhmacher (2012) extend this analysis and prove that, under certain conditions (to be explained in due course), a wide range of performance measures are monotonically increasing functions of the Sharpe ratio. Based on these findings, one might question the benefit of using more complex measures such as Omega and Sortino when the Sharpe ratio may suffice under certain conditions. It certainly makes sense that if several measures effectively do the same job then one will use the measure that is easiest to implement. Nevertheless, these findings should not be taken for granted in the general case. We focus our attention to this discussion in Chapter 6 where we consider situations outside of the restrictive framework laid out in Eling and Schuhmacher (2012). There, we present a few case studies to demonstrate that optimal portfolio weights and asset rankings are not always consistent between the Omega Index, Sortino ratio and the Sharpe ratio when we relax the assumptions in Eling and Schuhmacher (2012).

For the rest of this chapter, we concentrate solely on the Omega ratio and Sortino ratio and base our calculations on assuming that returns can be fully characterised by the Student t distribution. In sections (4.2) and (4.3), we present these alternative performance ratios as particular parameterisations of the OVR (4.1). For reference, we express Omega and Sortino analytically for Gaussian returns in section (4.5) before describing the Student t distribution in section (4.6). Subsequently, we present the analytically-reduced expressions for Omega and Sortino when the underlying random return is characterised by the Student t model.

4.2 The Omega Index

Keating and Shadwick (2000; 2002) first defined the Omega Index for applications in quantitative finance. It can be obtained as a particular parametrisation of (4.1) when setting $p = q = 1$[22]:

$$\Omega(b) = \phi(b, 1, 1) = \frac{\mathbb{E}[|(X-b)^+|]}{\mathbb{E}[|(X-b)^-|]} = \frac{\mathbb{E}[(X-b)^+]}{\mathbb{E}[(b-X)^+]}. \tag{4.2}$$

In Keating and Shadwick (2000), the Omega Index was expressed in terms of the cumulative distribution function of the underlying random variable:

$$\Omega(b) = \frac{\int_r^b (1 - F(x))dx}{\int_a^r F(x)dx}, \tag{4.3}$$

where x takes values in the range $[a, b]$. One can obtain (4.3) from (4.2) by writing expectations in terms of the probability distribution function of X and integrating by

[22]We use Ω to denote the Omega Index as done in Keating and Shadwick (2000).

parts (Shaw, 2011). Basing our analysis on the former representation, (4.3), Omega can be simplified to a ratio of the excess return and lower partial moment plus an additional unit constant:

$$\begin{aligned}\Omega(b) =& \frac{\mathbb{E}[(X-b)^+]}{\mathbb{E}[(b-X)^+]} \\ =& \frac{\mathbb{E}[(b-X)^+ - (b-X)]}{\mathbb{E}[(b-X)^+]} \\ =& 1 - \frac{\mathbb{E}[(b-X)]}{\mathbb{E}[(b-X)^+]} \\ =& \frac{\mathbb{E}[X]-b}{\mathbb{E}[(b-X)^+]} + 1.\end{aligned} \tag{4.4}$$

This simplification can also be found in Kappa (2004) and is obtained without the need to impose a specific distribution on the random variable X. As a point of interest, one can also define Omega as the ratio of the price of a European call option and the corresponding European put option where expectations are taken in the real world measure (Farinelli et al., 2006). Regardless of the form in which one expresses the Omega function, it can be described as the ratio of expected gains and losses (Keating & Shadwick, 2000; Keating & Shadwick, 2002; Avouyi-Dovi et al., 2004)

In the context of optimal asset allocation, optimising Omega implies maximising the upside potential whilst simultaneously minimising downside risk. Since $p = q$ for this particular performance measure, the Omega ratio allocates equal weight (has a symmetrical preference) to small and large deviations from the reference level. In other words, all returns above the benchmark are ‘good’ no matter the degree to which the return exceeds b. This is in view of the fact that using Omega as a means for optimal asset allocation implies concentrating on the likelihood of not meeting the pre-defined return threshold (Avouyi-Dovi et al., 2004). Analogously, returns lower that the specified reference level are all considered undesirable regardless of the distance between the two. For certain investors, this is a desirable property if their objective is, primarily, to reach a chosen target but not have to pay extra in terms of risk for generating higher returns.

It is worth noting a few mathematical properties associated with Omega since it is important to identify and understand what the output values signify, particularly when carrying out a comparison against better understood measures. Firstly, Omega takes the value 1 when b is set equal to the mean return (Keating and Shadwick, 2002). This is independent of the underlying distribution and can be seen immediately by substituting $b = \mathbb{E}[X]$ into (4.4):

$$\left.\frac{\mathbb{E}[|(X-b)^+|]}{\mathbb{E}[|(X-b)^-|]}\right|_{b=\mathbb{E}[X]} = \left.\frac{\mathbb{E}[X]-b}{\mathbb{E}[(b-X)^+]}+1\right|_{b=\mathbb{E}[X]} = 1. \tag{4.5}$$

This implies that, for any application of the Omega Index to any number of random variables, the level of deviations of Omega from 1 are indicative of the desirability of each asset relative to the others. In addition, equality in the Omega value between a number of assets for a particular return threshold implies an indifference point between those assets. Furthermore, the slope of the Omega Index as a function of the benchmark is a gauge for risk; the steeper the curve, the lower the risk of the asset under consideration (Keating and Shadwick, 2002).

There exist several references to the use of Omega in the context of quantitative finance. With particular significance to this thesis, optimisation of Omega over possible portfolio weights has been considered by several authors (see, for example, Avouyi-Dovi, Morin & Neto, 2004; Chaudry & Johnson, 2008; Kane et al., 2007; Mausser, Saunders & Seco, 2006; Passow, 2005; Shaw, 2011).

The fact that the Omega Index is capable of capturing higher moment information is brought to light by Avouyi-Dovi et al. (2004) who compare Omega-optimal portfolios with portfolios optimal in terms of the Sharpe ratio. In this case, the underlying investment universe consists of three stock market indices (US, UK and Germany). Avouyi-Dovi et al. (2004) demonstrate the disagreement between both portfolios. In particular, for smaller return thresholds, differences in optimal weights are more pronounced. It is for this reason that Omega is best-suited for minimising losses (Avouyi-Dovi et al., 2004). In Chapter 3 we tabulated optimal portfolio weights for different risk-adjusted performance measure-based objective functions. We showed that optimal weights based on these measures do not agree. This is further discussed in Chapters 5 and 6.

The analyses presented in this thesis all take the return threshold to be deterministic. Generalising this approach to allow for a stochastic threshold, such as a government treasury rate or the return of an index, is the subject of Mausser, Saunders and Seco (2006) for the purpose of constructing an Omega-optimal portfolio in a non-parametric setting. Alternatively, Passow (2005) follows the parametric approach to optimising Omega where returns are modelled with the flexible family of Johnson distributions (see Chapter 3 for an overview of this distribution). As we show with Student t returns in sections (4.9) and (4.10), Passow (2005) analytically expresses the Omega ratio for Johnson distributed returns. Here, it is illustrated how and why Johnson-Omega optimal portfolios can differ from the equivalent Sharpe-optimal portfolios. As Passow (2005) points out, strategies with low volatility do not necessarily imply low risk on account of the existence of higher moments and the effect of autocorrelation. This standpoint harmonises with the view that standard deviation should not be taken as a maxim for characterising risk.

From an industry perspective, Winton Capital (2003*a*), a UK-based hedge fund, use Omega to evaluate their performance as well as the performance of other diversified CTAs (Commodity Trading Advisors). They advocate the Omega Index on the premise that it is capable of using all information regarding the return distribution and is designed to account for any risk appetite. In a later paper, Winton (2003*b*) include the Sortino ratio in their analysis and use this measure as well as Omega to rank a number of CTAs to compare with the rankings based on the Sharpe ratio. We discuss their results in Chapter 5 as their findings relate to the work of Eling and Schuhmacher (2012) and the analyses we discuss later in this thesis.

4.3 The Sortino ratio

According to Satchell (2009), the Sortino ratio was created in the 1980's by Brian Rom at Investment Technologies. It is a potentially superior performance measure compared with the Sharpe ratio on account of defining risk as the the target semi-variance being used, as opposed to the standard deviation. Assigning $p = 1$ and $q = 2$ in (4.1) results in the Sortino ratio:

$$SR(b) = \phi(b, 1, 2) = \frac{\mathbb{E}[((X - b)^{+})]}{\mathbb{E}[|(X - b)^{-})^{2}|]^{1/2}}. \tag{4.6}$$

In practise, it has been frequently used as a performance measuring tool in the pension and mutual fund industry (Hoffman & Rollinger, 2013). In this context, portfolio and risk managers are required to guarantee a level of profitability whilst simultaneously meeting future liabilities. That is, the funding ratio (the ratio of the value of assets to the value of liabilities) needs to be maintained in order to meet solvability constraints. For this purpose, the benchmark b in (4.6) is often set equal to the 'Minimum Acceptable Return' (MAR) (Hoffman & Rollinger, 2013):

$$\phi(b = MAR, 1, 2) = \frac{\mathbb{E}[((X - MAR)^{+})]}{\mathbb{E}[|(X - MAR)^{-})^{2}|]^{1/2}}.$$

Optimising the Sortino ratio implies maximising the upside potential whilst minimising the downside semi-standard deviation. In other words, the Sortino ratio is a useful measure if we desire outcomes which yield returns exceeding the benchmark and simultaneously wish to penalise outcomes yielding relatively high volatility below the benchmark. Unlike Omega, the Sortino ratio does not exhibit a symmetrical preference for data occurring either side of the benchmark (Pedersen & Satchell, 1998). The downside risk component in the denominator gives further weight to extreme tail values than the denominator of Omega.

On a theoretical note, one must not confuse the upside potential return with the conditional mean of the excess return given the return exceeds the benchmark, (Satchell, 2009). The numerator of Sortino is simply a measure of excess return. As for the denominator, the use of downside variance was inspired by Fishburn (1977) and, as also mentioned by Sortino and Van der Meer (1991), reflects the rational attitude of risk aversion for underperformance. Taking the square root then makes it consistent (in terms of units) with the numerator.

As a means of measuring downside risk, Sortino and Van der Meer (1991) introduced the Sortino ratio for asset allocation purposes as an alternative to the classical mean-variance methodology. As they explain, mean-variance analysis can result in selecting two assets based on their net volatility being zero. However, this does not take into account for considering whether these assets will yield a desired return and does not, like the Sortino ratio, make a distinction between 'good' and 'bad' volatility.

In summary, the Omega Index and Sortino ratio offer alternative ways of evaluating portfolio performance measures to the classical Sharpe ratio. To advocate the use of Omega, Keating and Shadwick (2000) call attention to situations where the ranking of assets by Sharpe can differ dramatically to the ranking of assets by Omega. Furthermore, given a portfolio of differently distributed assets, measuring performance by Sharpe is not optimal, (Farinelli & Tibiletti, 2006). As often is the case, when the investor wishes to control the overperformance or underperformance, using the Sharpe ratio can lead to misleading results, (Farinelli et al., 2008).

4.4 General reduction of $\phi(b, p, q)$

As a point of interest and to lead us in to the subsequent analysis, we show that the one-sided variability ratio, $\phi(b, p, q)$, emerges as a function of the Sharpe ratio (given as $s = (\mu - b)/\sigma$ in this case) when the underlying random variable is a linear transformation of a random variable with zero mean and unit standard deviation. That is, we suppose that X in (4.1) is such that $X = \mu + \sigma W$ where W has zero mean and unit standard deviation. Substituting X for $\mu + \sigma W$ in the numerator of $\phi(b, p, q)$, we have,

$$\begin{aligned} \mathbb{E}[|(X-b)^+|^p]^{1/p} &= \mathbb{E}[|((\mu+\sigma W)-b)^+|^p]^{1/p} \\ &= \mathbb{E}[|\sigma(\tfrac{\mu-b}{\sigma}+W)^+|^p]^{1/p} \\ &= \mathbb{E}[\sigma^p|(s+W)^+|^p]^{1/p} \\ &= \sigma\mathbb{E}[|(s+W)^+|^p]^{1/p}, \end{aligned} \tag{4.7}$$

where, again, $s = (\mu - b)/\sigma$. Similarly, the denominator reduces to,

$$\begin{aligned} \mathbb{E}[|(X-b)^-|^q]^{1/q} &= \mathbb{E}[|(\mu+\sigma W-b)^-|^q]^{1/q} \\ &= \mathbb{E}[|\sigma(\tfrac{\mu-b}{\sigma}+W)^-|^q]^{1/q} \\ &= \mathbb{E}[\sigma^q|(s+W)^-|^q]^{1/q} \\ &= \sigma\mathbb{E}[|(s+W)^-|^q]^{1/q}. \end{aligned} \tag{4.8}$$

After cancelling the common σ term, the one-sided variability ratio is expressed as,

$$\phi(b,p,q) = \frac{\mathbb{E}[|(s+W)^+|^p]^{1/p}}{\mathbb{E}[|(s+W)^-|^q]^{1/q}}, \tag{4.9}$$

which depends on b through s.

