

Robust Portfolio Construction Using Sorting Signatures

Lillian Kasenene

A dissertation submitted to the Department of Actuarial Science, Faculty of Commerce, at the University of Cape Town, in partial fulfillment of the requirements for the degree of Master of Philosophy

23rd May 2014

*Master of Philosophy Specialising in Mathematical Finance,
University of Cape Town,
Cape Town*



The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy at the University of Cape Town. It has not been submitted before for any degree or examination at any other University.

23rd May 2014

Abstract

Mean-variance analysis introduced by Harry Markowitz has been criticised in the past mainly due to the counter-intuitive and unstable nature of the resultant portfolios from the optimisation. These disappointing results have been linked to the presence of estimation error in the estimates of the expected returns and covariances which serve as input to the optimisation. Several attempts have been made to produce more reliable estimates, with a significant amount of effort and resources placed in estimation of expected returns, which is generally a more difficult task than estimation of covariances. Almgren and Chriss (2006) provide a methodology for portfolio selection in which the order of expected returns replaces the numerical values of the returns. This framework allows full use of the covariance matrix, in a method analogous to mean-variance optimisation. We adopt this framework in our analysis together with the robust optimisation technique introduced by Golts and Jones (2009) which improves the estimate of the covariance matrix by direct modification in the optimisation process. Golts and Jones (2009) argue that a reduction of the angle between the input return forecasts and the output portfolio positions results in more investment relevant portfolios, inline with the investment manager's insights. They relate this angle to the condition number of the covariance matrix and use robust optimisation to improve the conditioning of this matrix. Assuming perfect alpha foresight of an investment manager, we apply a combination of the techniques of Almgren and Chriss (2006) and Golts and Jones (2009) to South African equity data and show that the resultant robust portfolios, though conservative in their risk-adjusted return statistics, are more diversified and exhibit lower leverage than mean-variance portfolios. We further show that independent of the optimisation method, there is a marginal difference in the performance of portfolios created using ordering information and actual returns.

Acknowledgements

I would like to express the deepest appreciation to my supervisor, Professor Dave Bradfield, for the constant engagement, direction, and support throughout this learning journey. In the same vein I acknowledge and appreciate the support of Geraldine Bailey, my co-supervisor, who helped me find momentum and solutions when there did not seem to be any. None of this would have been possible without the both of you.

I also thank Yashin Gopi from BNP Paribas Cadiz Securities in a special way for finding the time to provide the data that was the input for this work and for the useful comments and direction provided.

My sincere gratitude goes to Professor David Taylor for selecting this line of research for me and for the facilitation and direction throughout the course. Special mention and appreciation must go to HSBC for the bursary that went a long way to helping me to complete this course. I would like to thank all those at HSBC who continue to support this course and individuals like myself and encourage them to continue to do so.

To my colleagues and classmates who were great study companions, peers and friends, I thank you for making this journey possible and I cherish the new friendships I have made.

And as always this journey as all journeys in my life would not have been possible without the prayers, support, encouragement and love of my wonderful husband and family. I especially thank them and dedicate most of my work to them. Most of all, I thank God who provided me with life, health and willed this journey through the year.

Contents

1. Introduction	1
2. Literature Review	4
2.1 Ordering Information	4
2.2 Robust Portfolio Construction	6
3. Portfolios from Sorts	8
3.1 Portfolio Sort	8
3.2 Portfolio Preference Relation	9
3.3 Optimal Portfolio	10
3.4 Resultant Portfolios	12
4. Robust Optimisation and the Alpha-Weight Angle	13
4.1 Separation of Magnitude and Direction	14
4.2 Relationship between Alpha-weight Angle and Condition Number	15
4.3 Robust Optimisation and the Alpha-weight Angle	16
4.3.1 Spherical uncertainty region for alphas	16
4.3.2 Investment direction of robust weights	17
4.3.3 Investment magnitude of robust weights	18
4.3.4 Resultant portfolios	18
5. Empirical Analysis	20
5.1 Data and Methodology	20
5.2 Portfolio Performance Statistics	23
5.2.1 Return statistics	23
5.2.2 Risk statistics	24
5.2.3 Risk-adjusted return statistics	24
5.2.4 Beta	26
5.3 Condition Number and Alpha-Weight Angle	27
5.3.1 Condition number	27
5.3.2 Alpha-weight angle	29
5.4 Effect of Changes in χ	32
5.5 Portfolio Diversification and Concentration	34
5.5.1 Number of stocks in portfolio	34
5.5.2 Herfindahl-Hirschman Index (HHI)	35
5.5.3 Leverage	35
5.5.4 Comparison of Portfolio Weights	37

5.5.5	Comparison of Change in Portfolio Weights After Rebalancing	37
5.6	Performance Statistics for Varying IC	39
5.6.1	Risk statistics	39
5.6.2	Outperformance relative to benchmark	39
5.6.3	Sharpe and Information ratio	40
5.7	Summary of Pros and Cons of Each Technique	40
6.	Conclusion	42
	Bibliography	44
A.	Derivation of Formula for Computation of Centroid	46
B.	Matlab Code	48

List of Figures

3.1	Sorting signature for single complete sort of Top 100 ALSI stocks . .	11
3.2	Efficient frontiers of centroid and mean-variance portfolios	12
5.1	Backtesting methodology (<i>Source: Munro (2010)</i>)	21
5.2	Cumulative returns	23
5.3	Realised Beta	26
5.4	Eigenvalues: Case 1	27
5.5	Eigenvalues: Case 2	28
5.6	Eigenvalues: Case 3	28
5.7	Eigenvalues: Case 4	29
5.8	Alpha-weight angle for simple optimisation	30
5.9	Alpha-weight angle for robust optimisation	31
5.10	Outperformance statistics for varying values of χ	32
5.11	Fund Risk statistics for varying values of χ	33
5.12	Tracking Error statistics for varying values of χ	33
5.13	Herfindahl-Hirschman Index (HHI)	35
5.14	Portfolio leverage	36
5.15	Comparison of Portfolio Weights	37

List of Tables

4.1	Classical optimisation problems	14
4.2	Robust optimisation problems	16
5.1	Annualised outperformance	24
5.2	Fund risk statistics	24
5.3	Sharpe and Information ratios(annualised)	26
5.4	Average number of stocks	34
5.5	Average number of stocks with weights greater than 0.5%	34
5.6	Average leverage	36
5.7	Ad-hoc Calculation of Turnover	38
5.8	Risk statistics for varying IC	39
5.9	Annualised fund outperformance for varying IC	39
5.10	Risk-adjusted performance statistics for varying IC	40

Chapter 1

Introduction

Mean-Variance analysis, introduced by Harry Markowitz in his 1952 paper entitled '*Portfolio Selection*' established the foundations of what is today known as Modern Portfolio Theory. Within the Markowitz (1952) framework, a rational investor will select a portfolio based on the trade-off between expected return and risk. The portfolio offering the highest level of return for a given level of risk, or conversely, the portfolio with the lowest level of risk for a set level of expected return will be chosen as optimal. Theoretically, this framework is simple and intuitive but practically has well-known limitations that have led to heavy criticism by investment practitioners.

In its classical form, mean-variance analysis takes as input the expected return and covariance matrix of the assets in the universe, and produces as output, optimal portfolios. The inputs are calculated as statistical estimates from historical data and are therefore uncertain. Empirical research has highlighted several shortcomings of mean-variance optimal portfolios, justifying the limited use of this optimisation in the absence of practical enhancements. These portfolios have been shown to be counter-intuitive, with the resulting portfolio weights bearing little resemblance to the initial return forecasts. Mean-variance portfolios have also been criticised for not being well-diversified but generally concentrated in a limited number of assets/asset classes (see Jorion (1985)), moreover, equally-weighted portfolios have in practice been shown to outperform mean-variance portfolios. Furthermore, the optimisation has been shown to be highly sensitive to changes in the input parameters with small changes in the inputs producing completely different portfolios!

These limitations have been linked to the presence of estimation error in the inputs and what Michaud (1989) has termed '*error-maximisation*' property of the optimiser. The fact that the inputs are estimated with error is unknown to the optimiser leading it to overweight securities with high expected returns and low standard deviations and vice versa. Additionally, the optimisation involves the inversion of the

covariance matrix and Michaud (1989) highlights that an ill-conditioned covariance matrix will generally lead to unstable solutions. However, despite these limitations, mean-variance analysis still holds as the theoretical framework for portfolio selection albeit with various enhancements in order to make the analysis practically useful.

The use of factor models and Bayesian/Shrinkage estimators in place of the sample mean and covariance matrix have been proposed in order to develop more reliable estimates of the input parameters. Particularly, the generation of useful estimates for expected returns is a more difficult and resource-intensive task than the estimation of the covariance matrix (see Merton (1980)). Furthermore, Chopra and Ziemba (1993) show that errors in returns are approximately ten times as important as errors in variances and covariances in the optimisation process. This further emphasises the importance of the generation of reliable estimates of expected returns. In order to make the optimisation results more practically relevant, other enhancement techniques such as, the imposition of constraints in order to limit concentration, and the use of robust optimisation that takes estimation error into account directly in the optimisation process have been employed.

In this project we enhance mean-variance optimisation for South African equity data using a combination of practical techniques developed by Almgren and Chriss (2006) and Golts and Jones (2009). As highlighted above, the estimation of expected returns is a difficult and complicated task and therefore a methodology that simplifies this process is a huge opportunity for any investment manager. Almgren and Chriss (2006) provide a framework for portfolio selection, analogous to mean-variance analysis, in which the use of expected return data is replaced with information on the order of expected returns. Ordering information for expected returns is relatively simple to obtain as past research has shown that there exist correlations between expected returns and variables such as firm characteristics and past returns history. For the cases where perfect foresight is assumed, we use the future returns data to generate the order of returns and then employ the robust optimisation technique of Golts and Jones (2009) in the analysis. For the less-than-perfect foresight scenario, we generate forecasted returns based on the specified Information Coefficient (the methodology we employ to do this is explained in Chapter 5) and then employ these forecasts to create the sort. In formulation of the robust optimisation, we take the expected returns to be uncertain and specify a spherical uncertainty region around the return estimates. The optimisation modifies the covariance matrix directly, producing a better conditioned matrix, thereby leading to more stable portfolios.

The purpose of this research is to assess whether any benefits exist from the construction of portfolios using both ordering information and robust portfolio optimisation. We do this by comparing the performance and descriptive statistics of portfolios generated using classical mean-variance optimisation, ordering information alone in a mean-variance framework, robust optimisation, and a combination of ordering information and robust optimisation. Almgren and Chriss (2006) showed that the efficient frontiers generated when using ordering information and classical mean-variance analysis are the same, leading to similar performance characteristics. In this research, we shall further assess whether this is also true when using robust optimisation in place of classical mean-variance optimisation. Firstly, we assume perfect alpha foresight and then repeat the analysis for less than perfect foresight to investigate what changes there are to the performance.

Chapter 2

Literature Review

2.1 Ordering Information

Almgren and Chriss (2006) define ordering information as any set of inequality beliefs about expected returns. These beliefs form what is known as a ‘*sort*’ of the expected returns. Sorting can be established in a number of ways; it could be across a universe of stocks with one stock expected to outperform another, across sectors with stocks in one sector expected to outperform those in another, etc. Prior research has shown that there exist relationships between variables such as firm characteristics and past returns history, and expected returns. Information on these variables is usually easily available and so these relationships can be used to generate ordering information. Examples of these relationships already documented in the literature are highlighted below.

Firm characteristics may include among others, price-earnings ratios, firm size, and book-to-market equity. Basu (1977) investigated price-earnings ratios and showed that securities with lower ratios usually outperform securities with higher ratios. Banz (1981) and Reinganum (1981) showed that there existed a non-linear relationship between returns and total market value, with smaller firms having higher risk-adjusted returns on average, than larger firms. Fama and French (1992) showed that a combination of firm size and book-to-market equity captured the variation in the cross-section of expected stock returns. Jegadeesh and Titman (1993) investigated momentum and argued that strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past generate significant positive returns over 3 to 12-month holding periods. Lamont *et al.* (2001) showed that constrained firms earned lower returns than unconstrained firms, whereas Vassalou and Xing (2005) showed that firms which experience large increases in their default risk earn higher subsequent returns than firms that experience large decreases in their default risk. As highlighted above, not only one

variable can be used to generate a sort; various combinations of factors can also be used, giving rise to single and multiple sorts respectively. For the case where the sort is the only information available with respect to the expected returns, the challenge therefore is to establish a methodology for selection of an optimal portfolio.

