

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.

Portfolio Construction using Index Regression Models

by

Dirk Steyn (STYDIR003)

Dissertation presented in partial fulfillment of the requirements of the degree
Master in Financial Mathematics

at the

University of Cape Town



Acknowledgments

Firstly, I would like to thank my wife Rachelle for her patience and endless support in completing this dissertation. Secondly, I would like to thank Professor Casper Tröskic for the data, his enthusiasm and time spent in long consultations in his office. Thirdly, I would like to thank Allan Clark, for his advice and input in finalizing the dissertation. Next I would like to thank Professor Linda Haines for always making time to answer my many questions and supplying me with many good ideas and feedback.

Finally I would like to acknowledge the National Research Foundation for providing me with some financial support during the completion of this dissertation.

Abstract

In this dissertation we review the Sharpe Index Model and an innovation on this model introduced by Hossain, Troskie and Guo (2005b). These models are extended to the multi index framework. We then empirically investigate the impact of the models on portfolio creation over an extensive data set. Next we extend these models by modelling the regression residuals as ARMA and GARCH(1,1) processes and investigate the effect on the resulting portfolios. We then introduce the topic of bounded influence regression and apply it to financial data by down weighting extreme returns prior to regression. A new weighting function is introduced in this dissertation and the effects on the efficient frontiers and resulting market portfolios for the chosen set of shares are investigated. Then we investigate the creation of Principal Component Indices and the impact of using them in index regression models. Finally we use these and other models as investment strategies to build portfolios and simulate and compare them over the last 15 years using real market data. We draw conclusions regarding the significance of the index models compared to other investment strategies.

Contents

Contents	2
1 Introduction	7
2 The Sharpe Single Index Model	9
2.1 Classical Formulations	9
2.2 The Capital Market Line and Market Portfolio	12
2.3 Bivariate Normal Distribution and the Market Index	14
2.4 The Sharpe Index Model	17
2.5 The Improved Sharpe Index Model	21
2.6 Empirical Study: Sharpe Index and Improved Sharpe Index Models	23
2.6.1 The Data	23
2.6.2 Study Objectives	27
2.6.3 Methodology	27
2.6.4 Primary Findings	28
2.6.5 Conclusion	29
3 The Sharpe Multi Index Model	30
3.1 Improved Sharpe Multiple Index Model	34
3.2 Empirical Study: Multiple Index Model	36

3.2.1	The Data	36
3.2.2	Study Objectives	39
3.2.3	Methodology	40
3.2.4	Primary Findings	40
3.2.5	Conclusion	41
3.3	Conclusion	42
4	ARMA and GARCH Regression Models	44
4.1	Linear Regression Models with Autocorrelated Errors	44
4.2	GARCH Models	49
4.3	Empirical Study: Comparing a 9 share and a 20 share portfolio.	53
4.3.1	The Data	53
4.3.2	Study Objectives	54
4.3.3	Methodology	54
4.3.4	Primary Findings	54
4.3.5	Conclusion	57
4.4	Empirical Study: Regression models with ARMA and GARCH	57
4.4.1	The Data	58
4.4.2	Study Objectives	58
4.4.3	Methodology	58
4.4.4	Primary Findings	63
4.5	Conclusion	73
5	Bounded Influence Regression	76
5.1	Introduction	76

5.2	The Flat Weights Function	76
5.3	Mallows Weight Function	78
5.4	The ArcTan weighting function	83
5.5	Empirical Study: The ArcTan weighting functions	86
5.5.1	The Data	86
5.5.2	Study Objectives	87
5.5.3	Methodology	87
5.5.4	Primary Findings	87
5.6	Empirical Study: Comparing the weighting functions	88
5.6.1	The Data	88
5.6.2	Study Objectives	89
5.6.3	Methodology	89
5.6.4	Primary Findings and Results	90
5.7	Conclusion	96
6	Principal Components Regression Models	97
6.1	Theory of PCA	97
6.2	Sample Principal Components	99
6.3	Empirical Study: Regression with Principal Component Indices	102
6.3.1	The Data	102
6.3.2	Study Objectives	102
6.3.3	Methodology	103
6.3.4	Primary Findings	104
6.4	Conclusion	110