4.5 The Omega Index and Sortino ratio under Gaussian returns

In this section, we present an explicit formula for the Omega index and Sortino ratio when the underlying random variable is modelled in a Gaussian framework. Whilst Shaw (2011) recently presented the analytical form of the Omega Index in a Gaussian world, the author is unaware of an explicit formula for the Sortino in a Gaussian framework being documented. In a recent publication, Hall and Satchell (2010) present an analytic solution of the downside semi-variance (the squared denominator of the Sortino ratio) when the underlying is modelled by a Gaussian random variable. In their article, the portfolio return is denoted as r_p, p and t refer to portfolio and reference level, respectively and $\mu_t = \mu_p - t$, that is, the portfolio excess return. When $r_p \sim \mathcal{N}(\mu_p, \sigma_p^2)$, we have the following:

$$\begin{aligned} I(t) &= \int_{-\infty}^{t} (t - r_p)^2 pdf(r_p) dr_p \\ &= (\sigma_p^2 + \mu_t^2)\Phi\left(-\frac{\mu_t}{\sigma_p}\right) - \mu_t \sigma_p \phi\left(\frac{\mu_t}{\sigma_p}\right), \end{aligned} \tag{4.10}$$

where *pdf* denotes the Gaussian density function:

$$pdf(r_p) = \frac{1}{\sigma_p\sqrt{2\pi}} \exp\Big(-\frac{(r_p - \mu_p)^2}{2\sigma_p^2}\Big). \tag{4.11}$$

For the rest of this section, let X denote a Gaussian random variable with mean μ and variance σ^2. Therefore, X can be written as $X = \mu + \sigma Y$, where Y is a standard normal random variable ($Y \sim N(0, 1)$). As shown in Shaw (2010; 2011), in a Gaussian world, Omega can be simplified to the following expression:

$$\begin{aligned} \Omega(b) &= \frac{\mathbb{E}[(X - b)^+]}{\mathbb{E}[(b - X)^+]} \\ &= \frac{\mathbb{E}[b - X)^+ - (b - X)]}{\mathbb{E}[(b - X)^+]} \\ &= 1 - \frac{b - \mu}{\mathbb{E}[(b - X)^+]} \\ &= 1 - \frac{b - \mu}{\mathbb{E}[(b - \mu - \sigma Y)^+]} \\ &= 1 - \frac{b - \mu}{\sigma \mathbb{E}[(\frac{b-\mu}{\sigma} - Y)^+]} \\ &= 1 - \frac{z}{\mathbb{E}[(z - Y)^+]}, \end{aligned} \tag{4.12}$$

which depends on b through z where $z = (b - \mu)/\sigma$ (the negative Sharpe ratio).

To reduce this expression fully, we therefore only need to reduce the denominator in (4.12) to an analytical expression. In this section, functions $\Phi(.)$ and $\phi(.)$ will denote the standard normal cumulative distribution function and probability density function, respectively. The expectation in (4.12) can then be written as follows:

$$\begin{aligned}\mathbb{E}[(z-Y)^+] =&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}(z-y)e^{-\frac{1}{2}y^2}dy \\ =&z\Phi(z)+\phi(z).\end{aligned} \tag{4.13}$$

As a result, for a normally distributed random variable, the Omega index reduces to:

$$\Omega(b) = 1 - \frac{z}{z\Phi(z)+\phi(z)}. \tag{4.14}$$

Substituting $z = -s$ in (4.14) where $s = (\mu - b)/\sigma$, we have:

$$\Omega(b) = 1 + \frac{s}{\phi(s) - s(1-\Phi(s))}. \tag{4.15}$$

For reducing Sortino in the Gaussian framework, we make use of the general reduction given in (4.9) above where W is a standard normal random variable, $W = Z \sim N(0,1)$, and s denotes the Sharpe ratio as defined for the Gaussian-Omega above:

$$SR(b) = \frac{\mathbb{E}[(s+Z)^+]}{\mathbb{E}[((s+Z)^-)^2]^{1/2}}. \tag{4.16}$$

The numerator in (4.16) is then evaluated as follows:

$$\begin{aligned}\mathbb{E}[(s+Z)^+] =& \int_{-s}^{\infty}(s+z)\phi(z)dz \\ =& s\int_{-s}^{\infty}\phi(z)dz + \int_{-s}^{\infty}z\phi(z)dz \\ =& s\Phi(s)+\phi(s).\end{aligned} \tag{4.17}$$

The expression under the square root sign in the denominator of (4.16) is then:

$$\begin{aligned}\mathbb{E}[((s+Z)^-)^2] =& \int_{-\infty}^{-s}(s+z)^2\phi(z)dz \\ =& s^2\int_{-\infty}^{-s}\phi(z)dz + 2s\int_{-\infty}^{-s}z\phi(z)dz \\ &+ \int_{-\infty}^{-s}z^2\phi(z)dz \\ =& s^2-(s^2-1)\Phi(s)-s\phi(s).\end{aligned} \tag{4.18}$$

Therefore, the Gaussian-Sortino can be expressed as follows:

$$SR(s) \quad = \quad \frac{s\Phi(s)+\phi(s)}{\sqrt{s^2-(s^2-1)\Phi(s)-s\phi(s)}}. \tag{4.19}$$

Since the evaluation of the Omega Index and Sortino ratio in section (4.8) is based on the underlying random variable being modelled by the Student t distribution, we dedicate the next section to describing this model and its analytical relationship with the Gaussian distribution.

4.6 An overview of the Student t distribution

The Student t distribution was introduced into the statistical literature in 1908 by William Gosset. It is a particular case of the generalised hyperbolic distribution, a member of the mean-variance normal mixture distributions (Breymann, Luthi & Platen, 2009). Similar to the normal distribution, it belongs to the family of elliptical distributions and is therefore symmetrical. However, due to the denominator in the characterisation of a Student t statistic, this distribution is more likely to produce data points that lie far from the mean. As defined in Platen and Sidorowicz (2007), one can derive the Student t density function from a random variable modelled as a normal mean-variance mixture distribution defined as

$$X = m(W) + \sqrt{W}\sigma Z. \tag{4.20}$$

Here, Z is a standard normal random variable ($Z \sim N(0,1)$), $W \geq 0$ is a non-negative random variable independent of Z, $\sigma \in \mathbb{R}$ is a constant and $m : [0, \infty) \to \mathbb{R}$ is a measurable function. In Shaw (2008), when X in (4.20) above is a standard Student t random variable, we have $m(W) = 0$ and $W = \nu / C_\nu^2$ where C is a Chi-squared random variable with ν degrees of freedom. In this case, the associated probability density function of X is given by

$$f_X(x) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma[(v+1)/2]}{\Gamma[v/2]} \left(1 + \frac{x^2}{v}\right)^{-(v+1)/2}. \tag{4.21}$$

4.7 A note on the multivariate Student t

In multidimensional problems, we need to consider the multivariate Student t model. As mentioned in Shaw and Lee (2008) and Shaw (2011), there are several methods in which we can generalise to the multivariate t case.[23] The multivariate model we consider is given in Shaw (2011) and is of the form where the random variables X_i are given as follows:

$$X_i = (\chi^2_\nu/\nu)^{-1/2} \sum_{k=1}^{N} A_{ik} Z_k. \tag{4.22}$$

As before, Z_j represents a standard Gaussian random variable for each $j \in \{1, \ldots, N\}$, χ_ν is a Chi-squared random variable with ν degrees of freedom and A_{ik} is a mixing matrix. The covariance of this system is then as shown below,

$$\mathbb{E}[X_i X_j] = \frac{\nu}{\nu - 2}(A.A^T)_{ij}. \tag{4.23}$$

For weights, w_i, the random vector,

$$\mathbf{X} = \sum_{i=1}^{N} w_i X_i, \tag{4.24}$$

will also be a Student t distributed random variable with variance,

$$V(X) = \frac{\nu}{\nu - 2} \sum_{ij} w_i (A.A^T)_{ij} w_j. \tag{4.25}$$

The author wishes to elucidate the fact that by no means do we suggest that all financial returns should be modelled by the Student t distribution. Several alternatives

[23] The interested reader may wish to consult the study by Shaw and Lee (2008) on generating bivariate t distributions and discussions on the grouped t distributions and copulas.

to the Gaussian distribution exist, one example being the Johnson family (Johnson, 1949) discussed in Chapter 3. Nevertheless, numerous statistical evaluations of empirical data point to the Student t model as the most appropriate distribution for describing the behaviour of financial returns (Platen and Sidorowicz, 2007); Breymann, Luthi & Platen, 2009; Shaw, 2011).

4.8 The Omega Index and Sortino ratio for Student t distributed returns

To facilitate evaluations carried out in this chapter, the calculations presented are based on the following form of a general Student t distributed random variable with ν degrees of freedom, mean μ, scale σ and standard deviation $\sigma\sqrt{\nu/(\nu-2)}$:

$$T_\nu = \mu + \sigma X, \tag{4.26}$$

where X is a standard Student t random variable with zero mean and variance $\nu/(\nu-2)$ as described in section (4.6). To simplify notation, we denote the standard Student t probability density function, $\phi_X(x)$, and cumulative density function, $\Phi_X(x)$ as follows:

$$\phi_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} = \gamma(\nu)(\nu+x^2)^{-(\nu+1)/2}, \tag{4.27}$$

where Γ represents the Gamma function and $\gamma(\nu)$ is given below:

$$\gamma(\nu) = \frac{\nu^{\nu/2}}{\sqrt{\pi}}\frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)}. \tag{4.28}$$

The cumulative density function is defined as:

$$\Phi_X(x) = \int_{-\infty}^{x} \phi_X(x)(u)du = 1 - \frac{1}{2} I_{X_\nu}\Big(\frac{\nu}{2}, \frac{1}{2}\Big), \quad z(x) = \frac{\nu}{x^2 + \nu}, \tag{4.29}$$

where I is the regularised incomplete beta function.

The Sharpe ratio for Student t distributed returns which we will denote as S_{T_ν}, where a Student t random variable is denoted as in (4.26) and the excess return is measured relative to a generic benchmark, b, is then:

$$\begin{aligned} S_{T_\nu} &= \frac{\mathbb{E}[T_\nu] - b}{Vol[T_\nu]} \\ &= \frac{\mu - b}{\sigma\sqrt{\dfrac{\nu}{\nu - 2}}} \\ &= \frac{\mu - b}{\sigma\nu^*}, \end{aligned} \tag{4.30}$$

where $\nu^* = \sqrt{\nu/(\nu - 2)}$. The scalar ν^* is determined by the degrees of freedom and will become useful in what follows.

It is interesting to note that, regardless of the values we choose for parameters p and q in (4.1), when the underlying random variable is distributed according to the Student t model, one can reduce (4.1) as will be shown below.

For any p,

$$\begin{aligned} \mathbb{E}[|(T_\nu - b)^+|^p]^{1/p} =& \mathbb{E}[|(\mu + \sigma X - b)^+|^p]^{1/p} \\ =& \sigma\mathbb{E}[((\frac{\mu - b}{\sigma} + X)^+)^p]^{1/p} \\ =& \sigma\mathbb{E}[((\nu^* S_{T_\nu} + X)^+)^p]^{1/p} \end{aligned} \tag{4.31}$$

where S_{T_ν} is defined in (4.31). Similarly, for any q,

$$\begin{aligned}\mathbb{E}[((b-T_\nu)^+)^q]^{1/q} =& \mathbb{E}[|(b-\mu-\sigma X)^+|^q]^{1/q}\\ =& \sigma\mathbb{E}[((\frac{b-\mu}{\sigma}-X)^+)^q]^{1/q}\\ =& \sigma\mathbb{E}[((-\nu^* S_{T_\nu}-X)^+)^q]^{1/q}.\end{aligned} \tag{4.32}$$

After cancelling the common term, σ, in both (4.31) and (4.32), we obtain:

$$\frac{\mathbb{E}[|(T_\nu-b)^+|^p]^{1/p}}{\mathbb{E}[|(b-T_\nu)^+|^q]^{1/q}} = \frac{\mathbb{E}[((\nu^* S_{T_\nu}+X)^+)^p]^{1/p}}{\mathbb{E}[((-\nu^* S_{T_\nu}-X)^+)^q]^{1/q}}. \tag{4.33}$$

Therefore, Student-Omega and Student-Sortino can be expressed, respectively, as:

$$\Omega_{T_\nu} = \frac{\mathbb{E}[(\nu^* S_{T_\nu}+X)^+]}{\mathbb{E}[(-\nu^* S_{T_\nu}-X)^+]} = 1 + \frac{\nu^* S_{T_\nu}}{\mathbb{E}[(-\nu^* S_{T_\nu}-X)^+]}. \tag{4.34}$$

$$SR_{T_\nu} = \frac{\mathbb{E}[(\nu^* S_{T_\nu}+X)^+]}{\mathbb{E}[((-\nu^* S_{T_\nu}-X)^+)^2]^{1/2}}. \tag{4.35}$$

In the next two sections, we focus our efforts on expanding (4.34) and (4.35) further to explicit functions of the Student t Sharpe ratio.