Several ad hoc methods have been used in the past to generate portfolios from ordering information. An example is a technique roughly corresponding to the approach of Thorp (2003) where portfolio weights are set to be proportional to a *linear* vector that is long the highest-ranked asset and short the lowest. This method does not utilise the covariance matrix and therefore does not involve any optimisation. The methodology of Almgren and Chriss (2006) is analogous to mean-various analysis in that it takes as inputs the order of expected returns and the covariance matrix. The input vector of ordered returns specified by Almgren and Chriss (2006) is *non-linear* and the resultant portfolios generated in a mean-variance framework using this vector have been shown to outperform those generated using a linear sort. The incorporation of covariance information in this framework has also been shown to provide better risk-adjusted performance statistics than methods that utilised only ordering information.

In Chapter 12 of Satchell and Scowcroft (2003), a methodology different from that of Almgren and Chriss (2006), albeit in a mean-variance setting, is specified for calculation of the optimal rank portfolio. The difference in these two methodologies is in the calculation of the vector of ordered returns that serves as input to the mean-variance optimisation. In Satchell and Scowcroft (2003), a linear ordering of stocks based on past returns (in descending order) is specified at each time interval and then the expectation calculated in the usual way over the full time period. The criticism of this approach is that the optimal rank portfolios have been shown to be less efficient than classical mean-variance optimisation portfolios. In contrast, Almgren and Chriss (2006) calculate the input vector of ordered returns as the centroid of the geometric cone of return vectors consistent with the specified sort (e.g, descending order of past returns). Almgren and Chriss (2006) show that this vector is non-linear and its use as input to the optimisation results in generation of an efficient frontier identical to that of classical mean-variance analysis, leading us to believe that in this sense, centroid optimal portfolios are more superior.

2.2 Robust Portfolio Construction

It is well known from empirical evidence that mean-variance optimisation is often unstable and highly sensitive to changes in the input parameters. This instability can be partially explained by the presence of estimation error in the inputs (see Michaud (1989)). The optimisation does not know that the inputs are statistical estimates and so takes them as certain. In order therefore to generate improvements in the analysis, it is necessary to perform the optimisation under uncertainty, taking the estimation error into account directly in the optimisation process. This involves the specification of uncertainty regions or confidence intervals for the input parameters, leading to the generation of portfolios that can perform well under a range of different scenarios.

Initially, stochastic programming was used to perform optimisation under uncertainty. This required that detailed information about the probability distributions of the uncertain parameters be available. Specification of these distributions is usually a very difficult and complicated task which led to the limited use of these methods in portfolio construction, and in turn, the popularity of robust optimisation techniques. In robust optimisation, general assumptions of the probability distribution of the uncertain parameters are made, leading it to be easily adopted as a more tractable and cost-efficient alternative to stochastic programming (Fabozzi *et al.* (2007)).

Different forms of robust optimisation are documented in the literature, each with a unique specification of the uncertainty set of the input parameters. Costa and Paiva (2000) defined the uncertainty set of the expected returns and covariance matrix as a convex combination of some known vertices of a polytope. They solved the resultant robust formulations of the mean-variance optimisation problems as linear-matrix inequalities (LMI) using existing numerical programs. In contrast, Halldórsson and Tütüncü (2003) allowed for the return and covariance matrix information to be presented in the form of an interval, with restrictions placed on the specification of the intervals for the covariance matrix to ensure that it remains semi-definite. They then developed an algorithm that could be applied to solve the resulting optimisation problem. One of the criticisms of the work of Halldórsson and Tütüncü (2003) is that they do not provide an approach for specification of the parameter intervals in the uncertainty set. Goldfarb and Iyengar (2003) specify this in the presentation of an alternative approach for formulating robust optimisation problems in which a factor model is used for asset returns. They show that the uncertainty sets are defined naturally by the statistical procedures used to estim-

ate these returns, and formulate the corresponding robust optimisation problems as second-order cone programs (SOCPs) which can be solved very efficiently using interior-point algorithms. Later, Ceria and Stubbs (2006) presented a more generalised framework, applicable to practical situations, for the formulation of robust portfolio optimisation problems. This framework makes the analysis more effective by preventing the optimisation from adjusting the returns based on anything other than estimation error (specified constraints could otherwise cause further adjustments). In recent research, Golts and Jones (2009) present a robust optimisation formulation in which a spherical uncertainty region is defined for the expected returns and the sample covariance matrix is modified directly in the optimisation process. After some mathematical manipulation of the objective function, the resultant robust forms of classical mean-variance optimisation problems can be solved efficiently using existing numerical algorithms.

Performance and descriptive statistics of robust portfolios have been shown to be distinct from Markowitz (1952) optimal portfolios. Tütüncü and Koenig (2003) showed that robust portfolios tend to invest in fewer asset classes compared to mean-variance portfolios, and that robust asset allocation generally improved the worst-case behaviour of portfolios. However, increase in the size of the uncertainty sets led to under-performance in the *more likely* scenarios, suggesting that a trade-off needs to be made based on the risk attitude of investors. In addition, when analysing the change in portfolio composition over long periods, they report evidence that robust portfolios have significantly less turnover.

Chapter 3

Portfolios from Sorts

Research by Merton (1980) highlighted the difficulty in generating reliable estimates of expected returns for input into mean-variance optimisation. To date, estimation of expected returns is still a difficult and resource-intensive task for portfolio managers. There exist cases where either one cannot obtain enough data to make estimates of expected returns or one does not have sufficient confidence in the reliability of available estimates. In such cases, the onus is on the portfolio manager to modify the available estimates in order to obtain more useful estimates, which may be done objectively, subjectively or by a combination of both. Almgren and Chriss (2006) do this modification by simply replacing expected returns with ordering information and providing a framework analogous to mean-variance analysis in which both the ordering information and covariance matrix are utilised in the optimisation. Theoretical details of this methodology are presented in this section.

3.1 Portfolio Sort

A portfolio sort is defined as a set of inequality relationships between the expected returns of a set of assets. Suppose S_1, \dots, S_n is the available universe of n assets. For expected returns denoted by r_1, \dots, r_n , a portfolio sort could be defined by:

$$r_1 \geq r_2 \geq \dots \geq r_n$$

This is a *single complete sort*, and is the simplest and most common type of sort, which orders all the assets of the portfolio in descending order of expected returns. Alternative representations of sorts are:

$$r_1 \geq \dots \geq r_m \geq 0 \geq r_{m+1} \dots \geq r_n$$

$$r_1 - r_2 \geq r_3 - r_4 \geq \dots \geq r_{n-1} - r_n$$

As shown from the representations above, the possibilities are endless with regard to defining the inequality relationships, moreover, as mentioned in Chapter 2, both

single and multiple sorts can be defined. The methodology of Almgren and Chriss (2006) conveniently presents practical usability as it is generalised and can therefore be applied to both simple and more complicated types of sorts. For the purpose of this research however, and in the text that follows, we assume only a single complete sort which we generate from information on assets' future returns (assuming perfect foresight).

3.2 Portfolio Preference Relation

In the Markowitz (1952) framework, a rational investor selects a portfolio with the highest level of expected return for a given level of risk, or conversely lowest level of risk for a set level of return. When using ordering information in the absence of numerical values for expected returns, Almgren and Chriss (2006) present an alternative way of portfolio selection using *preference* relations. They postulate that an expected return vector $r = (r_1, \dots, r_n)$ is consistent with the sort if $r_1 \geq r_2 \geq \dots \geq r_n$. For a set Q of all consistent expected returns related to a sort, they define the portfolio preference relation as:

'If w and v are portfolios, then neglecting all constraints, an investor should prefer to hold w over v ($w \succeq v$) if the expected return of w is greater than or equal to that of v for every consistent expected return vector r , that is, for every $r \in Q$.'

Almgren and Chriss (2006) show mathematically that in order to produce optimal portfolios, the above definition is more than sufficient. A weaker but *sufficient* definition of the portfolio preference relation can therefore be stated as:

*'Neglecting all constraints, an investor should prefer w over v if the expected return of w is greater than or equal to that of v for a **greater fraction** of possible expected return vectors.'*

In order to mathematically compare portfolios based on this new definition of the preference relation, a radially symmetrical probability measure μ on the space Q of consistent expected returns is introduced (An uncorrelated normal distribution could qualify as one such distribution). μ assumes that each expected return direction is equally likely and therefore assigns equal probability to every direction in Q . The preference relation can now be expressed mathematically as:

$$w \succeq v \text{ if and only if } \mu(\{r \in Q | w \cdot r \geq v \cdot r\}) \geq \mu(\{r \in Q | v \cdot r \geq w \cdot r\})$$

3.3 Optimal Portfolio

The next step is to define a methodology for selection of an optimal portfolio when a budget constraint has been set (usually portfolio risk is used as the budget constraint). A portfolio is therefore defined to be optimal with respect to a sort if it is the *most preferable* under the above preference relation for a given level of risk. In order to calculate the optimal portfolio, Almgren and Chriss (2006) show that there is a vector c , defined as the *geometric centre of mass* or *centroid* of the set Q , with the following property:

$$w \succeq v \text{ if and only if } w \cdot c \geq v \cdot c$$

Conveniently, the centroid vector can be easily computed as it is exactly the same as the geometric centroid of the consistent cone, Q , of expected return vectors. The computation can be done analytically for simple sorts or using Monte Carlo methods for more complicated sorts. For the case of a single complete sort of n assets, the j th component of c is approximated to within 0.5% by:

$$c_{j,n} \approx \mathcal{N}^{-1} \left(\frac{n+1-j-\alpha}{n-2\alpha+1} \right) \quad (3.1)$$

where

$$\alpha = A - Bn^{-\beta}$$

$\mathcal{N}^{-1}(\cdot)$ is the inverse cumulative normal distribution, $A = 0.4424$, $B = 0.1185$ and $\beta = 0.21$. Refer to *Appendix A* for a detailed derivation of equation (3.1), including assumptions made, as outlined by Almgren and Chriss (2006)

Therefore, in order to find the optimal portfolio, given the centroid and portfolio risk budget constraint, the following optimisation problem is solved:

$$\max_w w \cdot c \text{ subject to } w' \cdot \Sigma \cdot w \leq \sigma^2 \quad (3.2)$$

where Σ is the assets' covariance matrix and σ^2 is the budget constraint. We can see clearly that equation (3.2) above is analogous to the classical optimisation problem in mean-variance analysis with the centroid replacing the vector of expected returns.

Almgren and Chriss (2006) show that the efficient portfolios generated (which they call '*centroid portfolios*' or '*portfolios from sorts*') are exactly the same as the Markowitz (1952) efficient portfolios for expected returns that are both consistent with the sort and sum to zero (refer to Almgren and Chriss (2006) for proof). The specified centroid vector in this methodology possesses both of these properties. It is

non-linear and similar to the image of a set of equally spaced points of the cumulative normal function with all of its elements summing to zero. Therefore the use of the centroid vector in place of expected returns results in the same efficient portfolios as in mean-variance analysis. Figure 3.1 is an example of the signature of the centroid vector for a single complete sort of Top 100 ALSI stocks.

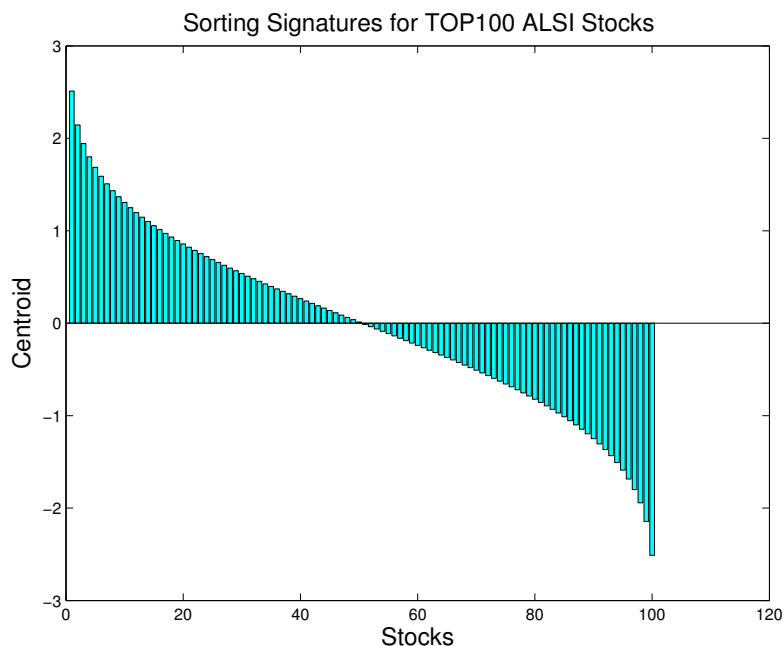


Fig. 3.1: Sorting signature for single complete sort of Top 100 ALSI stocks

To generate Figure 3.1 the following methodology was employed: We used weekly data on shares listed on the All Share Index (ALSI) and an *out-of-sample* period of January 2006 to November 2013 was specified. At the end of each month, the largest 100 shares on the ALSI were identified (note that in some cases, more than one share may have the same weights, making the number of shares to be used in the sample more than 100). We selected the first date in the *out-of-sample* period and used the shares' returns data to calculate the centroid analytically as per equation (3.1). We then created a plot of the centroid vector as per Figure 3.1.