7 Portfolio Runs	112
7.1 Introduction.....	112
7.2 Assumption for portfolio construction	113
7.3 The Data Set	114
7.4 Strategies for Portfolio Construction	114
7.4.1 Sharpe Index Model Strategy (SIM).....	114
7.4.2 Improved Sharpe Index Model Strategy (ISIM)	114
7.4.3 Sharpe Multi Index Model Strategy (SMIM)	115
7.4.4 Index Model with ARMA and GARCH(1,1) (ARGA SIM).....	115
7.4.5 ISIM with ArcTan weighted returns (ARCTAN ISIM).....	115
7.4.6 Equal weighted share portfolio	115
7.4.7 JSE Overall Index	117
7.4.8 "7 Best" Portfolio	117
7.4.9 "7 Worst" Portfolio	118
7.4.10 PCA 1	118
7.5 The Performance of the Strategies.....	119
7.5.1 Market Overview.....	119
7.5.2 Performance of the Index Model Portfolios	121
7.5.3 Performance of the Alternative Portfolios	123
7.6 Conclusion	125
8 Summary and Conclusions	127
References	130

A EViews Code	132
B MATLAB Code	133
C VBA Code	134
D Singular Value Decomposition	135

University of Cape Town

Chapter 1

Introduction

The aims of the dissertation are as follows:

- To compare and investigate the performance of the Sharpe Index Model and extensions thereof.
- To examine bounded influence for Index Models, to critically appraise existing weight functions and to develop one that performs well with financial data.
- To investigate the idea of replacing the indices in the Sharpe Index Models by selected principal components derived from those indices and from shares comprising the portfolios of interest.
- To present an extensive analysis of the performance of the Index Models considered in the mini-dissertation over a long period of time.

In this dissertation we investigate the use of index regression models in portfolio construction. In Chapter 2 we present the theory of the Sharpe Index Model and introduce the Improved Sharpe Index Models in the single index framework. The relative positioning of the resulting mean-variance efficient frontiers are investigated.

In Chapter 3 we extend these models to incorporate additional indices and compare the risk-return profile of these models with each other and with the single index case. These

multi index models highlight the fundamental shortcomings of the Sharpe Multi Index Model and the effect of using the improved version of this model.

In Chapter 4 we extend the standard Sharpe Multi Index model by introducing ARMA and GARCH (1, 1) processes to models the residuals of the regression. Combinations of the index and residual models are investigated and resulting efficient frontiers and market portfolios calculated and compared. The real effect of using ARMA and GARCH(1,1) models on the resulting portfolio is investigated for the first time and put into perspective.

In Chapter 5 we constrain the historic returns by, in effect, filtering them via a weighting function prior to regression. This method is termed Bounded Influence Regression. The Mallows (1973) weighting function is investigated and a new weighting function is introduced in this dissertation. The different weighting functions are investigated and the resulting models are compared in an empirical study.

In Chapter 6 we build Principal Components Analysis based indices to be used as independent variables in regression models. We then compare different combinations of indices and their influence on the efficient frontier and resulting market portfolio.

In Chapter 7 we construct portfolios using various strategies based on the models introduced in this dissertation and then simulate these portfolios over the last 15 years using real market data. The performance and risk measures associated with the resulting portfolios are evaluated over the different periods.

In Chapter 8 we summarise the findings of this study and draw conclusions on the relevance of these Sharpe type regression models based on the empirical studies of JSE shares over the last 15 years.

Chapter 2

The Sharpe Single Index Model

In this chapter we introduce the basic notation and concepts used in the remainder of the mini-dissertation. We rely heavily on the notes of Troskie (2004) and also use his notation.

2.1 Classical Formulations

Let p be the number of shares in our portfolio. Then the vector of return values of the shares is given by

$$\mathbf{R} = \begin{pmatrix} R_1 \\ \vdots \\ R_p \end{pmatrix}. \quad (2.1)$$

We assume these returns to be log returns, with expectation

$$E(\mathbf{R}) = \boldsymbol{\mu}.$$

The covariance matrix of the share returns contains the historic covariance structure of the portfolio which is a measure of the non-systematic risk of the portfolio and is defined as follows:

$$\begin{aligned} \boldsymbol{\Sigma} &= E[(\mathbf{R} - \boldsymbol{\mu})(\mathbf{R} - \boldsymbol{\mu})'] \\ &= \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix}, \end{aligned}$$

where $\sigma_{ii} = \sigma_i^2$ is the variance of the i^{th} share and σ_{ij} the covariance between the i^{th} and the j^{th} share.