4.9 The Omega Index as a function of the Sharpe ratio

In (4.34) above, we partially reduced the Omega Index for Student t returns. To reduce this expression even further, we need to expand the expectation in the denominator in

(4.34) to characterise Omega explicitly in terms of the Student t Sharpe ratio. Hence,

$$\begin{aligned}
\mathbb{E}[(-\nu^* S_{T_\nu} - X)^+] &= \int_{\infty}^{-\nu^* S_{T_\nu}} (-\nu^* S_{T_\nu} - x)\phi_x(x)dx \\
&= -\nu^* S_{T_\nu} \int_{-\infty}^{-\nu^* S_{T_\nu}} \phi_x(x)dx - \int_{-\infty}^{-\nu^* S_{T_\nu}} x\phi_x(x)dx \\
&= -\nu^* S_{T_\nu} \Phi_X(-\nu^* S_{T_\nu}) - \gamma(\nu) \int_{-\infty}^{-\nu^* S_{T_\nu}} x(\nu + x^2)^{-(\nu+1)/2} dx \\
&= -\nu^* S_{T_\nu} \Phi_X(-\nu^* S_{T_\nu}) - \frac{\gamma(\nu)}{(1-\nu)} \int_{-\infty}^{-\nu^* S_{T_\nu}} \frac{d}{dx}\Big((\nu + x^2)^{-(\nu-1)/2}\Big) dx \\
&= -\nu^* S_{T_\nu} \Phi_X(-\nu^* S_{T_\nu}) - \frac{\gamma(\nu)}{(1-\nu)} \Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2} \\
&= -\Big(\nu^* S_{T_\nu} \Phi_X(-\nu^* S_{T_\nu}) + \frac{\gamma(\nu)}{(1-\nu)} \Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2}\Big).
\end{aligned}$$

Therefore, we have:

$$\Omega_{T_v}(S_{T_\nu}) = 1 - \frac{\nu^* S_{T_\nu}}{\Big(\nu^* S_{T_\nu} \Phi_X(-\nu^* S_{T_\nu}) + \frac{\gamma(\nu)}{(1-\nu)} \Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2}\Big)}. \tag{4.36}$$

Note that $\Omega_{T_v}(S_{T_\nu})$ is a function of b through the Student t Sharpe ratio S_{T_ν}.

4.10 The Sortino ratio as a function of the Sharpe ratio

To fully reduce the Sortino ratio, we start with the expression in (4.35) and expand the numerator and denominator separately. In this fashion, the numerator is evaluated as follows:

$$
\begin{aligned}
\mathbb{E}[(\nu^* S_{T_\nu} + X)^+] =& \int_{-\nu^* S_{T_\nu}}^{\infty} (\nu^* S_{T_\nu} + x)\phi_X(x)dx \\
=& \nu^* S_{T_\nu} \int_{-\nu^* S_{T_\nu}}^{\infty} \phi_X(x)dx + \int_{-\nu^* S_{T_\nu}}^{\infty} x\phi_X(x)dx \\
=& \nu^* S_{T_\nu} \int_{-\infty}^{\nu^* S_{T_\nu}} \phi_X(x) + \int_{-\infty}^{\nu^* S_{T_\nu}} x\phi_X(x)dx \\
=& \nu^* S_{T_\nu} \Phi_X(\nu^* S_{T_\nu}) + \gamma(\nu) \int_{-\infty}^{\nu^* S_{T_\nu}} x(\nu + x^2)^{-(\nu+1)/2}dx \\
=& \nu^* S_{T_\nu} \Phi_X(\nu^* S_{T_\nu}) + \frac{\gamma(\nu)}{(1-\nu)}\Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2}.
\end{aligned}
\tag{4.37}
$$

The denominator is then the square root of the following:

$$
\begin{aligned}
\mathbb{E}[((-\nu^* S_{T_\nu} - X)^+)^2] =& \int_{-\infty}^{-\nu^* S_{T_\nu}} (\nu^* S_{T_\nu} + x)^2\phi_X(x)dx \\
=& (\nu^* S_{T_\nu})^2 \Phi_X(-\nu^* S_{T_\nu}) + 2\nu^* S_{T_\nu} \int_{-\infty}^{-\nu^* S_{T_\nu}} x\phi_X(x)dx \\
&+ \int_{-\infty}^{-\nu^* S_{T_\nu}} x^2\phi_X(x)dx \\
=& (\nu^* S_{T_\nu})^2 \Phi_X(-\nu^* S_{T_\nu}) + \frac{2\nu^* S_{T_\nu}\gamma(\nu)}{(1-\nu)}\Big(\nu + (\nu^* S_{T_\nu})^2)\Big)^{-(\nu-1)/2} \\
&- \int_{-\infty}^{-\nu^* S_{T_\nu}} x^2\phi_X(x)dx.
\end{aligned}
\tag{4.38}
$$

The third term above can be integrated by parts to give:

$$
\begin{aligned}
\int_{-\infty}^{-\nu^* S_{T_\nu}} x^2\phi_X(x)dx =& -\frac{\gamma(\nu)}{(1-\nu)}\nu^* S_{T_\nu}(\nu + (\nu^* S_{T_\nu})^2)^{-(\nu-1)/2} + \frac{1}{(1-\nu)}\Phi_X(-\nu^* S_{T_\nu}) \\
=& \frac{1 + (1-\nu)(\nu^* S_{T_\nu})^2}{(1-\nu)}\Phi_X(-\nu^* S_{T_\nu}) + \frac{\nu^* S_{T_\nu}\gamma(\nu)}{(1-\nu)}\Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2} \\
=& \frac{1 + (1-\nu)(\nu^* S_{T_\nu})^2}{(1-\nu)}\Phi_X(-\nu^* S_{T_\nu}) + \frac{\nu^* S_{T_\nu}}{(1-\nu)}(\nu + (\nu^* S_{T_\nu})^2)\phi_X(\nu^* S_{T_\nu}).
\end{aligned}
\tag{4.39}
$$

The Sortino ratio as an explicit function of the Student t Sharpe ratio is thus:

$$SR_{T_\nu}(S_{T_\nu}) = \frac{\nu^* S_{T_\nu}\Phi_X(\nu^* S_{T_\nu}) + \dfrac{\gamma(\nu)}{(1-\nu)}\Big(\nu + (\nu^* S_{T_\nu})^2\Big)^{-(\nu-1)/2}}{\sqrt{\dfrac{1+(1-\nu)(\nu^* S_{T_\nu})^2}{(1-\nu)}\Phi_X(-\nu^* S_{T_\nu}) + \dfrac{\nu^* S_{T_\nu}}{(1-\nu)}(\nu + (\nu^* S_{T_\nu})^2)\phi_X(\nu^* S_{T_\nu})}}. \tag{4.40}$$

Note that both (4.39) and (4.40) both depend on b through S_{T_ν}.

Chapter 5

Consistency of risk-to-reward measures

Empirically, one often finds that many risk-adjusted performance measures lead to the same rankings as the Sharpe ratio when used for comparing alternative investments (Eling & Schuhmacher (2005)). Hence, for a long time now, one has suspected that such findings can be justified theoretically. Eling and Schuhmacher (2012) present us with this theoretical justification provided certain conditions on the performance ratio and the underlying random returns are satisfied. In Eling and Schuhmacher (2012), it is shown that, under certain conditions, partial-moments-based performance measures, value-at-risk-based performance measures and other admissible performance measures are monotonically increasing functions of the Sharpe ratio. This generalises the work we presented in Chapter 4 where we reduced the Omega Index and the Sortino ratio to explicit functions of the Sharpe ratio based on Student t returns.

The motivation for the study in Eling and Schuhmacher (2012) stems from the research carried out by Meyer (1987) who showed that expected utility implies Sharpe ratio rankings if the location-scale (LS) property is satisfied. The conclusions reached in Eling and Schuhmacher (2012) are strongly based on the LS property. In short, two random variables satisfy the LS condition if they differ only by their location and scale parameters. Meyer (1987) defines the LS property in terms of the cumulative distribution functions. In this manner, two random variables are said to satisfy LS if their cumulative distribution functions, $G_1(.)$ and $G_2(.)$, differ only by their location and scale parameters. Mathematically, this implies that $G_1(x) = G_2(\alpha + \beta x)$ where x represents the random variable and α and β are the (constant) difference in their location and scale parameters, respectively. The LS property can also be defined in terms of the respective probability distribution functions (see Meyer (1987)). To summarise, two random variables satisfying the LS property, as described above, once de-meaned and de-scaled, will be equal in distribution.

The analysis carried out in Eling and Schuhmacher (2012) also builds on an earlier empirical study (Eling & Schuhmacher (2005)), where the rankings of hedge funds based on a set of performance measures (including the Sharpe ratio) were compared. They found that, despite the non-normal behaviour of the underlying returns, rankings between ten alternative performance measures (including Omega and Sortino) and the Sharpe ratio when applied to the individual hedge funds generated very high correlations. This case study prompted the need to lay out a theoretical framework in Eling and Schuhmacher (2012) for why these performance measures were ranking assets, by and large, in the same order. For clarity, and since their work is highly relevant to this thesis, we recapitulate the findings provided in Eling and Schuhmacher (2012) below.

Within the literature concerning performance and risk measures, several sets of axioms exist which outline a maxim that such a measure should satisfy (Artzner et al. (1999), Pedersen & Satchell (1998)). For example, in Artzner et al. (1999), a list of four properties a risk measure should possess in order to be defined as 'coherent' is proposed. Discussions on this subject can also be found in Fishburn (1977), Meucci (2009) and Stone (1973). The LS property is one of the conditions required in order to arrive at the conclusions in Eling and Schuhmacher (2012). In addition, the following conditions are required of the reward ($\pi(x)$) and risk ($\rho(x)$) measures:

Positive Homogeneity

- $\pi(k.x) = k.\pi(x)$
- $\rho(k.x) = k.\rho(x)$ for all $k > 0$ and $\rho(0) = 0$.

Functional Translation Invariance.

- $\pi(x + s) > \pi(x)$ for $s > 0$
- $\rho(x + s) \leq \rho(x)$ for $s > 0$.

Positive homogeneity is a condition commonly demanded of a risk measure, (see, for example, Meucci (2009)). As noted in Eling and Schuhmacher (2012), positive homogeneity corresponds to a performance measure being invariant to leverage. In addition, functional translation invariance should be intuitive; adding a positive constant to a random variable, in this case the asset return, is a deterministic advantage. Consequently, in terms of risk, adding a value that is known with certainty should not increase risk. Rather, it can only lower the risk measure.[24]

Our approach in Chapter 4 was to consider two specific performance measures; the Omega Index and the Sortino Ratio. Instead, the analysis carried out by Eling and Schuhmacher (2012) is centered on a more general performance measure, $P(X_i)$, expressed as the ratio of a reward measure, $\pi(X_i)$, and risk measure, $\rho(X_i)$. The performance measure, $P(X_i)$, embodies a large subset of known risk-to-reward ratios including the one-sided variability ratio (4.1). Other examples include performance ratios based on VaR and CVaR and the Generalised Rachev ratio defined in Eling and Schuhmacher (2012) as $\mathbb{E}[X|X \geq -VaR(X; 1-\beta)]/\mathbb{E}[-X|X \leq -VaR(X; \alpha)]$ where $VaR(X; p)$ defines VaR for probability level p. Since the set of admissible performance measures in their study includes (4.1), the Omega Index and Sortino ratio are, therefore, also included.

Proving such admissible measures, henceforth ES measures, are increasing functions

[24]Strictly speaking, invariance refers to the property that a function behaves sensibly under translation, that is, that $f(x+h) = f(x)$.

of the Sharpe ratio is a consequence of the positive homogeneity and functional translation invariance axioms shown above. Suppose X_1 and X_2 are two random variables representing asset returns. When the Sharpe ratios of X_1 and X_2 exhibit equality, inequality or second-degree stochastic dominance, the same relation manifests itself for all other ES performance measures[25]. To add to the above, if the Sharpe ratios of X_1 and X_2 are identical then so will their respective ES performance ratios. Furthermore, when their Sharpe ratios disagree, for example $S(X_1) < S(X_2)$, then the same relation will exist between their respective ES measures. As a result of the above, every ES measure is an increasing function of the Sharpe ratio. The proofs of the preceding statements can be found in Eling and Schuhmacher (2012).

In further support of the conclusions reached in Eling and Schuhmacher (2012), Chen, He and Zhang (2011) prove that risk-adjusted performance measures based on lower partial moments (3.9) which include both the Omega Index and the Sortino ratio, are equivalent to the Sharpe ratio when used in a modelling framework where the underlying return distribution belongs to the family of Q-radial distributions. In brief, the Q-radial distributions are a more general family of multivariate normal distributions which are capable of modelling the heavy tail phenomenon lacking in the normal Gaussian framework (see Chen, He and Zhang (2011) and references therein for a theoretical

[25] Second-degree stochastic dominance, in the context of portfolio theory, refers to the preference of one allocation over another if the former involves less risk and has a mean greater than, or equal to, the latter portfolio.

overview of this class of distributions). Provided Q-radially distributed returns are assumed, equivalence between the Sharpe ratio and the set of lower partial moments-based risk-adjusted performance measures leads Chen, He and Zhang (2011) to deduce that they are in fact monotonic transformations of the Sharpe ratio. Consequently, for returns following a Q-radial distribution, the Sharpe ratio, as a more familiar and less complex risk-adjusted performance measure is 'the' measure to rely on (Chen, He & Zhang (2011)). Proving that these measures are in fact monotonically increasing functions of the Sharpe given the distributional conditions mentioned above, relies on the following Lemma (Chen, He & Zhang (2011)):

Lemma 5.1. *Suppose that $\Psi_n(y)$ is a density function of a probability distribution, and the following function is finite*

$$h_k(z) := \int_{-z}^{\infty} (y+z)^k \Psi_n(y) dy, \quad k \geq 0. \tag{5.1}$$

Then,

$$(h_k)'(z) = k h_{k-1}(z), \tag{5.2}$$

and

$$h_k(z) \geq 0, \quad \forall z \in \mathbb{R}.$$

In other words, $h_k(.)$ is an increasing function, for any $k \geq 0$.