We further compared the Markowitz (1952) and centroid efficient frontiers to verify that they are indeed the same (see Figure 3.2 below). Again, we used the returns data and covariance matrix for the Top 100 ALSI stocks on the first date of our *out-of-sample* period. For the Markowitz efficient frontier, we did a simple unconstrained mean-variance optimisation using MATLAB with inputs as the returns

data and covariance matrix. The resulting optimal portfolio risk/return vectors were then used to plot the efficient frontier. For the centroid efficient frontier, we did the same optimisation in MATLAB but with inputs as the calculated centroid vector and covariance matrix. The resulting portfolio weights are then scaled by the original returns vector to generate the returns vector for the centroid optimal portfolio that is plotted against the risk vector on the efficient frontier.

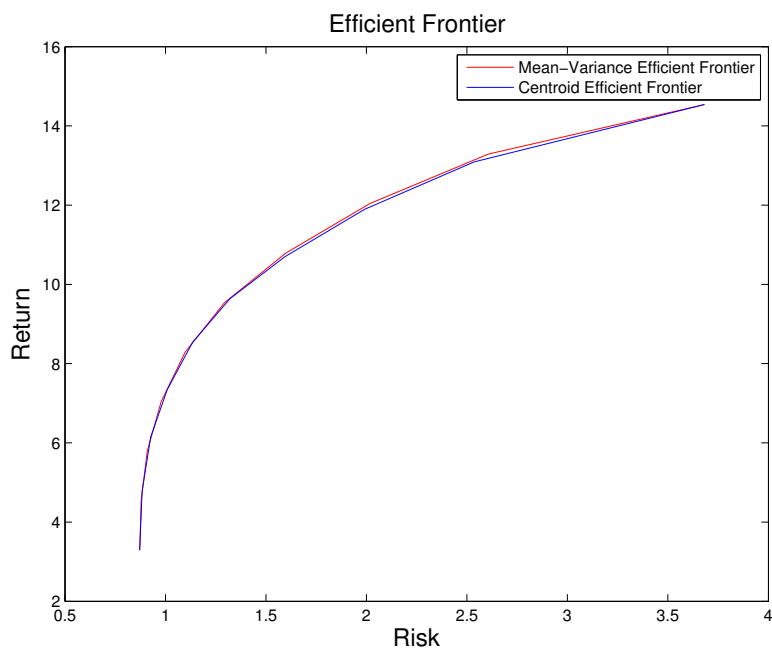


Fig. 3.2: Efficient frontiers of centroid and mean-variance portfolios

3.4 Resultant Portfolios

In their analysis, Almgren and Chriss (2006) compared the performance of the resultant portfolios to those generated using ad-hoc techniques that were not within the mean-variance framework (such as linear representation of sorts and generation of portfolios without use of covariance information). They showed that the use of covariance information improved risk-adjusted performance and that even without the use of covariance information, the centroid portfolios showed superior performance to those constructed using linear representations of sorts.

Chapter 4

Robust Optimisation and the Alpha-Weight Angle

Michaud (1989) argues that in a Markowitz (1952) framework, an ill-conditioned covariance matrix will generally lead to unstable solutions since the optimisation involves inversion of the covariance matrix. Several techniques have been proposed to regularise the covariance matrix, one of which is the robust optimisation methodology of Golts and Jones (2009). In their paper, '*A Sharper Angle on Optimisation*', Golts and Jones (2009), show that through robust optimisation, we can modify the input covariance matrix directly in the optimisation, making it better-conditioned where necessary. The resulting portfolios from this methodology are more stable with more intuitive portfolio positions in the sense that they reflect the manager's insights as specified in the input alphas. They introduce the concept of alpha-weight angle to quantify the alignment of the optimal portfolio positions to the input alphas; an angle of 90 degrees reflects that the portfolio has little to do with the manager's investment insights and therefore they argue that the angle should be made more acute.

By defining the relationship between the alpha-weight angle and the condition number of the covariance matrix, Golts and Jones (2009) show that ill-conditioned covariance matrices result in large alpha-weight angles whereas more regularised matrices ensure that the angle remains acute. Therefore, by modifying the condition number of the input covariance matrix during the optimisation process, we are able to constrain the alpha-weight angle to lie within specified levels.

In their methodology, Golts and Jones (2009) propose a theoretical separation of the mean-variance optimisation into two steps. The first step involves specification of the investment direction, and the second, scaling the magnitude based on the imposed constraints. They justify this proposition using the background theory

presented in the following section.

4.1 Separation of Magnitude and Direction

Consider n assets with expected single-period returns (alphas), $\alpha = (\alpha_1, \dots, \alpha_n)$ and the $n \times n$ covariance matrix denoted by Σ . Let $w = (w_1, \dots, w_n)$ denote the portfolio positions generated by the optimisation process. The expected return, r_p , and expected variance, σ_p^2 of the portfolio are respectively:

$$r_p = \sum_{i=1}^n \alpha_i w_i = \alpha' w \quad (4.1)$$

and

$$\sigma_p^2 = \sum_{i,j=1}^n w_i \Sigma_{ij} w_j = w' \Sigma w \quad (4.2)$$

where Σ_{ij} are the elements of the matrix Σ .

Table 4.1 below lists four classical Markowitz -type optimisation problems.

Number & Name	Problem	Solution
I. Risk Constraint	$\max_w r_p$ subject to $\sigma_p^2 \leq \sigma_0^2$	$w = \frac{\sigma_0}{\sqrt{\alpha' \Sigma^{-1} \alpha}} \Sigma^{-1} \alpha$
II. Return Constraint	$\min_w \sigma_p^2$ subject to $r_p \geq r_0$	$w = \frac{r_0}{\alpha' \Sigma^{-1} \alpha} \Sigma^{-1} \alpha$
III. Risk Aversion	$\max_w r_p - \lambda \sigma_p^2$	$w = \frac{1}{2\lambda} \Sigma^{-1} \alpha$
IV. Sharpe Ratio	$\max_w \frac{r_p}{\sqrt{\sigma_p^2}}$	$w = \frac{w' \Sigma^{-1} w}{\alpha' w} \Sigma^{-1} \alpha$

Tab. 4.1: Classical optimisation problems

For all the problems in the table above, it is easy to see that the solution can be written as $w = k \Sigma^{-1} \alpha$, where k is a constant. k can therefore be interpreted to represent the magnitude of the solution whereas $\Sigma^{-1} \alpha$ represents the direction. All the solutions in Table 4.1 above have the same direction, with the magnitude determined by the imposed constraints, that is, σ_0 in problem I, r_0 in problem II, and λ in problem III. Problem IV is magnitude independent. Golts and Jones (2009) thereby define:

1. Investment ‘direction’ by the unit vector, $\hat{w} = w/\sqrt{w'w}$.
2. Investment ‘magnitude’ by a norm of w which could be leverage, $\sum_i |w_i|$, or tracking error, $\sqrt{w'\Sigma w}$.

The magnitude and direction can thus be specified as independent problems, with different procedures used for each problem.

4.2 Relationship between Alpha-weight Angle and Condition Number

The alpha-weight angle is defined as the angle between the input vector of alphas and the output vector of portfolio positions. From the definition of the angle between two vectors:

$$\cos(\phi) = \frac{\alpha'w}{|w||\alpha|} \quad (4.3)$$

where ϕ is the alpha-weight angle. Recall that the condition number of a positive definite matrix Σ is defined as:

$$\text{Cond}(\Sigma) = \frac{\theta_{\max}(\Sigma)}{\theta_{\min}(\Sigma)} \quad (4.4)$$

where θ_{\max} and θ_{\min} are the largest and smallest eigenvalues of Σ respectively.

Bailey (2013) provides a rigorous proof to show that the alpha-weight angle and condition number of the covariance matrix are related by the equation below:

$$\cos(\phi) \geq \frac{\theta_{\max}\theta_{\min}}{\frac{1}{2}(\theta_{\max}^2 + \theta_{\min}^2)} = 2\sqrt{\frac{\kappa}{(\kappa + 1)^2}} \quad (4.5)$$

where κ is the condition number of the covariance matrix.

From equation (4.5) above, we can see that the larger the condition number κ , the larger the alpha-weight angle, ϕ . Therefore, an optimisation procedure that will reduce the condition number of the covariance matrix will result in a more acute alpha-weight angle and consequently, a more stable portfolio.

In order to keep the alpha-weight angle within tight bounds, robust portfolio construction is proposed. This is discussed in the following section.

4.3 Robust Optimisation and the Alpha-weight Angle

In robust optimisation, the uncertainty of expected returns and covariances is taken into consideration directly in the optimisation process, thus creating a more realistic model. This optimisation is done under a ‘worst-case’ scenario and the chosen portfolio is that which will perform well under a number of different scenarios, thereby providing some protection from estimation and model risk. The resultant portfolios are usually conservative, leading to a decrease in performance under the more likely scenarios (Fabozzi *et al.* (2007)). The challenge therefore when creating robust optimisation formulations is to design them cleverly in order to prevent overly conservative results. The first step in robust construction is to define the uncertainty set; in the technique of Golts and Jones (2009) that we shall adopt, this is defined as a spherical region around the input alphas.

4.3.1 Spherical uncertainty region for alphas

The alphas are presumed to be lying in a spherical uncertainty region, U_α , of radius $\chi|\alpha|$, for χ between 0 and 1. A portfolio is chosen to maximise utility in a worst-case scenario for realisation of $\alpha \in U_\alpha$, that is, for a minimum return. The classical mean-variance optimisation problems listed in Table 4.1 can be reformulated as the robust problems in Table 4.2.

Number	Problem
I	$\max_w \min_{U_\alpha} r_p$ subject to $\sigma_p^2 \leq \sigma_0^2$
II	$\min_w \sigma_p^2$ subject to $\min_{U_\alpha} r_p \geq r_0$
III	$\max_w \left(\min_{U_\alpha} r_p - \lambda \sigma_p^2 \right)$
IV	$\max_w \min_{U_\alpha} \frac{r_p}{\sqrt{\sigma_p^2}}$

Tab. 4.2: Robust optimisation problems

For the spherical uncertainty region defined as above, the assets’ single-period

return vector α_0 can be estimated as:

$$\alpha_0 = \alpha + \chi|\alpha|$$

The expected portfolio return, r_p can then be simplified as:

$$\begin{aligned} r_p &= \alpha'_0 w \\ &= (\alpha + \chi|\alpha|)'w \\ &= \alpha'w + \chi'w|\alpha| \end{aligned} \tag{4.6}$$

In order to evaluate the worst-case return, we need to find the minimum value of equation (4.6). Since $\alpha'w$ and $|\alpha|$ are constants, minimising equation (4.6) is equivalent to minimising the value of $\chi'w$. If δ is the angle between χ' and w , we can write:

$$\chi'w = |\chi||w| \cos(\delta)$$

$\chi'w$ is minimised when $\delta=2\pi$ so that $\chi'w = -|\chi||w|$. Since χ is positive (it has been defined to lie between 0 and 1), the worst-case return can be written as:

$$\begin{aligned} \min_{U_\alpha} r_p &= \alpha'w - \chi|w||\alpha| \\ &= |\alpha||w| \cos(\phi) - \chi|w||\alpha| \\ &= |\alpha||w|(\cos(\phi) - \chi) \end{aligned} \tag{4.7}$$

With the worst-case return defined as in equation (4.7) above, the robust optimisation problems in Table 4.2 can now be expressed in a standard form that can be accepted by available numerical optimisers. For example, robust optimisation problem I becomes:

$$\max_w |\alpha||w|(\cos(\phi) - \chi) \text{ subject to } \sigma_p^2 \leq \sigma_0^2 \tag{4.8}$$

For a positive return ($r_p \geq 0$), $\cos(\phi) \geq \chi$ and therefore by specifying the value of χ , we can constrain the angle ϕ to lie between 0 and $\cos^{-1}(\chi)$. The less we trust our alphas, i.e, the bigger the value of χ , the more we force our optimised weights to be closer to them.

4.3.2 Investment direction of robust weights

The robust optimisation problems in Table 4.2 have no closed form solutions generally but can be solved numerically. The investment direction \hat{w} however, can be

expressed as: (Golts and Jones (2009))

$$\hat{w} = \left(\chi \mathbb{I}_N + \frac{(\cos(\phi) - \chi)}{\hat{w}' \Sigma \hat{w}} \Sigma \right)^{-1} \hat{\alpha} \quad (4.9)$$

where $\hat{\alpha} = \alpha/|\alpha|$.