One will notice that $h_k(z)$ in (5.1) is equivalent to the non-normalised numerator of the one-sided variability ratio (4.1) when we set $b = -z$:

$$\begin{aligned} h_k(z) &= \int_{-z}^{\infty} (y+z)^k \Psi_n(y)dy \qquad (5.3)\\ &= \int_{b}^{\infty} (y-b)^k \Psi_n(y)dy \quad b \to -z \\ &= \mathbb{E}[((y-b)^+)^k]. \end{aligned}$$

Common to the analysis carried out in both Eling and Schuhmacher (2012) and Chen, He and Zhang (2011) is the fact that the underlying random returns are symmetrically distributed.

For the benefit of illustrating the differences between various parameterisations of the one-sided variability ratio in a practical setting, we calculated optimal portfolio weights based on employing four parameterisations of (4.1) as objective functions and compared the optimal weights with the Sharpe-optimal portfolios in Chapter 3. The results were tabulated in Table (3.3) and (3.4) for the 'Boom' period and 'Bust' period, respectively, where the underlying investment universe consisted of eighteen FTSE 100 stocks, (see Chapter 1 for further details). This analysis used empirical return data rather than simulating the underlying returns on the basis of assuming a particular distributional family. Here, we carry out the same analysis however, by using the same underlying data to simulate Student t distributed returns where the first two moments were taken from the empirical return dataset for each period and setting the degrees of freedom to be equal to four ($\nu = 4$), we place ourselves in the ES framework (see Tables (3.1) and (3.2) for the first four moments of each asset return for the Boom period and Bust period, respectively). That is, we impose that all asset returns are distributed such that each marginal distribution is a location and scale translation of any of the other seventeen marginal distributions. In carrying out the same optimisation analysis within this framework, we wanted to investigate whether we see any agreement in optimal asset weights and, with particular significance to the papers discussed in this chapter, asset rankings in order to corroborate the theoretical findings in Eling and Schuhmacher (2012).

Optimal asset weights based on optimising all five risk-adjusted performance measures ($\phi(b, 1, 1)$ (Omega), $\phi(b, 1, 2)$ (Sortino), $\phi(b, 1, 3)$, $\phi(b, 1, 4)$ and the Sharpe ratio) are tabulated in the appendix in Tables (9.3) and (9.4). What is of interest here is how each asset ranked when applying the performance measures to their individual returns. In Table (5.1) and (5.2), we present the rank correlations between the various parameterisations of (4.1) and the Sharpe ratio when applied to the Student t simulated asset returns corresponding to each asset.

In Table (5.2), we observe that the Omega ratio ranks assets in the same order as the Sharpe ratio. In Table (5.1), although the rankings are not equal, Omega still exhibits an almost-exact preference for each asset as the Sharpe ratio. Comparing the Sortino ratio and Sharpe ratio, we observe high rank correlations again however, differences in preferences occur. When we set $b = 3$ and $b = 4$ in (4.1), a different picture emerges. We do not see strong agreement in asset preferences. In this sense, we see the benefit of using

Table 5.1: Boom Period: Rank Correlations to the Sharpe ratio in the ES framework.

Performance Measure	Benchmark (% p.a.)	Rank Correlation (%)
Omega	5	99.59
Sortino	5	88.24
$\phi(b, 1, 3)$	5	49.43
$\phi(b, 1, 4)$	5	36.02
Omega	10	99.79
Sortino	10	88.85
$\phi(b, 1, 3)$	10	53.97
$\phi(b, 1, 4)$	10	37.67
Omega	15	99.79
Sortino	15	89.68
$\phi(b, 1, 3)$	15	55.83
$\phi(b, 1, 4)$	15	39.32

Table 5.2: Bust Period: Rank Correlations to the Sharpe ratio in the ES framework.

Performance Measure	Benchmark (% p.a.)	Rank Correlation (%)
Omega	5	100.00
Sortino	5	97.73
$\phi(b, 1, 3)$	5	71.31
$\phi(b, 1, 4)$	5	40.35
Omega	10	100.00
Sortino	10	98.56
$\phi(b, 1, 3)$	10	73.79
$\phi(b, 1, 4)$	10	45.72
Omega	15	100.00
Sortino	15	97.11
$\phi(b, 1, 3)$	15	69.04
$\phi(b, 1, 4)$	15	42.21

alternative risk-adjusted performance ratios, even when the underlying asset returns are homogeneous in distribution. As one would expect, when we turn to using these measures for the purpose of calculating optimal portfolios, discrepancies between Sharpe-optimal portfolios and portfolios optimal in the sense of alternative performance ratios will be more apparent. In practise, when the underlying asset returns exhibit varying distributions, these alternative risk-adjusted performance measures can potentially provide a more informative investigation into individual assets and resulting optimal portfolios.

To enrich this analysis, we apply the above performance ratios to a universe of six assets from non-equity asset classes and calculate optimal portfolios as carried out in the above case. The assets in this scenario include three currencies (EUR, GPM and JPY) and three commodities (Gold, Oil and Silver denominated in USD). The same time periods are covered and we use empirical asset returns for calculating both asset ranks and optimal portfolio weights using, again, $\phi(b, 1, 1)$ (Omega), $\phi(b, 1, 2)$ (Sortino), $\phi(b, 1, 3)$, $\phi(b, 1, 4)$ and the Sharpe ratio. The motivation for carrying out the same analysis in an empirical setting with these underlying assets is to investigate whether we observe strong agreement between any of the four parameterisations of (4.1) when we deviate from the ES framework. In the 18-asset case, applying the Omega ratio to the individual asset returns resulted in a preference ordering which was very highly correlated with the rankings based on the Sharpe ratio when both the empirical and Student t simulated returns were considered. This investigation is similar in spirit to that presented in Kaplan and Knowles (2004) where they explored how certain cases of the Kappa ratio (3.13), a ratio of excess return to the normalised lower partial moment relative to a benchmark, ranks assets when using empirical return data and Johnson SU-return data simulated to match the first four moments of the enpirical returns. Their goal was to investigate whether knowing the first four moments of a distribution is enough to determine Kappa values. In addition, sensitivity of Kappa to skewness and kurtosis was also considered. They found that discrepancies between empirical asset rankings and asset rankings produced from applying Kappa to the simulated data widened for levels of the threshold far below the mean. Similar to their analysis, we consider a negative benchmark. We also set the benchmark equal to zero and the mean of the underlying returns from each time period.

In this 6-asset case, we still see a high correlation in preferences between Omega and Sharpe however, for the Sortino ratio, $\phi(b, 1, 3)$ and $\phi(b, 1, 4)$ asset preferences differ dramatically as can be seen in tables (5.3) and (5.5). How correlated these asset rankings are with the Sharpe-rankings are presented in tables (5.4) and (5.6).

Whilst the framework underlying the latter analysis above, where the universe consisted of non-equity assets, deviated from the ES framework, our purpose here was to provide further evidence that alternative risk-adjusted performance ratios offer the practitioner an enriched set of tools for both ranking assets and calculating optimal portfolios. Even in the ES framework, for higher powers in the one-sided variability ratio, (4.1), asset

Table 5.3: Boom Period: Asset Rankings.

Asset	Benchmark (%)	Sharpe	Omega	Sortino	$\phi(b,1,3)$	$\phi(b,1,4)$
EUR	0	4	4	1	2	2
GBP	0	5	5	2	1	1
JPY	0	6	6	5	4	4
Gold (USD)	0	1	1	4	5	5
Oil	0	2	3	3	3	3
Silver (USD)	0	3	2	6	6	6
EUR	14.2	4	4	4	2	3
GBP	14.2	5	5	5	3	2
JPY	14.2	6	6	6	5	5
Gold (USD)	14.2	3	3	2	4	4
Oil	14.2	2	2	1	1	1
Silver (USD)	14.2	1	1	3	6	6
EUR	-1	4	4	1	2	2
GBP	-1	5	5	2	1	1
JPY	-1	6	6	5	4	4
Gold (USD)	-1	1	1	4	5	5
Oil	-1	2	3	3	3	3
Silver (USD)	-1	3	2	6	6	6

Table 5.4: Boom Period: Rank Correlations.

Performance Ratio	Benchmark (%)	Rank Correlation (%)
Omega	0	94.29
Sortino	0	-8.57
$\phi(b,1,3)$	0	-42.86
$\phi(b,1,4)$	0	-42.86
Omega	14.2	100.00
Sortino	14.2	82.86
$\phi(b,1,3)$	14.2	-2.86
$\phi(b,1,4)$	14.2	-8.57
Omega	-1	94.29
Sortino	-1	-8.57
$\phi(b,1,3)$	-1	-42.86
$\phi(b,1,4)$	-1	-42.86

Table 5.5: Bust Period: Asset Rankings.

Asset	Benchmark (%)	Sharpe	Omega	Sortino	$\phi(b,1,3)$	$\phi(b,1,4)$
EUR	0	4	4	3	2	1
GBP	0	5	5	5	5	4
JPY	0	6	6	6	6	6
Gold (USD)	0	1	1	1	1	2
Oil	0	3	3	4	3	3
Silver (USD)	0	2	2	2	4	5
EUR	10.7	4	4	4	4	2
GBP	10.7	5	5	5	5	5
JPY	10.7	6	6	6	6	6
Gold (USD)	10.7	1	1	1	1	1
Oil	10.7	3	3	3	2	3
Silver (USD)	10.7	2	2	2	3	4
EUR	-1	4	4	3	1	1
GBP	-1	5	5	5	5	4
JPY	-1	6	6	6	6	6
Gold (USD)	-1	1	1	1	2	2
Oil	-1	3	3	4	3	3
Silver (USD)	-1	2	2	2	4	5

Table 5.6: Bust Period: Rank Correlations.

Performance Ratio	Benchmark (%)	Rank Correlation (%)
Omega	0	100.00
Sortino	0	94.29
$\phi(b,1,3)$	0	77.14
$\phi(b,1,4)$	0	42.86
Omega	10.7	100.00
Sortino	10.7	100.00
$\phi(b,1,3)$	10.7	94.29
$\phi(b,1,4)$	10.7	77.14
Omega	-1	100.00
Sortino	-1	94.29
$\phi(b,1,3)$	-1	60.00
$\phi(b,1,4)$	-1	42.86

rankings differed dramatically to those given by the Sharpe ratio. Hence, alternative risk-adjusted performance ratios should not be discounted on the basis that they frequently rank assets in the same order as the Sharpe ratio under certain statistical scenarios. These results serve as a reminder that rankings of assets can change, sometime dramatically, depending on which risk-adjusted performance ratio we employ. As pointed out by Kaplan and Knowles (2004), this is largely due to their varying sensitivity to moments beyond the mean and standard deviation.

In the next chapter, we illustrate a few more cases where discrepancies between Sharpe ratio asset rankings and those calculated via alternative risk-adjusted performance measures arise. We achieve this by considering anisotropic distributions in a bivariate framework. In brief, these are simulated bivariate distributions as described by Shaw and Lee (2008) where inhomogeneous marginals (In our case different degrees of freedom) can be modelled. We describe these in more detail in the subsequent chapter. Accordingly, we show that we do not need to model highly exotic marginal distributions to exhibit situations where asset rankings differ amongst a range of risk-adjusted performance ratios.

Chapter 6

Optimal portfolios with anisotropic bivariate distributions and asymmetric dependency structures

In Chapter 5, we discussed the work of Eling and Schuhmacher (2012) who have neatly formalised what has long been suspected in industry; that ranking assets via the Sharpe ratio is frequently consistent with the equivalent computations in terms of other performance measures. Nevertheless, their theoretical justification of this observed phenomenon is founded on the hypothesis that the statistical characteristics of underlying asset returns satisfy the LS property and that certain properties of the performance measure are preserved. This supposes that all returns can be modelled by the same statistical distribution. When the constituents comprising the investment universe have returns that vary significantly in behaviour, it is erroneous to model their returns with the same distribution. Furthermore, their study ignores the effect of dependence between returns. For the purpose of comparing the individual assets over a certain period of time, selecting one on account of its Sharpe ratio ranking the highest disregards whether its performance was highly correlated with an asset whose Sharpe ratio was relatively low. In addition, in terms of quantifying risk, overlooking the true nature of dependence can result in a misleading distributional structure of the portfolio.

Based on the notion that returns can be modelled by the same distribution, the deductions of Eling and Schuhmacher (2012) may possibly lead readers to suppose that calculating the optimal portfolio by means of maximising the Sharpe ratio is the most favourable methodology or at least sufficient enough to not require the use of more complex ratios (satisfying the ES criteria) on the basis that the asset rankings determined by these performance measures are highly correlated. We stress, however, that their conclusions are not entirely universal and should be taken with caution.

In Chapter 5, we calculated several optimal portfolios and ranked individual assets based on a number of risk-adjusted performance ratios in an empirical and parametric setting where, in the latter case, we supposed that all marginal distributions satisfied the location-scale property. The purpose of this was to investigate whether, in practise, sorting and weighting assets via the Sharpe ratio provided highly correlated, if not the same, results to those derived from ranking assets and optimising portfolios in terms of other risk-adjusted performance measures. In our case, we considered four parameterisations of (4.1); $\phi(b, 1, 1)$ (Omega Index), $\phi(b, 1, 2)$ (Sortino ratio), $\phi(b, 1, 3)$ and $\phi(b, 1, 4)$. All of these belong to the measures satisfying the ES criteria in that they satisy positive homogeneity and translation invariance as described in Eling and Schuhmacher (2012). In both an empirical and parametric setting, we observed considerable differences in optimal portfolio weights. In agreement with the work in Eling and Schuhmacher (2012), we showed that, for low integer powers of the denominator of (4.1), asset rankings were highly correlated with those calculated via the Sharpe ratio. This was true when we used both empirical return data and simulated return data which was modelled to satisfy the location-scale property. For higher powers of the denominator of (4.1) however, asset rankings were far from perfectly positively correlated in all cases.

For the benefit of highlighting the capabilities of employing a wider range of risk-return measures in a portfolio context, we also calculated optimal portfolio weights and ranked assets according to the Sharpe ratio, the Omega Index and Sortino ratio when our underlying index consisted of non-equity assets. In this chapter we present a few more cases to exhibit scenarios under which asset rankings dictated by the Sharpe ratio, Omega Index and Sortino ratio do not necessarily agree. We also show that Sharpe-optimal portfolios differ considerably to portfolios optimal in terms on the Omega Index and Sortino ratio.

In these case studies, we consider i) anisotropic bivariate distributions and ii) asymmetric dependency structures for the purpose of asset ranking and optimal asset allocation. The term anisotropic refers, in this situation, to distributions whose properties differ beyond mean and variance and can be considered as the antithesis to the location-scale property described in Eling and Schuhmacher (2012). Traditionally, many multi-dimensional asset models have used simple multivariate distributions such as the Gaussian and the standard multivariate t. These have the appeal that they are tractable. However, they tend to be very isotropic in character after variance scaling. So for example, in the standard multivariate t, all dimensions have the same degrees of freedom and power law in the tails. However, we might wish to allow for different assets to have differing marginal distributions. Some might be close to Gaussian, others might be heavy tailed. This is what I mean by an anisotropic distribution.

The analysis carried out in the first part of this chapter below is based on the work of Shaw and Lee (2008). In their study, they proposed a method for generating anisotropic bivariate Student t distributions with canonical analytical functions. The motivation for carrying out this research was to highlight the ease with which one can produce anisotropic bivariate distributions with existing simulation techniques. Shaw and Lee (2008) show how to construct bivariate Student t distributions via techniques used in the simulation of t-copulas. In doing so, they propose a class of multivariate Student t distributions which allow for unequal degrees of freedom containing both the independent and dependent case, a property not satisfied in the common multivariate representation of the Student t distribution which is, as shown in Shaw and Lee (2008):

$$g_\nu(t_1, t_2, \ldots, t_d) \equiv g_\nu(\underline{t}) = \frac{\sqrt{|A|}}{\sqrt{\nu^d \pi^d}} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma\frac{\nu}{2}} \left(1 + \frac{\underline{t}^T A \underline{t}}{\nu}\right)^{-(\nu+d)/2}. \tag{6.1}$$

Various analytical examples of anisotropic Student t bivariate distributions have been given by Shaw and Lee (2008), where an explicit mechanism is given for varying the degrees of freedom between the marginals. Having the flexibility to model a set of returns with varying degrees of freedom is of interest in multiple disciplines, in addition to mathematical finance and has obvious practical appeal. Focusing on the bivariate case, Shaw and Lee (2008) are able to explicitly characterise a canonical bivariate distribution for the Student-normal case with a "naturally associated copula" and density function. For the more general Student - Student case, there also exists a canonical representation of the bivariate distribution with certain restrictions in order to be able to extract the case of zero correlation. We employ these in two of our case studies further on in this chapter.

The theory underpinning Shaw and Lee's (2008) paper relies on the following representation of a Student t random variable

$$T = \frac{Z}{\sqrt{C^2/a}}, \tag{6.2}$$

where Z is normally distributed and C^2, independent of Z, has a χ^2 distribution with a degrees of freedom.[26] From this representation, it is an easy task to arrive at the univariate density for the Student t:

$$f_a(t) = \frac{1}{\sqrt{a\pi}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \frac{1}{(1 + t^2/a)^{\frac{a+1}{2}}}. \tag{6.3}$$

In terms of characterising the linkage between the two distributions, a parameter, θ, is introduced. This can be considered as a rotation angle corresponding to the relationship between the marginals. The correlation is then defined in terms of θ.

Since we will be making use of the Student-normal and Student-Student bivariate anisotropic distributions in what follows, we describe how these are formulated in sections (6.1) and (6.2) below.

6.1 The Student-normal case

For the case where we want one Student t marginal (T_1) and one Gaussian marginal (T_2), our bivariate anistropic distribution can be characterised as follows:

$$T_1 = \frac{Z_1}{\sqrt{C_1^2/a}}, \quad T_2 = Z_1 \sin(\theta) + Z_2 \cos(\theta), \tag{6.4}$$

where Z_1 and Z_2 are two independent normal random variables and C is as before. By rearranging the above expression in terms of the normal variables, Z_1 and Z_2, and using the standard normal density, Shaw and Lee (2008) show that the bivariate density

[26] Note that a does not need to be an integer (Shaw & Lee, 2008).

function is,

$$f(t_1,t_2) = \frac{e^{\gamma^2/2-(1/2)\sec^2(\theta)t_2^2}\alpha^{-a/2-1/2}}{\sqrt{2a}\Gamma(\frac{a}{2})\pi\cos(\theta)} \times \left\{\sqrt{2}\gamma\Gamma\left(\frac{a}{2}+1\right){}_1F_1\left(\frac{1}{2}-\frac{a}{2};\frac{3}{2};-\frac{\gamma^2}{2}\right)+\Gamma\left(\frac{a+1}{2}\right){}_1F_1\left(-\frac{a}{2};\frac{1}{2};-\frac{\gamma^2}{2}\right)\right\}, \tag{6.5}$$

where f depends on t_1 and t_2 through α, γ and β which are given in Shaw and Lee (2008) as below:

$$\alpha = 1+\frac{t_1^2}{a\cos^2(\theta)}, \quad \beta = \frac{t_1 t_2 \sin(\theta)}{\sqrt{a}\cos^2(\theta)}, \quad \gamma = \frac{\beta}{\sqrt{\alpha}}. \tag{6.6}$$

The correlation in this case is given by

$$\rho = \sin(\theta)\frac{\Gamma(\frac{a-1}{2})}{\Gamma(\frac{a}{2})}\sqrt{\frac{a}{2}-1}. \tag{6.7}$$

Several density contour plots of the Student-normal case are shown in Figure (6.1). These plots represent the bivariate Student-normal anisotropic densities where a takes values $a \in \{1,2,4,20\}$ and $\theta \in \{0,-\pi/4,\pi/4\}$ in (6.5).

6.2 The Student-Student case

For the Student-Student case, the marginals are formulated as follows:

$$T_1 = \frac{Z_1}{\sqrt{C_1^2/a}}, \quad T_2 = \sqrt{\frac{a+b}{C_1^2+C_2^2}}[Z_1\sin(\theta)+Z_2\cos(\theta)], \tag{6.8}$$

where a and b are the degrees of freedom of T_1 and T_2, respectively, Z_1 and Z_2 are as before and C_1^2 and C_2^2 are χ^2 distributed which, in this case, are coupled. There appears

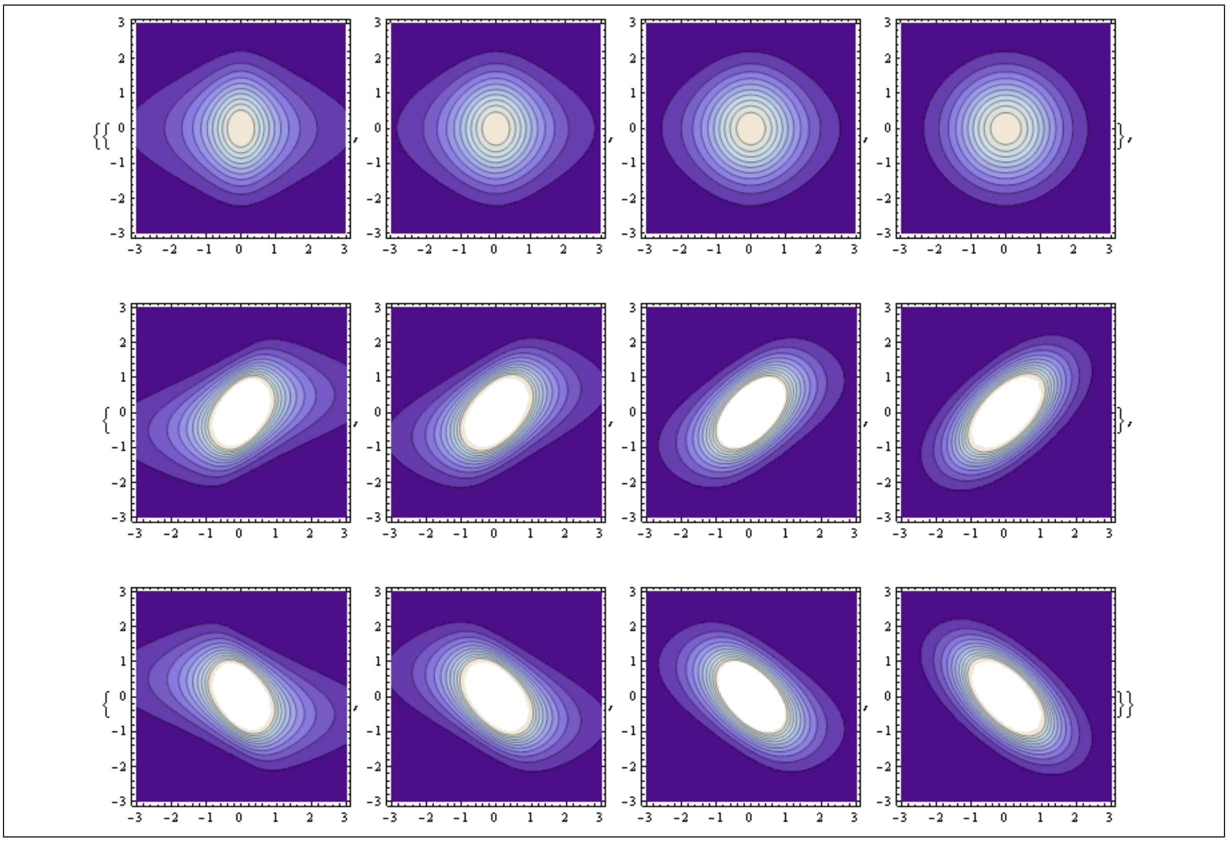

Figure 6.1: *Contour Plots of Student-normal anisotropic bivariate densities given by (6.5) for $a \in \{1, 2, 4, 20\}$ (columns from left to right) and $\theta \in \{0, -\pi/4, \pi/4\}$ (rows from top to bottom).*

to be no closed form expression for the density in this case. Shaw and Lee (2008) did however explicitly formulate the correlation for the Student-Student anisotropic case as follows:

$$\rho = \sin(\theta)\frac{\Gamma(\frac{a-1}{2})}{\Gamma(\frac{a}{2})}\sqrt{\frac{a}{2}-1}\frac{\Gamma(\frac{a+b-2}{2})}{\Gamma(\frac{a+b-1}{2})}\sqrt{\frac{a+b}{2}-1}. \tag{6.9}$$

6.3 Testing for agreement with Sharpe ratio asset rankings

We consider two cases where a bivariate asset return distribution can be modelled by one of Shaw and Lee's (2008) anisotropic bivariate distributions and investigate whether asset rankings agree when various risk-adjusted performance ratios are applied to the marginals for various levels of the benchmark. We also calculate optimal portfolio weights based on these same performance ratios and benchmark values. Our first case considers a Student-normal bivariate anisotropic distribution. We then simulate an anisotropic bivariate Student-Student distribution. Both these cases require employing simulation techniques suggested in Shaw and Lee (2008). To enrich this analysis further, we consider a third case; a bivariate anisotropic distribution coupled with an exotic dependency structure. In this instance, we suppose that the dependency structure can be modelled by a Crash Clayton copula.

6.3.1 The anisotropic Student-normal bivariate distribution

For this case, we suppose a bivariate return dataset can be modelled by the Student-normal case described in section (6.1). In this hypothetical case, we suppose the Gaussian marginal (T_2) has mean 4% and standard deviation 10% and the Student t marginal (T_1) has mean 8% and standard deviation 20% with 3 degrees of freedom. We simulate these marginals based on (6.4) above. The correlation coefficient between T_1 and T_2 is then determined by the value we assign to θ in the following expression for ρ in (6.7):

$$\rho = \sin(\theta)\frac{\Gamma(\frac{a-1}{2})}{\Gamma(\frac{a}{2})}\sqrt{\frac{a}{2}-1} = \sin(\theta)\frac{\Gamma(1)}{\sqrt{2}\Gamma(3/2)}. \tag{6.10}$$

Based on this parametrisation, we calculate optimal weights based on maximising the Sharpe ratio, the Omega index and Sortino ratio when we set the correlation to 0 and 0.5. Here, we find that even when marginal densities are elliptical and we impose one Gaussian return, our optimal portfolios and asset rankings for the two levels of correlation (by varying θ) do not agree across all values of the benchmark.

The reader can consult tables (**??**) and (**??**) for optimal weights when $\rho = 0$ and $\rho = 0.5$. Of particular interest are the values we obtain from applying the three risk-adjusted performance ratios to the individual marginals. In the zero correlation case in Table (**??**), one will observe that Sortino ranks assets in the opposite order to the Sharpe ratio and Omega Index when $b = 1\%$. Disagreement in asset preferences are also seen in the case where $b = 0\%$ and $b = 1\%$. A similar picture emerges in the $\rho = 0.5$ case where optimal asset weights also differ significantly.