From section 4.1, we defined the investment direction for classical mean-variance optimisation \hat{w} by $\Sigma^{-1}\alpha$. In comparing this to equation (4.9) above, we see that the robust optimisation modifies the input covariance matrix to:

$$\left(\chi \mathbb{I}_N + \frac{(\cos(\phi) - \chi)}{\hat{w}' \Sigma \hat{w}} \Sigma \right)^{-1}$$

The above form is similar to that of a shrinkage estimator. The use of shrinkage methods has been suggested as a way to improve maximum likelihood estimates of the covariance matrix. A general form of a shrinkage estimator is a convex combination of the empirical estimator (Σ) with some suitable target, T :

$$\Sigma(t) = (1 - t)T + t\Sigma$$

In our robust optimisation, this suitable target is the identity matrix, \mathbb{I}_N , and the modified covariance matrix is dependent on the values of χ and α , that is, $\Sigma(\chi, \alpha)$. The robust optimisation procedure therefore shrinks the empirical covariance matrix towards the identity matrix which has a condition number of 1. This shrinkage therefore reduces the condition number of the covariance matrix, resulting in a sharper alpha-weight angle.

4.3.3 Investment magnitude of robust weights

As in the classical optimisation problems, the investment magnitude can be determined by a separate procedure and is determined by desired leverage, tracking error, or target return constraints. In the empirical analysis, the relevant constraints are specified as detailed in chapter 5.

4.3.4 Resultant portfolios

Golts and Jones (2009) show that the resulting portfolios from the robust optimisation exhibited both a reduced *ex-ante* Sharpe ratio and leverage of the input alphas. The reduction in overall portfolio leverage can be explained mathematically as below:

From equations (4.1) and (4.3):

$$r_p = |w||\alpha| \cos(\phi)$$

Given a fixed level of expected portfolio return r_p :

If ϕ is large (making $\cos(\phi) \approx 0$), and the investment direction chosen as \hat{w} , then we can see that r_p will not be generated with sufficient magnitude unless the portfolio is leveraged in the \hat{w} direction. The reverse is true if ϕ is small and therefore we will realise a reduction in leverage for smaller ϕ .

Chapter 5

Empirical Analysis

The empirical analysis involves evaluation of performance and descriptive statistics of portfolios constructed using the below techniques:

Case 1: Classical mean-variance optimisation using expected returns and sample covariance matrix

Case 2: Mean-variance optimisation using sorted returns and sample covariance matrix (as per Almgren and Chriss (2006))

Case 3: Robust optimisation using expected returns and sample covariance matrix (as per Golts and Jones (2009))

Case 4: Robust optimisation using sorted returns and sample covariance matrix

Henceforth, we refer to classical mean-variance optimisation as *simple optimisation*.

5.1 Data and Methodology

Weekly data on shares included in the All Share Index (ALSI) listed on the Johannesburg Stock Exchange (JSE) from January 2002 to December 2013 is used. The backtesting algorithm used is the method of Munro (2010). The *in-sample* period, used to estimate the covariance matrix, is defined to cover 170 weekly returns up to the portfolio construction date whereas the *out-of-sample* period runs from January 2006 to November 2013. This is illustrated in Figure 5.1 below for a portfolio constructed in December 2005.

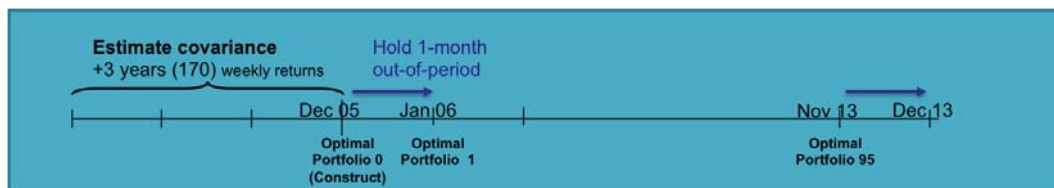


Fig. 5.1: Backtesting methodology (*Source: Munro (2010)*)

We assume *perfect foresight* and use the following period returns as the alpha forecasts. Later, we look at a more realistic scenario when we perform the analysis assuming less than perfect forecast which we quantify using an Information Coefficient (IC) of 0.1.

The optimisation is run from the start of the out-of-sample period and thereafter rebalancing is done monthly. Each month, the 100 largest shares of the ALSI are identified and the relevant weights of which are used as constituents of the benchmark sample against which performance is evaluated (note that in some cases, more than one share may have the same weights, making the number of shares to be used in the sample more than 100). As stated above, we used 170 observations for calculation of the sample covariance matrix. We do this in order to ensure that the number of observations used is greater than the number of shares listed, therefore generating a fairly good approximation. The output of the backtesting process is various out-of-sample fund performance statistics.

Classical mean-variance and robust optimisation is setup using MATLAB's `fmincon` function; particularly, the interior-point algorithm is used. The interior-point algorithm is chosen because it is a large-scale algorithm which can handle large, sparse, as well as small dense problems. The algorithm satisfies bounds at all iterations and can recover from MATLAB's 'NaN' or 'Inf' results. We solve both simple and robust optimisation problems for a specified risk constraint which in our analysis is 4% monthly tracking error and perform the optimisation subject to the portfolio standard deviation equalling this value of tracking error. Additional constraints that we specify are:

- ⇒ χ set to 0.5 for the robust optimisation in order to constrain the value of the alpha-weight angle to be no more than 60 degrees.
- ⇒ Lower limits and upper limits of stock weights set to 0 and 1 respectively.

For cases 2 and 4, a single complete portfolio sort is specified using future returns

data (perfect foresight), and the centroid vector calculated analytically.

For the scenario where we assume less than perfect foresight, we use the ideas presented by Grinold and Kahn (1999) and Ye (2008) to generate new values of alphas that serve as input to the optimisation. Grinold and Kahn (1999) define the *fundamental law of active management* as:

$$IR = IC \cdot \sqrt{BR} \quad (5.1)$$

where

- ⇒ IR is the *Information Ratio*, defining the amount of portfolio excess return (or active return) generated by excess risk taken relative to a benchmark
- ⇒ BR is the strategy's *Breadth*, defined as the number of independent investment decisions made each year
- ⇒ IC is the *Information Coefficient*, measuring the manager's skill. It is the correlation of each forecast with the actual returns

Ye (2008) goes on to generalise this law by stating that the active return is not only dependent on the breadth and signal quality (measured by the IC), but also by variation of the signal quality and constraints. She provides insight on how managers should select portfolios depending on whether their signal quality is stable or varied from one period to the next. For the empirical analysis, we assume a constant IC of 0.1 for the less-than-perfect-foresight scenario and use the below formula to generate the alpha forecasts: (Grinold and Kahn (1999))

$$\alpha = IC \cdot [IC \cdot \theta + \omega \cdot \sqrt{1 - IC^2} \cdot z] \quad (5.2)$$

where:

- ⇒ θ is the residual return
- ⇒ ω is the standard deviation of the residual returns
- ⇒ z is a random number with mean 0 and variance 1

From equation (5.2) above, we see that when IC=1, that is, the manager has perfect foresight, the alphas are exactly equal to the residual returns.

For cases 2 and 4 where we are replacing the alpha values with the centroid, we first calculate the new alpha value for the specified IC level and then use the resulting value to calculate the centroid vector.

5.2 Portfolio Performance Statistics

We explore various out-of-sample performance statistics for the resultant fund in detail in this section. We measure the risk and return statistics of the fund, as well as the risk-adjusted return statistics (Sharpe and Information ratios). In each of the scenarios, fund risk is measured as the standard deviation of the returns. Recall that the benchmark is defined monthly as the Top 100 ALSI shares.

5.2.1 Return statistics

Figure 5.2 and Table 5.1 show the resultant cumulative returns and outperformance (relative to benchmark) statistics respectively. The annualised outperformance is measured as:

$$\text{AnnOutperformance} = \text{AnnFundRets} - \text{AnnBmarkRets} \quad (5.3)$$

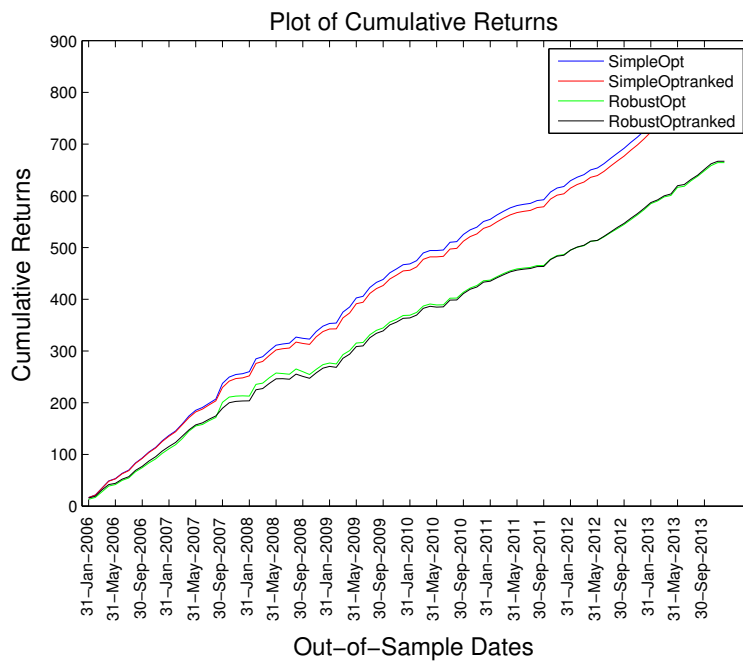


Fig. 5.2: Cumulative returns

Portfolios generated using simple optimisation (case 1 and 2) consistently outperform those generated using robust optimisation (case 3 and 4). As discussed in chapter 4, robust optimisation is done for a worst-case realisation of the input parameters (minimum value of expected return in our case) and so this result is hardly

surprising. Furthermore, we see that the difference in performance of portfolios generated using actual returns (case 1 and 3) and ordering information (case 2 and 4) is marginal. This can be explained by the fact that the efficient portfolios generated are similar. Almgren and Chriss (2006) proved this for mean-variance optimisation and our results show that the same is true for robust optimisation. Additionally, we see that in contrast to the simple portfolios, the robust portfolios from sorts outperform the robust portfolios generated using actual returns.

	Out-performance
Case 1	169.0641
Case 2	164.0975
Case 3	119.7424
Case 4	120.9949

Tab. 5.1: Annualised outperformance

5.2.2 Risk statistics

The fund risk is measured as the standard deviation of the realised returns whereas the tracking error (TE) is measured as the standard deviation of the realised returns over the benchmark. Generally, we see that the portfolios generated using robust optimisation are less risky. Additionally, in comparing portfolios using actual returns and ordering information, we note interestingly that portfolios from sorts exhibit lower risk statistics.

	Total Risk		TE
Case 1	19.3931	Case 1	5.5691
Case 2	18.6654	Case 2	5.3601
Case 3	19.0065	Case 3	5.4581
Case 4	17.2333	Case 4	4.9488

(a) Total risk

(b) Monthly tracking error

Tab. 5.2: Fund risk statistics

5.2.3 Risk-adjusted return statistics

We measure the fund's risk-adjusted performance statistics using their Sharpe and Information Ratios. The Sharpe Ratio is the most commonly used measure of risk-

adjusted performance. It measures the excess return (or risk premium) over the risk-free rate per unit of risk and is used to identify the reward an investor would expect from investing in a risky asset. Generally, the higher the Sharpe ratio, the better the risk-adjusted fund performance. For the analysis we used the yield of a 5-year South African government bond as a proxy for the risk-free rate of return and used the formula below to calculate the Sharpe ratio:

$$\text{Sharpe Ratio} = \frac{\mathbb{E}[R_p - R_f]}{\sqrt{\text{Var}(R_p)}} \quad (5.4)$$

where:

⇒ R_p and R_f denote the portfolio return and risk free rate of return respectively

⇒ $\text{Var}(\cdot)$ denotes the variance

⇒ $\mathbb{E}[\cdot]$ denotes the expected value

The Information Ratio (IR) on the other hand, defines the amount of portfolio excess return(or active return) generated by excess risk taken relative to the benchmark(or tracking error). We calculate the information ratio as:

$$\text{Information Ratio} = \frac{\mathbb{E}[R_p - R_b]}{\sqrt{\text{Var}(R_p - R_b)}} \quad (5.5)$$

where:

⇒ R_p and R_b denote the portfolio return and benchmark return respectively

⇒ $\text{Var}(\cdot)$ denotes the variance

⇒ $\mathbb{E}[\cdot]$ denotes the expected value

As detailed in Table 5.3, the robust portfolios exhibit lower risk-adjusted performance as observed from the lower Sharpe and Information ratios. The lower Sharpe ratios are in-line with Golts and Jones (2009) research where they observe that the robust optimisation procedure generally results in a reduction in the Sharpe ratio. Interestingly, just like the risk statistics, we observe that the portfolios from sorts exhibit better risk-adjusted performance.

	Sharpe Ratio	Information Ratio
Case 1	5.4028	8.7635
Case 2	5.4994	8.8377
Case 3	4.3637	6.3331
Case 4	4.8322	7.0579

Tab. 5.3: Sharpe and Information ratios(annualised)

5.2.4 Beta

Beta (β) is a measure of volatility of a portfolio in comparison to a defined benchmark (the benchmark could be the entire market). We calculate the fund beta using the equation below:

$$\beta = \frac{\text{Cov}(R_p, R_b)}{\text{Var}(R_b)}$$

where R_b is the benchmark return and $\text{Cov}(\cdot)$ and $\text{Var}(\cdot)$ represent the covariance and variance respectively.

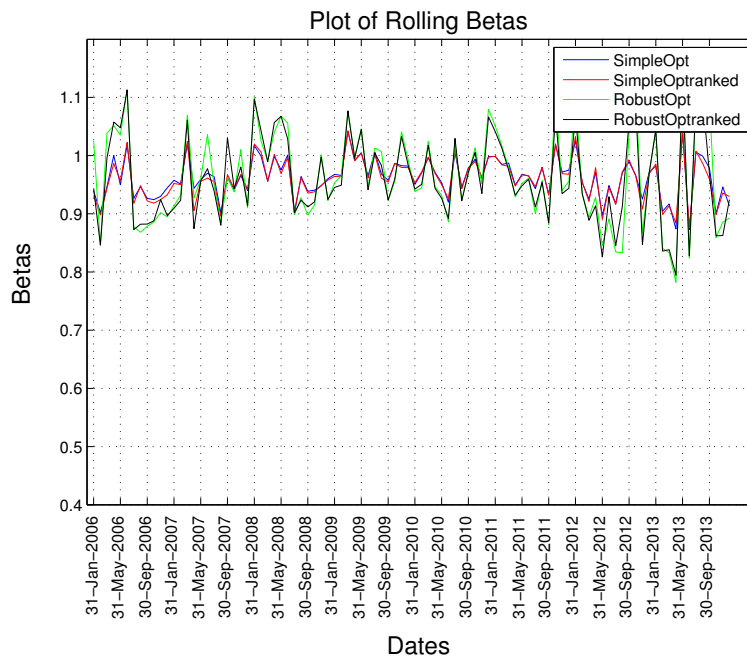


Fig. 5.3: Realised Beta

In Figure 5.3 the *in-sample* period is used as the rolling window. The betas

for all scenarios generally oscillate around 0.95, indicating that the volatility of the portfolios is approximately equivalent to the benchmark volatility. However, we notice that the movements in the betas for cases 3 and 4 are greater than for cases 1 and 2, suggesting that the betas for the simple portfolios are slightly more stable.

5.3 Condition Number and Alpha-Weight Angle

As discussed in chapter 4, the robust optimisation technique of Golts and Jones (2009) leads to a modification of the condition number of the input covariance matrix, and consequently, the resulting alpha-weight angle. We compare these values for both the simple and robust optimisation scenarios.

5.3.1 Condition number

We calculate the condition number using equation (4.4). Figure 5.4 to 5.7 show the sorted eigenvalues and condition numbers for each of the scenarios on the last rebalance date (note the difference in scale for simple and robust optimisation in the figure). Robust optimisation dramatically reduces the ratio of the largest to smallest eigenvalues, leading to a better conditioned covariance matrix (the condition number for robust optimisation is more than 600 times smaller than that for simple optimisation!). We also note that the condition number for case 4 is significantly lower than for the rest.

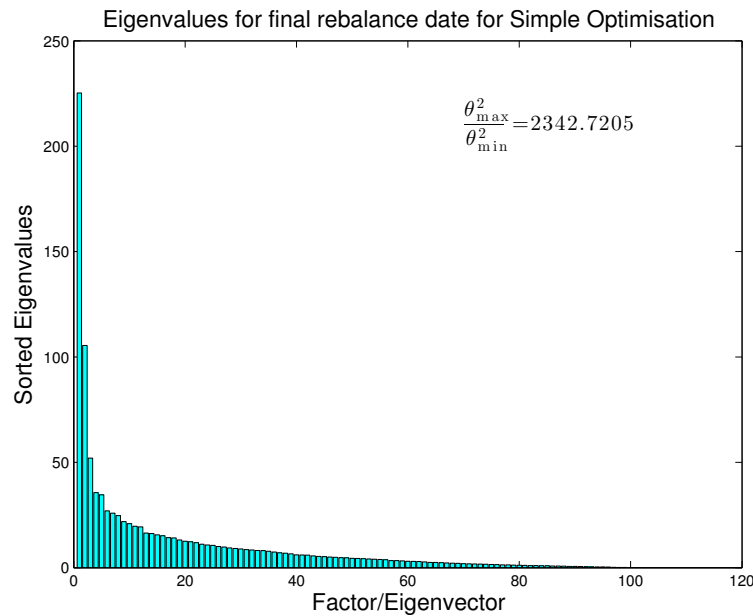
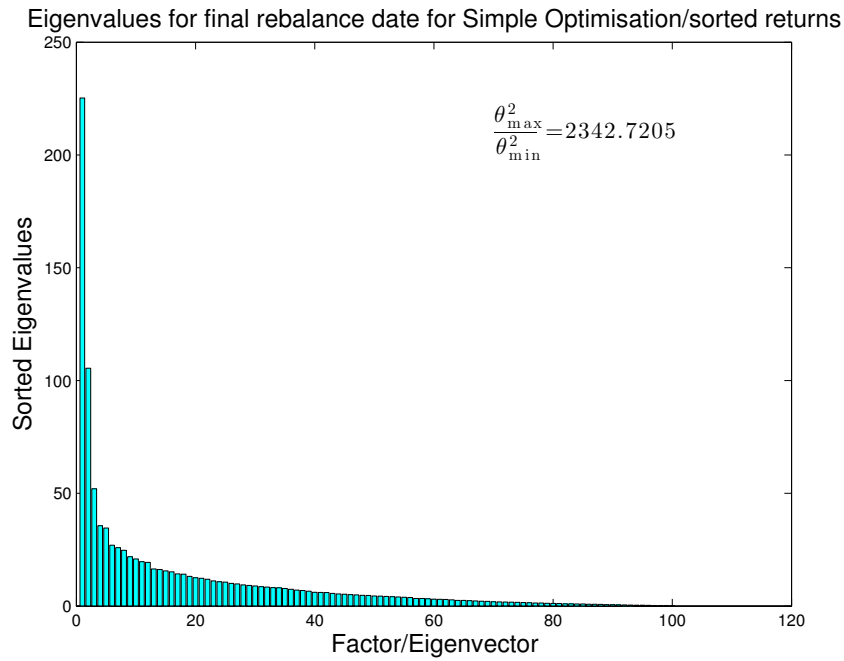
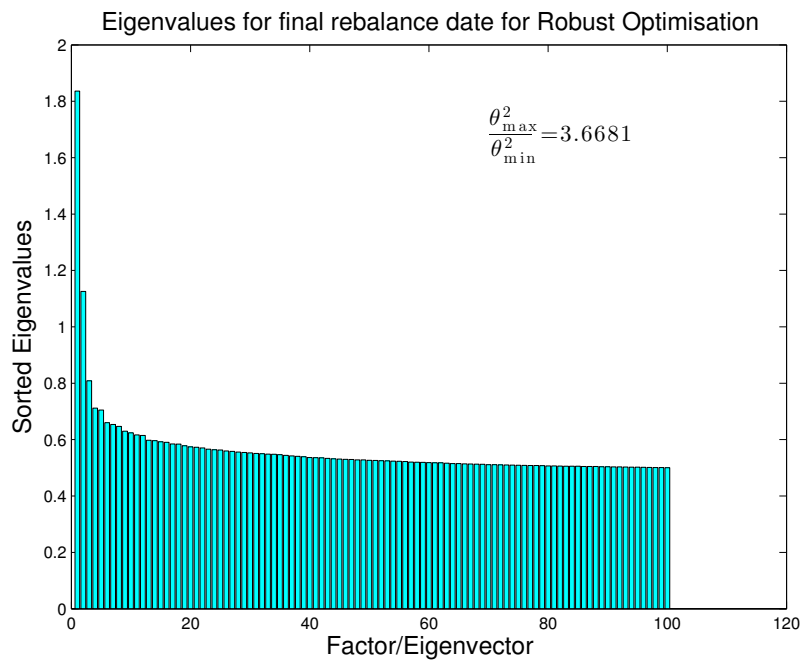


Fig. 5.4: Eigenvalues: Case 1

**Fig. 5.5:** Eigenvalues: Case 2**Fig. 5.6:** Eigenvalues: Case 3

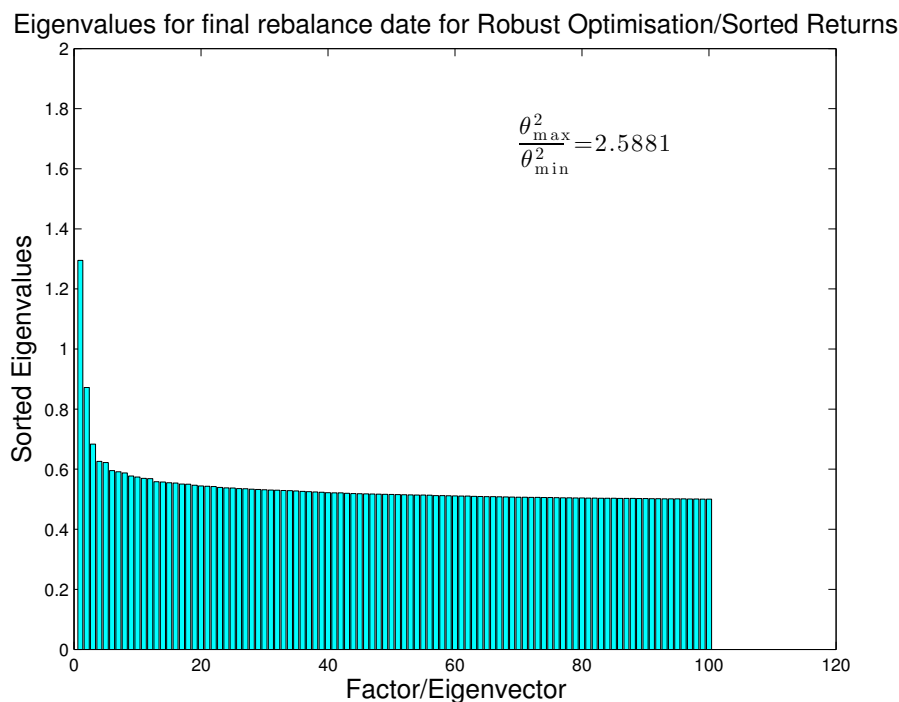
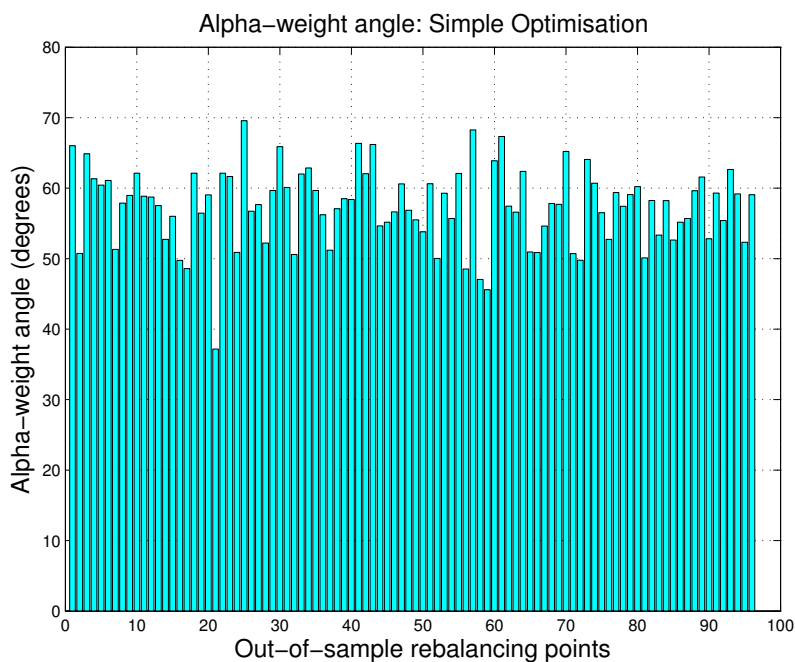


Fig. 5.7: Eigenvalues: Case 4

5.3.2 Alpha-weight angle

The alpha-weight angle is calculated using equation (4.3). For cases 2 and 4, the alphas are replaced by the centroid vector in the calculation.