6.3.2 The anisotropic Student-Student bivariate distribution

Based on the Student-Student anisotropic model given by (6.8) in section (6.2), we give an example where discrepancy between the Sharpe-, Omega- and Sortino-optimal portfolios

b (%)	Performance Ratio	T1	T2
-1%	Sharpe	68.646%	31.354%
-1%	Omega	60.258%	39.742%
-1%	Sortino	76.766%	23.234%
0%	Sharpe	66.322%	33.678%
0%	Omega	56.280%	43.720%
0%	Sortino	75.014%	24.986%
1%	Sharpe	62.803%	37.197%
1%	Omega	51.271%	48.729%
1%	Sortino	72.907%	27.093%
2%	Sharpe	56.795%	43.205%
2%	Omega	43.609%	56.391%
2%	Sortino	70.145%	29.855%
3%	Sharpe	44.241%	55.759%
3%	Omega	30.742%	69.258%
3%	Sortino	65.997%	34.003%

Table 6.1: *Optimal Portfolio weights for the Student-Normal Anisotropic Bivariate case with zero correlation.*

b (%)	Performance Ratio	T1	T2
-1%	Sharpe	1st	2nd
-1%	Omega	2nd	1st
-1%	Sortino	1st	2nd
0%	Sharpe	1st	2nd
0%	Omega	2nd	1st
0%	Sortino	1st	2nd
1%	Sharpe	2nd	1st
1%	Omega	2nd	1st
1%	Sortino	1st	2nd
2%	Sharpe	2nd	1st
2%	Omega	2nd	1st
2%	Sortino	2nd	1st
3%	Sharpe	2nd	1st
3%	Omega	2nd	1st
3%	Sortino	2nd	1st

Table 6.2: *Individual asset rankings in the Student-Normal Anisotropic Bivariate case with zero correlation.*

b (%)	Performance Ratio	T1	T2
-1%	Sharpe	72.362%	27.638%
-1%	Omega	54.500%	45.500%
-1%	Sortino	54.498%	45.502%
0%	Sharpe	65.434%	34.566%
0%	Omega	40.103%	59.897%
0%	Sortino	40.102%	59.898%
1%	Sharpe	53.873%	46.127%
1%	Omega	19.302%	80.698%
1%	Sortino	19.303%	80.697%
2%	Sharpe	30.673%	69.327%
2%	Omega	0.000%	100.000%
2%	Sortino	0.000%	100.000%
3%	Sharpe	0.000%	100.000%
3%	Omega	0.000%	100.000%
3%	Sortino	0.000%	100.000%

Table 6.3: *Optimal Portfolio weights for the Student-Normal Anisotropic Bivariate case with positive (*0.5 *correlation.*

b (%)	Performance Ratio	T1	T2
-1%	Sharpe	1st	2nd
-1%	Omega	2nd	1st
-1%	Sortino	1st	2nd
0%	Sharpe	2nd	1st
0%	Omega	2nd	1st
0%	Sortino	1st	2nd
1%	Sharpe	2nd	1st
1%	Omega	2nd	1st
1%	Sortino	1st	2nd
2%	Sharpe	2nd	1st
2%	Omega	2nd	1st
2%	Sortino	2nd	1st
3%	Sharpe	2nd	1st
3%	Omega	2nd	1st
3%	Sortino	2nd	1st

Table 6.4: *Individual asset rankings in the Student-Normal Anisotropic Bivariate case with positive (0.5) correlation.*

and asset rankings exists. In this hypothetical situation, we assume that one marginal (T_1) has mean 8%, standard deviation 20% and 3 degrees of freedom ($\nu = 3$) and the second marginal (T_2) has mean 4% and standard deviation 10% with 10 degrees of freedom ($\nu = 10$). These are given in Table (6.5).

	Mean	Standard Deviation	ν
T_1	8%	20%	3
T_2	4%	10%	10

Table 6.5: *Student-student Case.*

We calculate optimal portfolio weights for these two assets as well as rank the assets for various benchmark values, performance ratios (specifically the Sharpe ratio, Omega Index and Sortino ratio) and parametrisations of θ. In particular, for $\rho = 0.28$, we observe inconsistencies in both allocations and rankings.

As can be seen in Table (**??**), for $b = 1\%$, whilst Sharpe- and Sortino optimal weights broadly agree, the Omega-optimal weights differ by over 10%. Similarly for other values of the benchmark. In particular, Omega-optimal portfolios focus on one asset for lower levels of the benchmark than the Sharpe and Sortino equivalent weights.

If we now consult Table (6.7) containing asset rankings, although asset rankings agree for three out of four levels of the benchmark, we see that, for $b = 0\%$, employing

b (%)	Performance Measure	T1 (%)	T2 (%)
1	Sharpe	39.1	60.9
1	Omega	55.3	44.7
1	Sortino	31.8	68.2
2	Sharpe	51.2	48.8
2	Omega	75.2	24.8
2	Sortino	36.9	63.1
3	Sharpe	80.2	19.8
3	Omega	100	0
3	Sortino	45.1	54.9
4	Sharpe	100	0
4	Omega	100	0
4	Sortino	64.9	35.1

Table 6.6: *Optimal Portfolio Weights for the Student-Student Anisotropic Case.*

the Omega ratio to sort assets results in a different order to those given by the Sharpe ratio and Sortino Index. In light of the results presented for the optimal weights, it is apparent that asset rankings do not explain the full picture. Furthermore, in situations where our asset universe consists of more than two stocks (a likely and more realistic scenario), we may observe more inconsistencies than we observe for this illustrative case.

These results indicate that we don't have to consider extreme alternative distributional models to find situations where the conclusions in Eling and Schuhmacher (2012) don't ring true anymore. This is just one example based on one particular parametrisation

	Omega		Sortino		Sharpe	
b	T_1	T_2	T_1	T_2	T_1	T_2
0%	2nd	1st	1st	2nd	1st	2nd
1%	1st	2nd	1st	2nd	1st	2nd
2%	1st	2nd	1st	2nd	1st	2nd
3%	1st	2nd	1st	2nd	1st	2nd

Table 6.7: *Asset Rankings: Student-Student Case.*

of the bivariate family defined above.

6.3.3 LS marginals coupled with a Clayton copula

A limitation of the theory set out in Eling and Schuhmacher (2012) is that they ignore dependency. To add to the previous analysis and to conclude this section, we investigate the impact that a less conventional dependency structures has on the ranking of assets and consequent optimal portfolios. As before, we consider two assets which are both Student t distributed. However, we now suppose that their dependency structure can be modelled by a Clayton Crash copula. In this framework, we test whether consistency between Sharpe- and Omega-optimal portfolios prevails.

In Chapter 2, we gave a brief overview of copulas. Here, we focus on the Clayton copula as a model of dependency, a family of dependency structures belonging to the asymmetric Archimedean copula group. Its profile exhibits a higher concentration of dependency on the downside than the upside, hence it is often referred to as the Clayton Crash copula.

We briefly introduced the Archimedean copula family in Chapter 2. In the bivariate case, an Archimedean copula function is given by

$$C(u,v) = \psi^{-1}(\psi(u) + \psi(v)). \tag{6.11}$$

For the Clayton copula, the generator function $\psi(.)$ takes the form:

$$\psi(t) = (t^{-\theta} - 1). \tag{6.12}$$

Therefore, the inverse is given by

$$\psi^{-1}(t) = (1+t)^{-1/\theta}, \tag{6.13}$$

and the copula is defined as

$$C_\alpha(u,v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0), \quad \theta \in [-1,\infty) \setminus \{0\}. \tag{6.14}$$

For most applications, $\theta > 0$, therefore $C_\alpha(u,v)$ reduces to

$$C_\alpha(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\alpha} \quad \theta \in (0,\infty). \tag{6.15}$$

The parameter, θ, governs the strength of negative dependency between the underlying marginals. Its corresponding probability density function is then given by

$$c(u,v) = (1+\theta)(uv)^{(\theta+1)}(u^{-\theta} + v^{-\theta} - 1)^{-(2\theta+1)/\theta}, \tag{6.16}$$

and the relationship between the dependency parameter, θ, and Kendall's tau, τ, is as follows:

$$\tau(\theta) = \theta/(\theta + 2). \tag{6.17}$$

We simulate Student t returns and model their dependency structure with a Clayton copula. We then calculate optimal Sharpe- and Omega portfolios for different values of the benchmark.

Before we proceed, we briefly outline the necessary steps required to simulate Clayton random variables. The principal objective in copula simulation is to simulate uniformly distributed random variables from the unit hypercube with the desired dependency structure. Generating $\sim U[0, 1]$ random variables is fundamental for simulating other number sequences and is a huge subject in its own right. Leaving aside the numerous historical approaches for such simulations,[27] several numerical and arithmetic methods exist and the reader should consult a dedicated source for a thorough grounding in this field.

Assuming we have a uniform generator with all the desirable properties,[28] simulating

[27]Such approaches include throwing dice, dealing cards and methods based on electric circuits or π.

[28]A good uniform random number generator should produce a sequence of numbers which behave as the realisation of a set of independent and identically uniformly distributed (i.i.d) numbers over the unit interval. For practical purposes, the generator should be able to produce a sequence of numbers, as described above, quickly without demanding vast amounts of storage. Nowadays, most programming languages have an in-built system for producing uniform random variables as they form the basis for simulating numerous other random variable sequences. For example, the Box-Muller transform, or its polar variation, can be used to generate two independent normal random variables in its original form,

a copula amounts to carrying out a transformation on these uniformly distributed random variables. The form of the transformation dictates the characteristics of the dependency structure. Hence, to simulate Clayton random variables, we apply the Clayton generator function to our uniform random variates. More specifically, Aas (2004) suggests a method which involves making use of the fact that the inverse of the Clayton generator, (6.13), is equal to the Laplace transform of a Gamma variate. For this reason, a procedure for generating Clayton random variables entails first simulating a Gamma variate ($X \sim Gamma(1/\theta, 1)$), then simulate n independent uniform random variables $V_1, V_2, \ldots, V_n$. Generating Clayton random variables results from transforming these uniform random variables and the Gamma variate, X,[29]

$$\mathbf{U} = ((1 - \tfrac{\log V_1}{X})^{1/\theta}, (1 - \tfrac{\log V_2}{X})^{1/\theta}, \ldots, (1 - \tfrac{\log V_n}{X})^{1/\theta}). \tag{6.18}$$

A visualisation of various bivariate Clayton plots are shown in Figure (6.2).

For our final analysis, we calculate optimal portfolio weights when employing a Clayton crash dependency model and we impose the condition that marginal densities satisfy the location-scale property. In this case, we suppose that two marginals are

and in Bailey's variation of it, single Student t variates (Bailey, 1994).

[29]As described in Aas (2004), this method is derived from the fact that, if ψ is the inverse of the Laplace transform of a distribution function, G, on $\mathbb{R}^+$ with $G(0) = 0$ then an Archimedean copula can be simulated by first simulating a random variable from the distribution, G as well as simulating (independently) n uniform variates and then applying them to the expression, $\underline{U} = (\phi^{-1}(-\log(V_1)/X), \ldots, \phi^{-1}(-\log(V_d)/X))$.

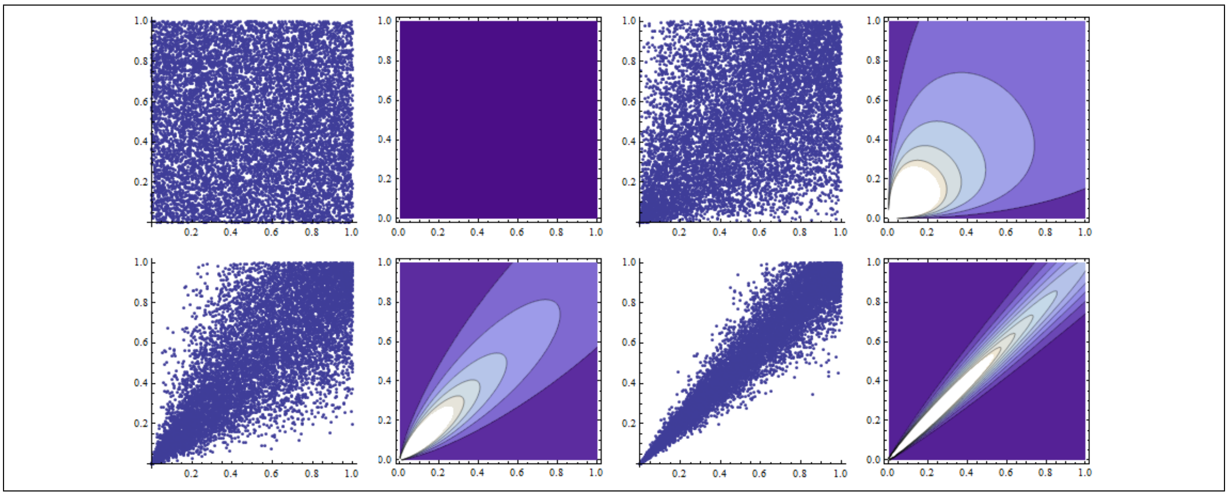

Figure 6.2: *Clayton Bivariate Dependency Plots.*

distributed according to the Student t model with means, standard deviations and degrees of freedom parameters tabulated in Table (6.8). We begin by simulating $10,000$ samples ($20,000$ for the two assets) from a Clayton Crash model, with dependency parameter set equal to 4 ($\alpha = 4$ in (6.15))[30]. Due to Sklar's theorem, the dependency structure and the underlying marginal characteristics can be separated completely so that once we have modelled the dependency structure all that is left is for us to feed these Clayton-dependent uniforms through the quantile functions of the marginal distributions. For calculating optimal portfolios, we can now feed the simulated returns through the respective objective function (see Chapter 1 for a description of the optimisation process). We run three optimisations for various benchmark levels for maximising Omega, Sortino and the Sharpe ratio. The optimal portfolio weights are given in Table (6.9). We also calculate the order in which each performance measure ranks each asset. Interestingly, all asset rankings

[30] This is purely hypothetical, in general the lower α the more dependent the assets are on the downside.

agree in this case as can be seen in Table (6.10).