In line with Golts and Jones (2009) research, we observe that the robust optimisation indeed leads to a reduction in the alpha-weight angle. Furthermore, we notice that the alpha-weight angles for the portfolios generated using Almgren and Chriss (2006) technique in case 4 are more acute - indicating that the investment weights generated better reflect the input alphas. A possible explanation for this can be derived from the discussion in section 4.3.2 where we noted that the modified covariance matrix from the robust optimisation is dependent on the values of χ and α . When using sorts, the alphas are replaced by the centroid, the magnitude of which is less than the magnitude of the alphas. This could be a contributing factor to this result.

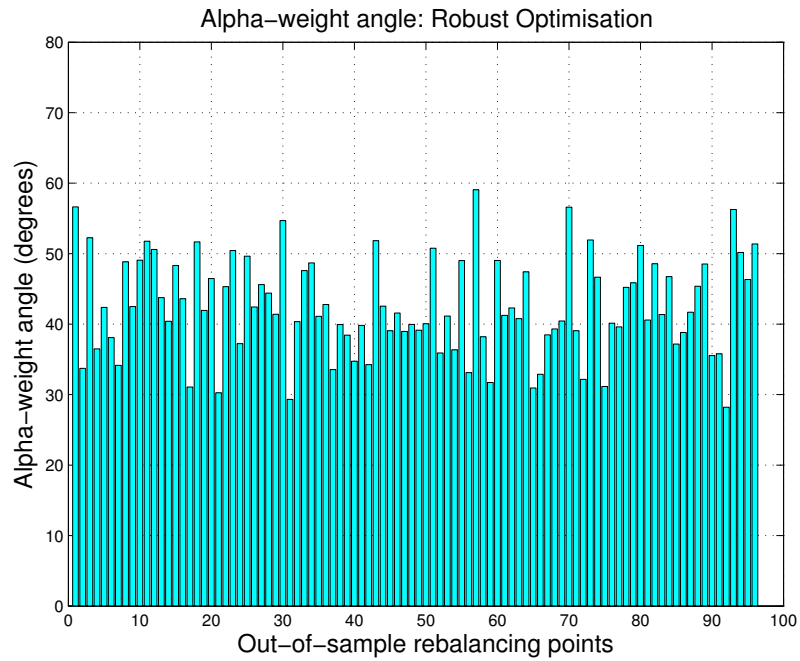


(a) Case 1

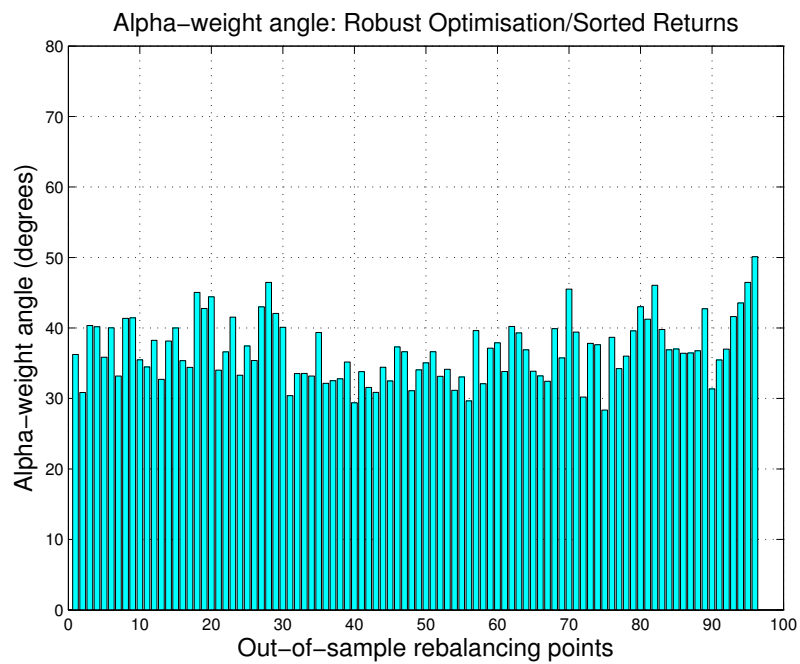


(b) Case 2

Fig. 5.8: Alpha-weight angle for simple optimisation



(a) Case 3



(b) Case 4

Fig. 5.9: Alpha-weight angle for robust optimisation

5.4 Effect of Changes in χ

In cases 3 and 4, the value of χ is specified in order to restrict the alpha-weight angle to a pre-determined level. For the analysis thus far, we have set χ to 0.5. The case of χ equal to 0 is equivalent to simple optimisation; we therefore observe the fund performance statistics for χ ranging from 0.25 to 1. Varying the value of χ in this way is equivalent to increase in the size of the uncertainty region and in essence, increase in ‘robustness’ of the portfolios.

We realise a reduction in both risk and outperformance with increasing values of χ . In other words, the resultant portfolios become more conservative as we increase the size of the uncertainty set. As highlighted in section 2.2, Tütüncü and Koenig (2003) suggested a trade-off when using robust optimisation based on the risk appetite of investors; in line with this, an optimal value of χ may need to be selected.

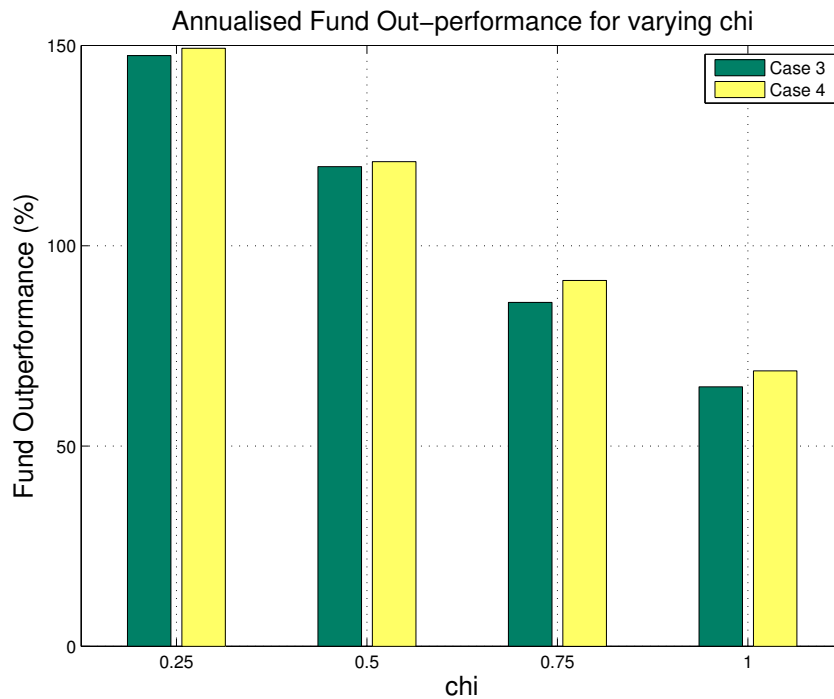


Fig. 5.10: Outperformance statistics for varying values of χ

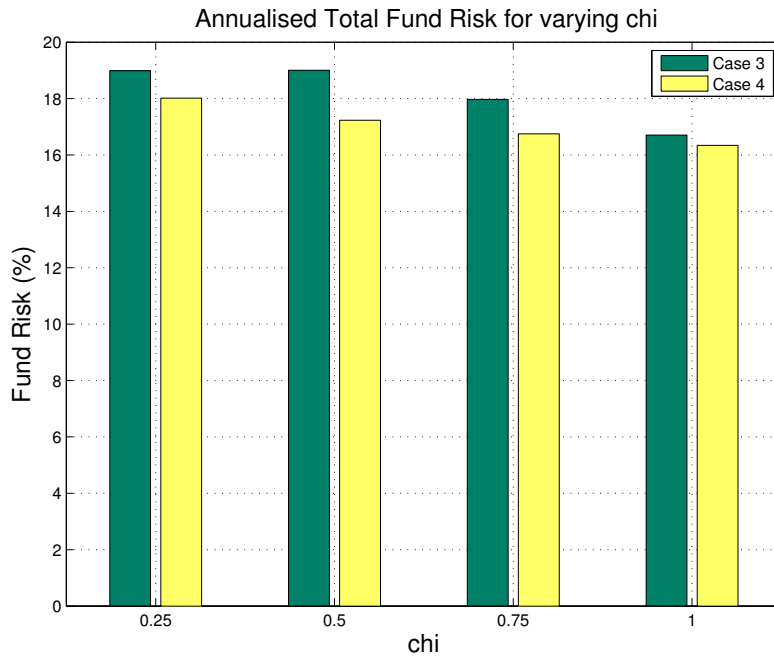


Fig. 5.11: Fund Risk statistics for varying values of χ

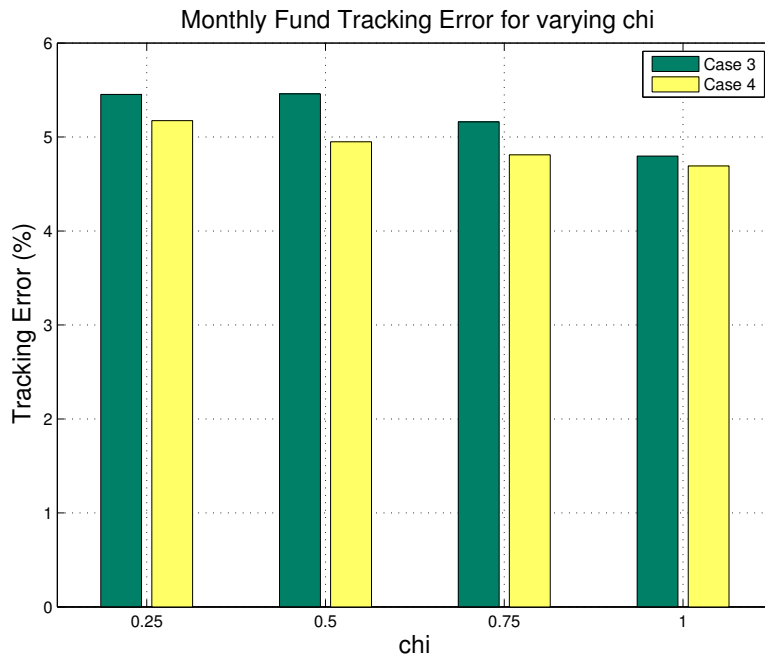


Fig. 5.12: Tracking Error statistics for varying values of χ

5.5 Portfolio Diversification and Concentration

We gain insight into the descriptive statistics of the resultant fund using different measures of portfolio diversification and concentration. Whereas diversification is often used as a strategy to minimise risk, concentration is used to enhance returns. In this section, we explore the out-of-sample diversification/concentration statistics of the fund.

5.5.1 Number of stocks in portfolio

According to research by Statman (1987), a well diversified stock portfolio consists of at least 30 stocks for a borrowing investor and 40 stocks for a lending investor. Table 5.4 shows the average number of stocks for each of the scenarios generated.

Based on this definition of diversification, all scenarios result in well-diversified portfolios; however, we observe that the robust portfolios are more diversified than the simple portfolios. Also, there is a significant increase in the diversification of the portfolios in going from actual returns to the use of sorted returns.

	Avg. Num. Stocks
Case 1	72.9688
Case 2	99.6458
Case 3	88.1875
Case 4	100.3646

Tab. 5.4: Average number of stocks

We also looked at the average number of stocks with weights greater than 0.5% for each of the scenarios. The results are displayed in Table 5.5.

	Avg. Num. Stocks with Weight > 0.5%
Case 1	17
Case 2	17
Case 3	20
Case 4	21

Tab. 5.5: Average number of stocks with weights greater than 0.5%

5.5.2 Herfindahl-Hirschman Index (HHI)

The Herfindahl-Hirschman Index (HHI) is a commonly accepted measure of market concentration calculated by evaluating the sum of the square of investment weights of each share in a portfolio.

$$HHI = \sum_{i=1}^n w_i^2$$

The HHI number can range from 0 to close to 1 with lower values of HHI indicating well diversified portfolios.

In line with the results in section 5.5.1 above, as per the HHI measure, the robust portfolios are more diversified than the simple optimised portfolios.