	Mean	Standard Deviation	d.o.f
Asset 1	9%	24%	3
Asset 2	5%	9%	3

Table 6.8: *Asymmetric Dependency coupled with LS-marginals.*

	Sharpe		Omega		Sortino	
b	T_1	T_2	T_1	T_2	T_1	T_2
1%	2%	98%	0%	100%	0%	100%
2%	13%	87%	0%	100%	0%	100%
3%	46%	54%	100%	0%	100%	0%
4%	100%	0%	100%	0%	100%	0%

Table 6.9: *Optimal portfolio weights: LS marginals with asymmetric dependency.*

In this chapter, we set out to provide a number of analyses to check whether the findings in Eling and Schuhmacher (2012) are true when we relax one or a few of the conditions that they impose. In the first two cases, we employed Shaw and Lee's (2008) method of simulating anisotropic bivariate distributions for modelling two asset returns. In both the Student-normal case and the Student-Student case, we found that when we ranked the assets according to the Sharpe ratio, Omega Index and Sortino ratio, we did

	Sharpe ratio		Omega Index		Sortino ratio	
b (%)	T1	T2	T1	T2	T1	T2
-1	2nd	1st	2nd	1st	2nd	1st
0	2nd	1st	2nd	1st	2nd	1st
1	2nd	1st	2nd	1st	2nd	1st
2	2nd	1st	2nd	1st	2nd	1st
3	1st	2nd	1st	2nd	1st	2nd
4	2nd	1st	2nd	1st	2nd	1st
5	2nd	1st	2nd	1st	2nd	1st

Table 6.10: *Asset Rankings: LS marginals combined with a Clayton crash copula.*

not observe a universal agreement in preferences. We also calculated optimal portfolio weights by maximising these risk-adjusted performance measures, all of which belong to the family of measures satisfying the conditions set out in Eling and Schuhmacher (2012). In the latter example presented in this chapter, we considered the case where two marginals satisfied the LS property however, we supposed that their dependency structure could be modelled by the Clayton crash copula. In this instance, we did in fact find 100% correlation between asset rankings despite the fact that we observed rather different optimal portfolio weights. In this case, we verified the conclusions laid out in Eling ad Schuhmacher (2012). However, it would be interesting to see if this is still true when we carry out the same investigations on a much larger or diverse universe of assets or a wider variety of dependence structures.

Chapter 7

Conclusion

For applications in portfolio optimisation, we advocate a broader use of alternative performance and risk measures. Notwithstanding the large number of quantitative tools available, traditional methods and measures continue to be used despite their well-known drawbacks. The focus of this thesis was to take a closer look at alternative reward-to-risk measures such as those belonging to the class of one sided variability ratios, $\phi(b, p, q)$, proposed by Farinelli and Tibiletti (2002). Concentrating on the Omega Index and Sortino ratio, we showed how these measures can be reduced to analytical functions of the Sharpe ratio based on Student t returns. In a more general framework, Eling and Schuhmacher (2012) proved that reward-to-risk measures satisfying positive homogeneity and translation invariance are monotonically increasing functions of the Sharpe ratio conditional on the location-scale property being satisfied.

The implications of the analytic representations of Omega and Sortino in Chapter 4 as well as the conclusions in Eling and Schuhmacher (2012) discussed in Chapter 5 suggest that many supposed mathematically superior and more complex performance ratios achieve the same result as the universally familiar Sharpe ratio when used for comparing alternative investments. In the latter part of this thesis, however, we provide situations outside of the framework of Eling and Schuhmacher (2012) where portfolio rankings and optimal portfolio weights based on the Sharpe ratio, Omega Index and Sortino ratio yield different results. On the other hand, when we considered the case where we impose the LS condition on two marginal return distributions yet incorporate an asymmetric dependency structure, in this case the Clayton copula, our findings coincided with those of Eling ad Schuhmacher (2012).

The motivation of this thesis was to explore a richer class of risk-adjusted performance ratios capable of incorporating all the information regarding the return distributions. Whilst the results in Eling and Schuhmacher (2012) theoretically show that the Sharpe ratio suffices as a means of comparing alternative investments under certain conditions, we wish to emphasise that there will be situations where a more general measure capable of capturing a wide range of distributional shapes should yield more realistic results than those measures that rely on a small number of statistical characteristics of the underlying dataset.

The author acknowledges that there are clear limitations contained in the analyses presented in this thesis. We centered the analytics in Chapter 4 on the assumption that returns can be modelled by the Student t model. We should consider whether a similar analysis can be achieved based on alternative distributions, preferably those capable of modelling additional moments. We have already mentioned that Passow (2005) claims to have analytically reduced the Omega Index when the underlying return is modelled by the Johnson SU distribution. Furthermore, investigating the merits of various parameterisations of the one-sided variability measure in an optimisation setting over multiple time periods is also of interest.

To conclude this thesis, our general philosophy is that one should not base a portfolio optimisation model on one single measure. Rather, a number of performance and risk measures should be considered both ex ante and ex post for selecting the optimal portfolio. As Sortino and Forsey (1996) summarises, 'No one risk measure is the be-all and end-all.'

Chapter 8

References

Bibliography

[1] Aas, K. 2004. Modelling the dependence structure of financial assets: A survey of four copulas. [Online]. Norwegian Computing Centre, report SAMBA/22/04. Available from: `http://www.nr.no/files/samba/bff/SAMBA2204c.pdf`.

[2] Artzner, P., Delbaen,F., Eber, J.M., and Heath, D. 2004. Coherent measures of risk. Mathematical Finance, Vol. 9, pp. 203-228.

[3] Avouyi-Dovi, S., Morin, S. and Neto, D. 2004. Optimal asset allocation with omega function. [Technical report]. Banque de France.

[4] Bailey, R.W. 1994. Polar Generation of Random Variates with the t-Distribution. Mathematics of Computation, Vol. 62, No. 206, pp. 779-781.

[5] Benartzi, S., Thaler, R.H. 2001. Naive diversification strategies in defined contribution saving plans. The American Economic Review, Vol.91, No. 1, pp. 19-98.

[6] Blatberg, R.C., and Gonedes, N.J. 1974. A comparison of the stable and student distributions as statistical models for stock prices. The Journal of Business, Vol. 47, No. 2, pp. 244-280.

[7] Breymann, W., Luthi, D. and Platen, E. 2009. Empirical behaviour of a world stock index from intra-day to monthly time scales. The European Physical Journal B, Vol. 71, No. 4.

[8] Chaudry, A., Johnson, H.L. 2008. The efficacy of the Sortino ratio and other benchmarked performance meaures under skewed return distributions. Australian Journal of Management, Vol. 32, No. 3, pp. 485-502.

[9] Chen, L., He, S. and Zhang, S. 2011. When all risk-adjusted performance measures are the same: In praise of the Sharpe ratio. Quantitative Finance, Vol. 11, No. 10.

[10] Choi, P. 2001. Estimation of value at risk using Johnson SU-normal distribution. [Online]. Available from: `http://gloria-mundi.com/library_journal_view.asp?journal_id=5574`.

[11] Cogneau, P., Hubner, G. 2009. 101 Ways to Measure Portfolio Performance. [Online]. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract\textunderscore_id=1326076.

[12] Cont, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, Vol. 1, pp. 223-236.

[13] Daul, S., Giorgi, E., Lindskog, F., McNeil, A. 2003. The grouped t-copula with an application to credit risk. [Online]. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract\textunderscoreid=1358956.

[14] Demarta, S., McNeil, A.J. 2005. The t copula and related copulas. International Statistical Review, Vol. 73, No. 1, pp. 111-129.

[15] Devroye, L. 1986. Non-Uniform Random Variate Generation. Springer-Verlag, New York. [Online]. Available from: http://luc.devroye.org/rnbookindex.html.

[16] DiMiguel, V., Garlappi, L., Uppal, R. 2009. Optimal versus naive diversification: how efficient is the 1/N portfolio strategy? The Review of Financial Studies, Vol. 22, No. 5, pp. 1915-1953.

[17] Durrleman, V., Nikeghbali, A., Roncalli, T. 2000. Which copula is the right one? [Online]. Available from: http://www.thierry-roncalli.com/download/copula-choice.pdf.

[18] Eling, M., and Schuhmacher, F. 2012. A decision-theoretic foundation for reward-to-risk performance measures. Journal of Banking and Finance, Vol. 36, No.7, pp. 2077-2082.

[19] Eling, M., and Schuhmacher, F. 2005. Does the choice of performance measure influence the evaluation of hedge funds. Journal of Banking and Finance, Vol. 31, No. 9, pp 2632-2647.

[20] Eling, M., and Schuhmacher, F. 2007. Sufficient conditions for expected utility to imply drawdown-based measures. Journal of Banking and Finance, Vol. 35,No. 9, pp. 2311-2318.

[21] Fama, E. 1965. The behaviour of stock prices. The Journal of Business, Vol. 38, No. 1, pp 34-105.

[22] Farinelli, S. and Tibiletti, L. 2002. Sharpe thinking with asymmetrical preferences. [Online].Available from: `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=338380`.

[23] Farinelli, S. and Tibiletti, L. 2006. Sharpe thinking in asset ranking with one-sided measures. Europena Journal of Operational Research, Volume, 185, No.3, pp. 1542-1547.

[24] Farinelli, S., Ferreira, M., Theone, M., Rosello, D. and Tibiletti, L. 2006. Optimal asset allocation aid system: from "one-size" vs "tailor-made" per-

formance ratio. [Online]. Available from: `http://sunshine.dma.unive.it/QuantitativeFinance2007/viewpaper.php?id=93`.

[25] Farinelli, S., Ferreira, M., Rosello, D.,Theone, M., and Tibiletti, L. 2008. Beyond Sharpe ratio: optimal asset allocation using different performance ratios. Journal of Banking & Finance, Vol. 32, No. 10. pp. 2057-2063.

[26] Feldstein,M.S. 2006. Mean-variance analysis in the theory of liquidity preference and portfolio selection. The Review of Economic Studies, Vol. 36, No. 1, pp. 5-12.

[27] Fergusson, K., Platen, E. 2006. On the distributional characterization of log-returns of a world stock index. Applied Mathematical Finance, Vol.13, No. 1, pp 19-28.

[28] Rom, B.M., Ferguson, K. 1994. Post-modern portfolio theory comes of age. The Journal of Investing, Vol. 3, No. 3, pp. 11-17.

[29] Fishburn, P.C. 1977. Mean-risk analysis with risk associated with below-target returns. The American Economic Review, Vol. 67, No. 2.

[30] Hall, T., Satchell, S. 2010. Computing optimal mean/downside risk frontiers: the role of ellipticity. In: Optimizing optimization, the next generation of optimization applications and theory. Quantitative Finance Series.

[31] Harlow, W.V. 1991. Asset allocation in a downside-risk framework. Financial Analysts Journal, Vol. 47, No. 5.

[32] Harlow, W.V., Rao, K.S.Ramesh. 1989. Asset pricing in a generalized mean-lower partial moment framework: theory and evidence. The Journal of Financial and Quantitative Analysis, Vol. 24, No. 3.

[33] Hoffman, S.T., Rollinger, T.N. 2013. Sortino: A 'Sharper' Ratio. [Online]. Available from: `http://www.redrockcapital.com/assets/RedRock_Sortino_white_paper.pdf`.

[34] Johnson, N.L. 1949. Systems of frequency curves generated by methods of translation. Biometrika, Vol. 36, No. 1/2, pp. 149-176.

[35] Jondeau, E., Poon, S., Rockinger, M. 2007. Financial modeling under non-gaussian distributions. 1st Ed. London: Springer Finance.

[36] Kane, S.J., Batholomew-Biggs, M.C., Cross, M., Dewar, M. 2007. Optimizing omega. The Journal of Global Optimization, Vol. 45, No. 1.

[37] Kaplan, P.D., Knowles, J.A. 2004. Kappa: A generalized downside risk-adjusted performance measure. Journal of Performance Measurement, (2004).

[38] Keating, C. and Shadwick, F. 2000. An introduction to omega. The Finance Development Centre. [Online]. Available from: `https://faculty.fuqua.duke.edu/~charvey/Teaching/BA453_2005/Keating_An_introduction_to.pdf`.

[39] Keating, C. and Shadwick, F. 2002. A universal performance measure. The Finance Development Centre. [Online]. Available from: https://faculty.fuqua.duke.edu/~charvey/Teaching/BA453_2004/Keating_A_universal_performance.pdf.

[40] Lewis, A.L. 1990. Semivariance and the performance of portfolios with options. Financial Analysts Journal, Vol. 46, No. 4, pp. 67-76.

[41] Leyland, H.E. 1999. Beyond mean-variance: performance measurement in a nonsymmetrical world. Financial Analysts Journal, Vol. 55, pp. 27-36.

[42] Maillard, S., Roncalli, T., Teiletche, J. 2010. On the properties of equally-weighted risk contribution portfolios. The Journal of Portfolio Management, Vol. 36, No. 4, pp. 60-70.