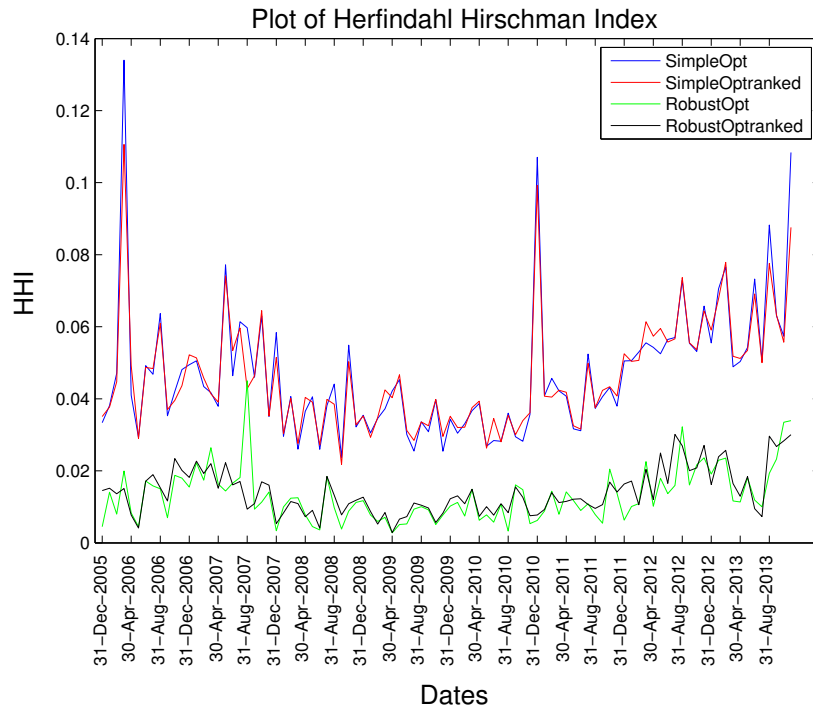


Fig. 5.13: Herfindahl-Hirschman Index (HHI)

5.5.3 Leverage

We measure the portfolio leverage as the sum of the absolute value of the active portfolio weights.

$$\text{Leverage} = \sum |(Active\ weights)_i|$$

From Figure 5.14 and Table 5.6 we see that the robust portfolios exhibit lower leverage with a reduction of over 30% for both cases of actual and sorted returns. We also notice that the portfolios from sorts have slightly higher leverage than the corresponding mean-variance/robust portfolios. In section 4.3.4 we explained mathematically why the robust portfolios exhibit lower leverage. When using sorted returns, the magnitude of the centroid vector is much less than that of the actual returns and this contributes to an increase in leverage of these portfolios.

	Avg. Leverage
Case 1	1.1078
Case 2	1.1314
Case 3	0.7043
Case 4	0.7672

Tab. 5.6: Average leverage

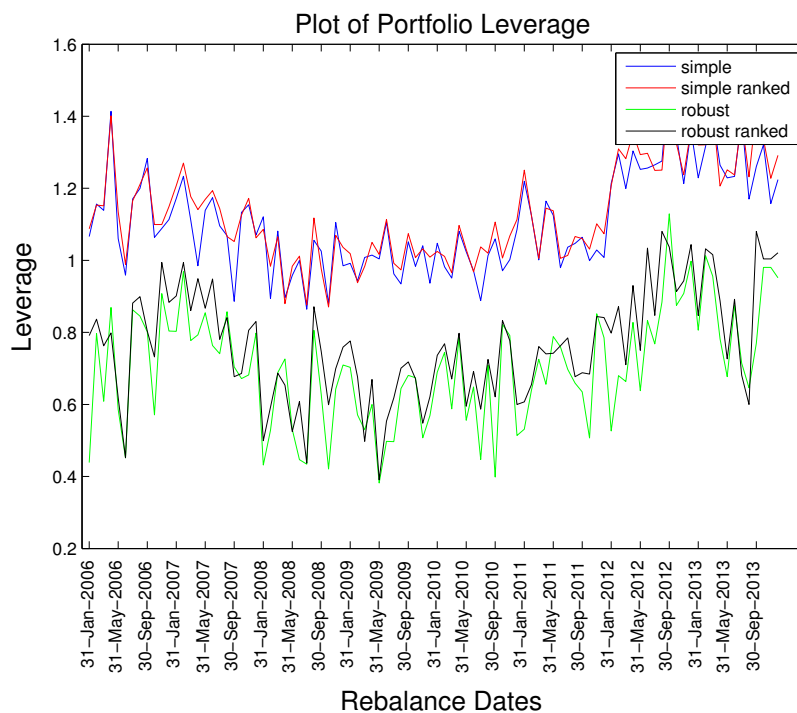


Fig. 5.14: Portfolio leverage

5.5.4 Comparison of Portfolio Weights

We compare the difference in portfolio weights for the simple (case 1 and 2) and robust (Case 3 and 4) scenarios by measuring the sum of the square of the difference in investment weights of the two portfolios.

$$\sum (w_1 - w_2)^2$$

In Figure 5.15 we compare this measure for the simple (case 1 and 2) and robust portfolios (case 3 and 4). For the robust portfolios, this measure is approximately zero, showing that the use of ordering information hardly modifies the investment weights when using robust optimisation.

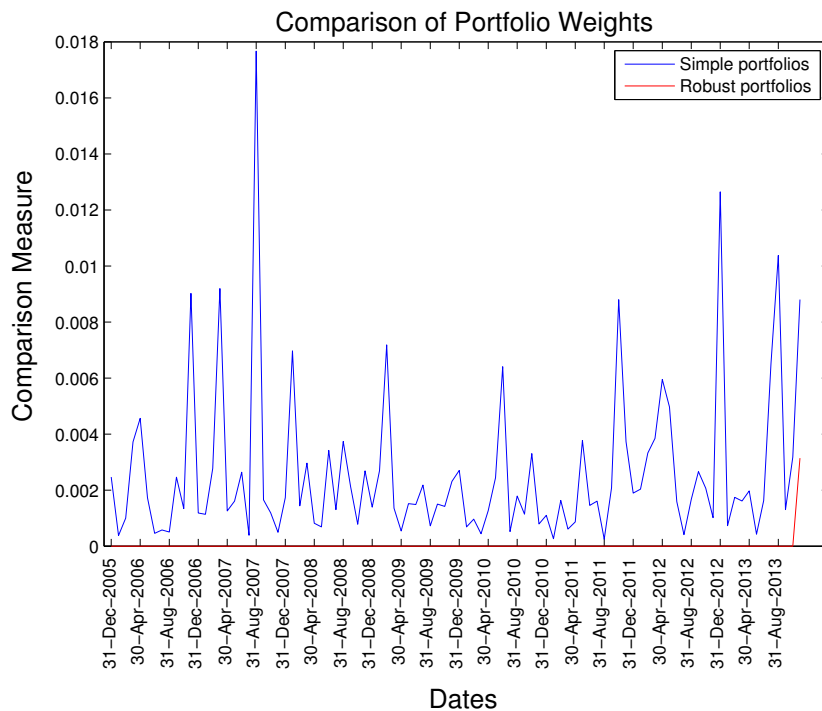


Fig. 5.15: Comparison of Portfolio Weights

5.5.5 Comparison of Change in Portfolio Weights After Rebalancing

In the preceding sections, we highlighted that robust portfolio construction leads to more stable portfolios than classical mean-variance optimisation. In order to demonstrate this stability, the turnover of the resultant portfolios over time will need to be

	Δ Robust Weights $>$ Δ Simple Weights
Month 1 to 2	49 %
Month 2 to 3	44 %
Month 3 to 4	39 %
Month 4 to 5	41 %
Month 5 to 6	40 %
Month 6 to 7	37 %
Month 7 to 8	52 %
Month 8 to 9	44 %
Month 9 to 10	31 %
Month 10 to 11	42 %
Month 11 to 12	41 %

Tab. 5.7: Ad-hoc Calculation of Turnover

measured. Portfolio turnover is a measure of how frequently assets within a fund are bought and sold. It is calculated by taking either the total amount of new securities purchased or the amount of securities sold - whichever is less - over a particular period, divided by the total net asset value (NAV) of the fund. It is usually reported for a 12-month time period. Generally, funds with lower turnover rates are desirable as this translates into lower transaction fees.

Due to limitations in the data set used for this analysis (such as lack of information on the individual stock prices), we did not calculate turnover as above but used a more ad-hoc approach to demonstrate stability. We looked at changes in the portfolio weights at each rebalance date over the first year for both the simple (case 1) and robust portfolios (case 3). We computed this change month on month and for each period used the frequentist probability approach to check the percentage of time the change in the robust portfolio weights was greater than that of the simple portfolio. This data is shown in Table 5.7. This shows that on average, the change in weights of robust portfolios is greater than that for simple portfolios 42% of the time. This figure is borderline on convincing that the robust portfolios are more stable - more detailed analysis will need to be done before any reasonable conclusions can be made. Particularly an updated data set will be required in order to correctly compute the turnover. The analysis period also needs to be long enough to make reasonable conclusions.

5.6 Performance Statistics for Varying IC

Thus far, we have assumed perfect foresight of an investment manager and therefore used a constant IC of 1. In this section we assume a more realistic IC of 0.1 and use equation (5.2) to calculate the new alphas. For each of the scenarios, we evaluate the performance statistics of the resultant portfolios. As in section 5.2, the benchmark is defined monthly as the Top 100 ALSI shares. It should be noted that only one simulation was done and so the results may be susceptible to bias.

5.6.1 Risk statistics

As in section 5.2.2, we use standard deviation as a measure of risk. For all scenarios, we do not realise a significant reduction in the risk of the portfolios when the IC is decreased to 0.1. The trend is still the same as for the perfect foresight case, although more significant, with the robust portfolios being less risky.

	IC=0.1	IC=1		IC=0.1	IC=1
Case 1	5.0962	5.5691	Case 1	17.7466	19.3931
Case 2	5.1004	5.3601	Case 2	17.7612	18.6654
Case 3	4.7687	5.4581	Case 3	16.6059	19.0065
Case 4	4.7387	4.9488	Case 4	16.5016	17.2333

(a) Tracking error

(b) Total risk

Tab. 5.8: Risk statistics for varying IC

5.6.2 Outperformance relative to benchmark

We use equation (5.3) to measure the annualised fund outperformance relative to the benchmark.

	IC=0.1	IC=1
Case 1	17.0657	169.0641
Case 2	17.5586	164.0975
Case 3	16.5783	119.7424
Case 4	16.7847	120.9949

Tab. 5.9: Annualised fund outperformance for varying IC

We notice a drastic degradation in performance for an IC of 0.1 for all scenarios. However the difference in performance between the simple and robust portfolios becomes less significant with the reduced IC. Additionally, we see that the portfolios from sorts slightly outperform portfolios generated using actual returns for both cases of simple and robust optimisation.

5.6.3 Sharpe and Information ratio

Sharpe and Information ratios for all scenarios are reduced drastically for IC=0.1. The general reduction in information ratios for all scenarios is in line with the fundamental law of active management defined by Grinold and Kahn (1999) - see equation (5.1). If the number of stocks in the portfolio is assumed to be the breadth, then the IR for an IC of 0.1 should approximately be 1 for 100 stocks.

Furthermore, in contrast to the case of perfect foresight, the Sharpe and Information ratios for the robust optimisation are greater than those for the simple optimisation. This shows that in realistic situations, portfolios generated using the robust optimisation should exhibit better risk-adjusted performance than mean-variance portfolios.

	IC=0.1	IC=1
Case 1	0.9667	8.7635
Case 2	0.9938	8.8377
Case 3	1.0036	6.3331
Case 4	1.0225	7.0579

(a) Information ratio

	IC=0.1	IC=1
Case 1	0.9776	5.4028
Case 2	1.0009	5.4994
Case 3	1.0072	4.3637
Case 4	1.0231	4.8322

(b) Sharpe ratio

Tab. 5.10: Risk-adjusted performance statistics for varying IC

5.7 Summary of Pros and Cons of Each Technique

Generally, simple optimisation is relatively straightforward, simple to implement, and less computationally resource-intensive than robust optimisation. The introduction of ordering information further complicates the implementation and so a trade-off between performance and complexity will have to be made.

For the analysis based on an IC of 1, we see that the portfolios from sorts are superior in terms of risk-adjusted performance. Particularly, the simple portfolios

from sorts exhibit the best risk-adjusted performance. Also, we see the portfolios from sorts emerging superior in terms of diversification and concentration. However, in this case, robust portfolios are the most diversified.

In using a more realistic IC of 0.1, we see a change in the statistics with the robust portfolios from sorts emerging superior in terms of risk-adjusted performance. However, the differences in these statistics for all the cases are rather marginal. As highlighted above, the implementation of robust portfolio construction is rather complex and the use of sorting signatures further adds to this complexity. In looking at the more realistic scenario with $IC=0.1$, we realise marginal differences in risk-adjusted performance among all scenarios which brings up the question: “Is the implementation of case 4 is warranted taking into consideration its complexity?”. In fact, the performance of portfolios in case 2 seems reasonably good, given its simplicity. Perhaps, in order to get slightly better performance, case 3 would be an even better approach to adopt than case 4.

Chapter 6

Conclusion

This dissertation has explored a new and exciting area of research in creation of portfolios from sorts using the robust optimisation technique introduced by Golts and Jones (2009). As is evident from the empirical analysis, the difference in performance of robust portfolios constructed using ordering information and actual returns is at most, marginal. The values of the performance and descriptive statistics of these portfolios consistently lie within the same range. This presents a huge opportunity for investment managers as generation of reliable estimates for expected returns is often a complicated and resource-intensive task. The opportunity therefore to use ordering information in place of expected returns in a robust optimisation setting without degradation of performance is a promising addition.

Furthermore, additional benefits do exist from using ordering information in a robust setting. The data necessary to create sorts for expected returns is usually readily available and easy to calculate making the estimation process of the inputs to the optimisations relatively simpler and more cost-effective. Moreover, the portfolios generated from sorts have been shown to *slightly* outperform robust portfolios generated using actual returns in terms of return, risk, and risk-adjusted performance. These portfolios are also more diversified, intuitive, and less sensitive to changes in the input parameters as seen from the significantly lower values of the condition number and resulting alpha-weight angle.

In moving from an ‘ideal’ world to a more realistic scenario defined by an information coefficient of 0.1, we see the robust portfolios exhibiting superior risk-adjusted performance. Their return statistics are almost at par with those of the simple portfolios, making their risk adjusted performance better owing to the lower risk statistics. Higher Sharpe ratios for the robust portfolios in this scenario is in contrast to Golts and Jones (2009) results where the reduction in Sharpe ratios is highlighted as a consequence of the use of robust optimisation. Additionally, for both cases of

simple and robust optimisation, we see the portfolios from sorts now outperforming portfolios generated using actual returns. In real-world practical cases, we should therefore expect portfolios generated using ordering information in a robust optimisation setting as we defined, to exhibit superior performance.

In relation to Case 3 (robust optimisation using expected returns), our results vastly differ from those of Bailey (2013). This is due to the use of completely different methodologies in the specification of the input alphas and limit constraints. While we assumed perfect foresight in the specification of the alphas, Bailey (2013) used ‘convexity’ of the relationship between SA stocks and EU equity market as an alpha (basically trying to find an optimal portfolio of stocks that had the best upside to downside relationship with the EU equity market). The lower and upper limits were set to 0 and 1 respectively, except for the cases where sufficient data was not available to calculate a stock’s convexity for that point in time, in which case the limits were set to the benchmark weight. Additionally, we used more updated data in our research (up to November 2013) whereas Bailey (2013) used data up to July 2012.

An area of further research could be in the enhancement of the robust formulation in order to generate less conservative results as we noticed a fairly significant decrease in the outperformance relative to the benchmark in going from simple to robust optimisation. Additionally, this analysis could be repeated with more complicated types of sorts used in the formulation.

Bibliography

- Almgren, R. and Chriss, N. (2006). Optimal portfolios from ordering information, *Journal of Risk* **9**(1): 1–47.
- Bailey, G. (2013). *Robust portfolio construction: Controlling the alpha-weight angle*, Master's thesis, University of Cape Town.
- Banz, R. W. (1981). The relationship between return and market value of common stocks, *Journal of Financial Economics* pp. 3–18.
- Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios, *The Journal of Finance* **XXXII**(3): 663–682.
- Blom, G. (1958). *Statistical Estimates and Transformed Beta Variables*, John Wiley & Sons, Inc.
- Ceria, S. and Stubbs, R. A. (2006). Incorporating estimation errors into portfolio selection: Robust portfolio construction, *Research paper*, Axioma.
- Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice, *Journal of Portfolio Management* pp. 6–11.
- Costa, O. L. V. and Paiva, A. C. (2000). Robust portfolio selection using linear-matrix inequalities, *Journal of Economics Dynamics and Control* pp. 1–21.
- Fabozzi, F. J., Kolm, P. N., Pachamanova, D. A. and Focardi, S. M. (2007). *Robust Portfolio Optimization and Management*, John Wiley & Sons, Inc.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns, *The Journal of Finance* **XLVII**(2): 427–465.
- Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems, *Mathematics of Operations Research* **28**(1): 1–38.
- Golts, M. and Jones, G. (2009). A sharper angle on optimization. Working Paper.
- Grinold, R. C. and Kahn, R. N. (1999). *Active Portfolio Management*, McGraw-Hill.
- Halldórsson, B. V. and Tütüncü, R. H. (2003). An interior-point method for a class of saddle-point problems, *Journal of Optimization Theory and Applications* **116**(3): 559–590.

- Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* **XLVIII**(1): 65–90.
- Jorion, P. (1985). International portfolio diversification with estimation risk, *Journal of Business* **58**(3): 259–278.
- Lamont, O., Polk, C. and Saá-Requejo, J. (2001). Financial constraints and stock returns, *The Review of Financial Studies* **14**(2): 529–554.
- Markowitz, H. (1952). Portfolio selection, *The Journal of Finance* **7**(1): 77–91.
- Merton, R. C. (1980). On estimating the expected return on the market, *Journal of Financial Economics* **8**: 323–361.
- Michaud, R. O. (1989). The markowitz optimization enigma: Is ‘optimized’ optimal?, *Financial Analysts Journal* **45**(1): 31–42.
- Munro, B. (2010). Covariance estimation for south african equity portfolios, *Presentation*, Cadiz.
- Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings’ yields and market values, *Journal of Financial Economics* **9**(1): 19–46.
- Satchell, S. and Scowcroft, A. (2003). *Advances in Portfolio Construction and Implementation*, Butterworth - Heinemann Finance.
- Statman, M. (1987). How many stocks make a diversified portfolio?, *Journal of Finance and Quantitative Analysis* **22**(3).
- Thorp, E. (2003). A perspective on quantitative finance: Models for beating the market, *Quantitative Finance Review* pp. 33–38.
- Tütüncü, R. H. and Koenig, M. (2003). Robust asset allocation.
- Vassalou, M. and Xing, Y. (2005). Abnormal equity returns following downgrades. Draft.
- Ye, J. (2008). How variation in signal quality affects performance, *Financial Analysts Journal* **64**(4): 48–61.

Appendix A

Derivation of Formula for Computation of Centroid

Let x be an n -vector of independent samples from a distribution with density $f(x)$ and cumulative distribution function $F(x)$. The density is assumed to be standard Gaussian making the density of x spherically symmetric. Let y be the vector consisting of the components of x sorted in descending order. Then, the density of the j th component $y_{j,n}$ is

$$\text{Prob}\{w < y_{j,n} < w + dw\} = \frac{n!}{(j-1)!(n-j)!} F(w)^{n-j} (1-F(w))^{j-1} f(w) dw$$

The centroid component $c_{j,n}$ is the mean of this distribution:

$$\begin{aligned} c_{j,n} &= \frac{n!}{(j-1)!(n-j)!} \int_{-\infty}^{\infty} w F(w)^{n-j} (1-F(w))^{j-1} f(w) dw \\ &= \frac{n!}{(j-1)!(n-j)!} \int_0^1 F^{-1}(z) z^{n-j} (1-z)^{j-1} dz \\ &= \mathbb{E}_g(F^{-1}(z)) \end{aligned}$$

where $\mathbb{E}_g(\cdot)$ denotes expectation under the probability density

$$g(z) = \frac{n!}{(j-1)!(n-j)!} z^{n-j} (1-z)^{j-1}$$

When j and n are large, this distribution is narrow. Thus, reasonable approximations to the integral are either $F^{-1}(z_{\text{mean}})$ or $F^{-1}(z_{\text{max}})$, where the mean and the peak of the distribution are

$$z_{\text{mean}} = \frac{n-j+1}{n+1}, z_{\text{max}} = \frac{n-j}{n-1}$$

For the normal distribution, these formulas are special cases, with $\alpha=0,1$, of “Blom’s approximation” (Blom (1958)).

$$c_{j,n} \approx \mathcal{N}^{-1} \left(\frac{n+1-j-\alpha}{n-2\alpha+1} \right)$$

Blom (1958) shows that the values of $\alpha = 0.33$ and 0.50 provide lower and upper bounds for the true value of $c_{j,n}$, and he suggests that $\alpha = 0.375$ is a reasonable approximation for all values of j, n . However, by comparison with numerical

integration, Almgren and Chriss (2006) found that an excellent approximation is $\alpha = A - Bj^{-\beta}$, with $A = 0.4424$, $B = 0.1185$, and $\beta = 0.21$. This gives centroid components with maximum fractional error of less than 0.5% when n is very small, decreasing rapidly as n increases.

Appendix B

Matlab Code

```
%Analytical Calculation of Centroid Vector

[~,ix]=sort(Alphas,'descend');

A=0.4424;

B=0.1185;

Beta=0.21;

n=size(stCov(1,1).fCovFull,2);

alpha_centroid=A-B*n^(-Beta);

Centroid=norminv((n+1-(1:n)-alpha_centroid)/(n-2*alpha_centroid+1));

Centroid(ix)=Centroid;
```

```

%Robust Portfolio Optimisation

function [strOut, pStats, pWts, checkVar, pVol, pExpRet] =
Robust(vecExpRets, VCV, Ann, strConstraints, targetAnnVol,
strBlnOpts)

if strBlnOpts.blnDisplay == 1
    tic;
end

nAssets = size(VCV,2);

fminconOptions = optimset('TolFun',1e-8,'TolX',1e-
8,'Algorithm','interior-point','MaxFunEvals',1e8,'MaxIter',2000,
'display','off');

if isempty(targetAnnVol)
    nFrontierPts = 500;
else
    nFrontierPts = size(targetAnnVol,2);
    if strBlnOpts.blnInclMaxVol == 1
        nFrontierPts = nFrontierPts + 1;
    end
end

pWts = zeros(nFrontierPts, nAssets);
pVol = zeros(nFrontierPts, 1);
pExpRet = zeros(nFrontierPts, 1);
pAngle = zeros(nFrontierPts, 1);
Aeq = ones(1,nAssets);

% Budget constraint -- active bets must sum to zero if active or 1 if
not
if strBlnOpts.blnActive == 1
    beq = 0;
    iniWt = zeros(1,nAssets);
    minVarWts = zeros(1,nAssets);
else
    beq = 1;
    iniWt = ones(1,nAssets).*(1/nAssets);

    minVarWts = fmincon(@(pWts) pWts*VCV*pWts', iniWt,
strConstraints.A, strConstraints.b, Aeq, beq,...
    strConstraints.lb, strConstraints.ub, [], fminconOptions);
end

minVar = minVarWts*VCV*minVarWts';

maxRetWts = fmincon(@(tempWts) -tempWts*vecExpRets', iniWt,
strConstraints.A, strConstraints.b, Aeq, beq,...
    strConstraints.lb, strConstraints.ub, [], fminconOptions);

maxVar = maxRetWts*VCV*maxRetWts';

if strBlnOpts.blnInclMaxVol == 1
    targetAnnVol = [targetAnnVol, sqrt(maxVar)*sqrt(Ann)];
end

```

```

VarIncr = (maxVar - minVar)/(nFrontierPts-1);

iniWt = minVarWts;

for iPt = 1:nFrontierPts

    if isempty(targetAnnVol)
        targetVar = minVar + (iPt-1)*VarIncr;
    else
        targetVar = (targetAnnVol(1, iPt)^2)/Ann;
    end

    Chi = 0.5;

    [tempWts] = fmincon(@(tempWts)
        ((Chi*norm(tempWts)*norm(vecExpRets))-(dot(tempWts,vecExpRets))),
        iniWt, strConstraints.A, strConstraints.b, Aeq, beq,...
        strConstraints.lb, strConstraints.ub, @(tempWts)
        nonlconVar(tempWts, vecExpRets, targetVar, Ann, VCV),
        fminconOptions);

    pWts(iPt,:) = tempWts;

    pAngle(iPt,1) =
        (acosd(dot(tempWts,vecExpRets)/(norm(tempWts)*norm(vecExpRets))));

    pVol(iPt,1) = sqrt((tempWts)*VCV*(tempWts)')*sqrt(Ann);

    pExpRet(iPt,1) = dot(tempWts,vecExpRets);

    iniWt = tempWts;
    if strBlnOpts.blnDisplay == 1
        progressbar(iPt/nFrontierPts);
    end
end

pStats = [pVol, pExpRet];
checkVar = minVar:VarIncr:maxVar;
checkVar = sqrt(checkVar')*sqrt(Ann);
strOut.pAngle = pAngle;
strOut.pStats = pStats;
strOut.pWts = pWts;
strOut.checkVar = checkVar;
strOut.pVol = pVol;
strOut.pExpRet = pExpRet;

if strBlnOpts.blnDisplay == 1
    toc;
end

end

function [c,ceq] = nonlconVar(pWts, eRet, targetVar, Ann, VCV)

ceq=((pWts)*VCV*(pWts)')-targetVar;
c = [];
end

```