[43] Mandelbrot, B. 1963. The variation of certain speculative prices. The Journal of Business, Vol. 36, No. 4, pp. 394-419.

[44] Markowitz, H.M. 1952. Portfolio selection. The Journal of Finance, Vol. 7, No. 1, pp. 77-91.

[45] Markowitz, H.M. 1956. The optimization of a quadratic function subject to linear constraints. Naval Research Logistics Quarterly, Vol. 3, pp. 111-133.

[46] Markowitz, H.M.. 1959. Portfolio selection: efficient diversification of investments. New York, John Wiley & Sons.

[47] Markowitz, H.M. 1991. Foundations of portfolio theory. The Journal of Finance, Vol. 46, No. 2, pp. 469-477.

[48] Markowitz, H.M., Todd, T., Xu, G., Yamane, Y. 1993. Computation of mean-semivariance efficient sets by the critical line algorithm. Annals of Operations Research, Vol. 45, pp. 307-317.

[49] Markowitz, H.M. Portfolio theory: as I still see it. Annual Review of Financial Economics, Vol. 2, No. 1, pp. 1-23.

[50] Mausser, H., Saunders, D., Seco, L. Optimizing omega. [Online] Available from: `http://www.nag.co.uk/IndustryArticles/OptimizingOmegaPaper.pdf`.

[51] Meucci, A. 2009. Risk and asset allocation. 1st Ed. Berlin: Springer Finance.

[52] Meucci, A. 2011. A new breed of copulas for risk and portfolio management. [Online]. Available from: `http://www.symmys.com/node/335`.

[53] Meyer, J. 1987. Two-moment decision models and expected utility maximization. The American Economic Review, Vol. 77, No. 3, pp. 421-430.

[54] Modigliani, F., Modigliani, L. 1997. Risk-Adjusted Performance. The Journal of Portfolio Management, Vol. 23, No. 2, pp. 45-54.

[55] Nelsen, R.B. 2006. An introduction to copulas. Springer Series in Statistics, 2nd Ed. New York: Springer Finance.

[56] Nguyen-Thi-Thanh, H. 2007. Assessing hedge fund performance: does the choice of measures matter? [Online]. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.496.3097.

[57] Passow, A. 2005. Omega portfolio construction with Johnson distributions. [Online]. Available from: http://m.risk.net/digital_assets/179/risk_0405_technical_passow.pdf.

[58] Pederson, C.S., Satchell, S.E. 1998. Choosing the right measure of risk: a survey. The Current State of Economic Science.

[59] Pederson, C.S., Satchell, S.E. 2002. On the foundation of performance measures under asymmetric returns. Quantitative Finance, Vol. 2, No. 3.

[60] Platen, E., Sidorowics, R. 2007. Empirical evidence on student-t log-returns of diversified world stock indices. Quantitative Finance Research Centre, No. 194. [Online]. Available from http://www.qfrc.uts.edu.au/research/research_papers/rp194.pdf.

[61] Ravev, Z., Racheva-Iotova, B., Stoyanov, S. 2010. Capturing fat tails. Risk Magazine, 30 April.

[62] Rockafellar, R.T., Uryasev, S. 2002. Conditional value-at-risk for general loss distributions. Journal of Banking and Finance, Vol. 26, No. 7, pp. 1443-1471.

[63] Satchell, S. 2009. Optimizing optimization. 1st Ed. USA: Academic Press.

[64] Samuelson, P.A. 1970. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. The Review of Economic Studies, Vol. 34, No. 4, pp. 537-542.

[65] Sharpe, W.F. 1966. Mutual fund performance. The Journal of Business, Vol. 39, No.1, pp. 119-138.

[66] Shaw, W., Lee, K.T.A. 2008. Bivariate student t distributions with variable marginal degrees of freedom and independence. Journal of Multivariate Analysis, pp 1276-1287.

[67] Shaw, W. 2008. Share price movements in the post-credit crunch environment. [Online]. Available from: `http://www.mth.kcl.ac.uk/~shaww/web_page/papers/AssetPriceMovements_PCCE1.2.pdf`.

[68] Shaw, W. 2010. Monte carlo portfolio optimization for general investor risk-return objectives and arbitrary return distributions: a solution for long-only portfolios. [Online]. Available from: `http://papers.ssrn.com/sol3/papers.cfm?abstractid=1680224`.

[69] Shaw, W. 2011. Portfolio optimization for VaR, CVaR, omega and utility with general return distributions: a monte carlo approach for long-only and bounded short portfolios with optional robustness and a simplified approach to covariance

matching. [Online]. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstractid=1856476.

[70] Shaw, W. 2011. Risk, VaR, CVaR and their associated portfolio optimizations when asset returns have a multivariate student t distribution. [Online]. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1772731.

[71] Simaan, Y. 1998. What is the opportunity cost of mean-variance investment strategies? Management Science, Vol. 39, No. 5, pp. 578-587.

[72] Simons, K. 1998. Risk-adjusted performance of mutual funds. New England Economic Review, pp. 33-48.

[73] Sortino, F., Price, L.N. 1994. Performance measurement in a downside risk framework. The Journal of Portfolio Management, Vol. 3, No. 3, pp 59-64.

[74] Sortino, F., Van der Meer, R. 1991. Downside risk. The Journal of Portfolio Management, Vol. 17, No. 4, pp 27-31.

[75] Sortino,F., Forsey, H.J. 1996. On the use and misuse of downside risk. The Journal of Portfolio Management, Vol. 22, No. 2, pp 35-42.

[76] Stone, B.K. 1973. A general class of three-parameter risk measures. Journal of Finance, Vol. 28, No. 3, pp 675-685.

[77] STUDENT. 1908. The probable error of a mean. Biometrika, Vol. 6, No.1, pp 1-25.

[78] Twain, M. 1894. Pudd'nhead wilson. Charles L.Webster & Company.

[79] Wiesinger, A. 2010. Risk-Adjusted Performance Measurement - State of the Art. [Online]. Available from: http://www.iorcf.unisg.ch/Forschung/~/media/0D77AF25603347E795C77DDA1DD83676.ashx.

[80] Winton Capital Management (2003) Assessing CTA Quality with the Omega Performance Measure. [Online]. Available from: http://www.performance-measurement.org/Winton2003.pdf.

[81] Winton Capital Management (2003) Case Studies of CTA Assessment Using the Omega Performance Measure. [Online]. Available from: http://www.performance-measurement.org/Winton2003a.pdf.

[82] Zakamouline, V. 2011. The Performance Measure You Choose Influences the Evaluation of Hedge Funds. Journal of Performance Measurement, Vol. 15, No. 3, pp 48-64.

Chapter 9

Appendix

6-Asset Case: Asset return first four moments

Table 9.1: 6-asset case (Boom period): First four moments.

Statistic/Asset	EUR	GBP	JPY	GOLD	Oil	SILVER
Mean (% p.a.)	6.30	5.45	0.37	15.58	25.17	27.02
Standard Deviation (% p.a.)	8.81	8.36	8.76	16.97	29.44	32.24
Skewness	-0.09	-0.07	-0.20	-0.44	-0.07	-0.77
Kurtosis	3.72	3.52	4.23	5.58	4.16	8.98

18-Asset Case: Optimal Portfolios

Table 9.2: 6-asset case (Bust period): First four moments.

Statistic/Asset	EUR	GBP	JPY	GOLD	Oil	SILVER
Mean (% p.a.)	0.00	-0.05	-0.08	0.23	0.18	0.32
Standard Deviation (% p.a.)	0.12	0.12	0.12	0.23	0.38	0.46
Skewness	0.21	0.03	-0.31	-0.13	0.06	0.01
Kurtosis	5.78	6.72	6.68	6.03	6.74	9.59

Table 9.3: Boom period optimal portfolios: Simulated Student t Returns.

Performance Measure	Benchmark (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Sharpe Ratio	5	0.0012	0.2056	0.0201	0.1890	0.0000	0.0484	0.1492	0.0000	0.0000	0.0000	0.0002	0.0537	0.1981	0.0007	0.0700	0.0001	0.0637	0.0000
Omega	5	0.0210	0.2478	0.0000	0.1714	0.0001	0.0010	0.1617	0.0010	0.0001	0.0000	0.0002	0.0275	0.1727	0.0000	0.1521	0.0010	0.0421	0.0003
Sortino	5	0.0099	0.0097	0.0006	0.0816	0.0001	0.0017	0.1609	0.0167	0.0213	0.0000	0.0267	0.0366	0.1657	0.0018	0.1372	0.0052	0.2370	0.0874
$\phi(b,1,3)$	5	0.0028	0.0000	0.0426	0.1417	0.0000	0.0012	0.1054	0.0002	0.0002	0.0007	0.0001	0.0000	0.1871	0.0020	0.2093	0.0001	0.3003	0.0061
$\phi(b,1,4)$	5	0.0000	0.0005	0.0006	0.1063	0.0423	0.0495	0.1046	0.0307	0.0081	0.0001	0.1169	0.0236	0.1782	0.0009	0.1522	0.0042	0.1732	0.0082
Sharpe Ratio	10	0.0001	0.1709	0.0000	0.1974	0.0000	0.0086	0.1845	0.0000	0.0000	0.0000	0.0013	0.0665	0.3637	0.0000	0.0070	0.0000	0.0000	0.0000
Omega	10	0.0000	0.3797	0.0000	0.1512	0.0000	0.0029	0.0000	0.0025	0.0047	0.0000	0.0000	0.0729	0.3819	0.0032	0.0009	0.0001	0.0000	0.0000
Sortino	10	0.0020	0.0000	0.0000	0.2173	0.0000	0.0065	0.0789	0.0002	0.0147	0.0000	0.0005	0.0398	0.3085	0.0000	0.2988	0.0003	0.0000	0.0325
$\phi(b,1,3)$	10	0.0000	0.0000	0.0004	0.0000	0.0002	0.0000	0.1516	0.0006	0.0382	0.0000	0.0833	0.0510	0.1825	0.0004	0.1837	0.0000	0.2840	0.0240
$\phi(b,1,4)$	10	0.0035	0.0002	0.0434	0.2006	0.0209	0.0000	0.1050	0.0000	0.0074	0.0207	0.1285	0.0004	0.1631	0.0001	0.1141	0.0000	0.1532	0.0388
Sharpe Ratio	15	0.0000	0.1718	0.0001	0.2168	0.0000	0.0000	0.0003	0.0003	0.0048	0.0002	0.0000	0.0963	0.5045	0.0000	0.0006	0.0000	0.0044	0.0000
Omega	15	0.0035	0.2032	0.0000	0.0379	0.0000	0.0000	0.0802	0.0000	0.0000	0.0001	0.0000	0.1180	0.5136	0.0000	0.0055	0.0000	0.0182	0.0197
Sortino	15	0.0030	0.0738	0.0023	0.1577	0.0000	0.0000	0.0707	0.0000	0.0000	0.0000	0.0024	0.1004	0.3455	0.0005	0.1518	0.0007	0.0885	0.0027
$\phi(b,1,3)$	15	0.0141	0.0007	0.0000	0.0001	0.0002	0.0007	0.1261	0.0483	0.0036	0.0000	0.0015	0.0585	0.2300	0.0000	0.2295	0.0000	0.2701	0.0166
$\phi(b,1,4)$	15	0.0005	0.0008	0.0001	0.0010	0.0019	0.0072	0.0010	0.0002	0.2132	0.0043	0.0058	0.0805	0.1916	0.0062	0.1376	0.0000	0.2925	0.0556

Table 9.4: Bust period optimal portfolios: Simulated Student t Returns.

Performance Measure	Benchmark (%)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Sharpe Ratio	5	0.0000	0.4649	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5350
Omega	5	0.0000	0.4246	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.5753
Sortino	5	0.0000	0.5854	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4143
$\phi(b,1,3)$	5	0.0000	0.3859	0.0011	0.0000	0.0000	0.3175	0.0149	0.0000	0.0000	0.0000	0.0000	0.1678	0.0000	0.0001	0.0000	0.0017	0.0000	0.1110
$\phi(b,1,3)$	5	0.0000	0.2562	0.0002	0.0000	0.0000	0.4973	0.0000	0.0018	0.0000	0.0001	0.0006	0.2435	0.0000	0.0000	0.0001	0.0000	0.0003	0.0000
Sharpe Ratio	10	0.0000	0.3126	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6873
Omega	10	0.0000	0.2914	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.7085
Sortino	10	0.0000	0.5151	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.4844
$\phi(b,1,3)$	10	0.0022	0.4570	0.0000	0.0000	0.0000	0.2648	0.0000	0.0000	0.0000	0.0000	0.0000	0.1078	0.0000	0.0000	0.0000	0.0000	0.0002	0.1679
$\phi(b,1,4)$	10	0.0000	0.2826	0.0000	0.0000	0.0000	0.4637	0.0000	0.0001	0.0000	0.0000	0.0000	0.2523	0.0000	0.0000	0.0000	0.0012	0.0000	0.0001
Sharpe Ratio	15	0.0000	0.0104	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9896
Omega	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Sortino	15	0.0000	0.4765	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.5234
$\phi(b,1,3)$	15	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0001	0.0000	0.0000	0.3161	0.2772	0.0000	0.0001	0.0000	0.0000	0.4050
$\phi(b,1,3)$	15	0.0004	0.2535	0.0002	0.0000	0.0000	0.4897	0.0000	0.0000	0.0000	0.0000	0.0000	0.2555	0.0004	0.0000	0.0000	0.0000	0.0000	0.0002