A portfolio of shares is a proportional investment, say w_i in each share, so that

$$\sum_{i=1}^p w_i = 1. \quad (2.2)$$

Let

$$\mathbf{W} = \begin{pmatrix} w_1 \\ \vdots \\ w_p \end{pmatrix} \quad (2.3)$$

be the vector of weights invested in each share; then the return of the portfolio is

$$P = \mathbf{W}'\mathbf{R} = \sum_{i=1}^p w_i R_i.$$

The expected return of the portfolio is given by:

$$E(P) = \mathbf{W}'E(\mathbf{R}) = \mu_p$$

and the historic variance of the portfolio is:

$$\text{var}(P) = \mathbf{W}'\Sigma\mathbf{W} = \sum_{i=1}^p \sum_{j=1}^p w_i w_j \sigma_{ij} = \sigma_p^2.$$

If we assume in addition that $\mathbf{R} \sim N(\boldsymbol{\mu}, \Sigma)$ then it follows from normal theory that $P \sim N(\mu_p, \sigma_p^2)$.

We now assume, as we do throughout this mini-dissertation, that the members of the portfolio are fixed. The only way to optimize the portfolio is by adjusting the weights w_i invested in each share. Obviously, we can set the weights of some shares to zero thereby excluding them from the portfolio. Changing the weights will then change the value of the portfolio's expected return μ_p and variance σ_p^2 . Ideally we would want to maximize the expected return $E(P) = \mu_p$ of the portfolio. Unfortunately this is not possible, in general, without increasing the variance of the portfolio as well. The variance σ_p^2 (and standard

deviation σ_p) of the portfolio is a measure of the risk of the portfolio and in general should be minimized.

Thus we want to choose the weights w_i such that the expected return $E(P) = \mu_p$ is a maximum but also at the same time that the risk or variance σ_p^2 is a minimum. This is the classical portfolio formulation due to Markowitz (1952). Markowitz, also introduced the concept of an efficient frontier.

Definition1 A portfolio is called efficient if:

- For a given amount of risk, the expected return is maximized, or
- for a given amount of return, the expected risk is minimized.

The solution of this portfolio optimization problem will lie on the efficient frontier line. Specifically the efficient frontier is a line plot in the two dimensional mean-variance space. Finding the efficient frontier is a non-linear quadratic programming (QP) problem and can be solved using well-known numerical techniques. To solve the problem one needs either to fix the return μ_p and then minimize the variance σ_p^2 , or fix the variance and then maximize the return. They will both lead to the same answer. If we choose to fix the return,

then we need to solve the following constrained QP problem.

$$\begin{aligned} \text{Min } \sigma_p^2 &= \mathbf{W}'\Sigma\mathbf{W} & (2.4) \\ &= \sum_{i=1, j=1}^p w_i w_j \sigma_{ij} \end{aligned}$$

$$\begin{aligned} \text{subject to } \mu_p &= \mathbf{W}'\boldsymbol{\mu} = \mu_{fixed} \\ \sum_{i=1}^p w_i &= 1, \text{ and } 0 \leq w_i \leq 1, & (2.5) \end{aligned}$$

where μ_{fixed} is a fixed expected return on the μ_p . By varying μ_{fixed} a sufficient number of times we can generate the efficient frontier.

An implementation in MATLAB based on the routine in Ruppert (2004) can be found in the Appendix B

2.2 The Capital Market Line and Market Portfolio

Each point on the efficient frontier represents a different portfolio. If the use of a risk free asset is introduced, an optimal portfolio on the efficient frontier can be found. Suppose it is possible to borrow or lend any amount of money at the fixed interest rate R_f . The Capital Market Line (CML) follows by drawing a straight line out from the risk free rate R_f with zero variance into the (μ_p, σ_p^2) space. This line is then swung (upwards or downwards) until it is tangent to the efficient frontier. This then yields the point marked with a star and called the Market Portfolio in Figure 2.1. For a full discussion on the market portfolio and efficient frontier, see Elton, Gruber, Brown, Goetzmann (2003, p. 86). Clearly if R_f or the efficient frontier moves then the market portfolio will move accordingly.

Points between R_f and the market portfolio represents lending portfolios (you are lending money to the bank). The portfolios are constructed by varying proportions of the risk-free asset R_f and the market portfolio.

Points on the extension of the CML line, that lie above the market portfolio represents borrowing portfolios. Their creation requires borrowing money at rate R_f to increase total investable capital. The total investable capital is then invested in the market portfolio which means that both return and risk (variance) is increased along the CML.

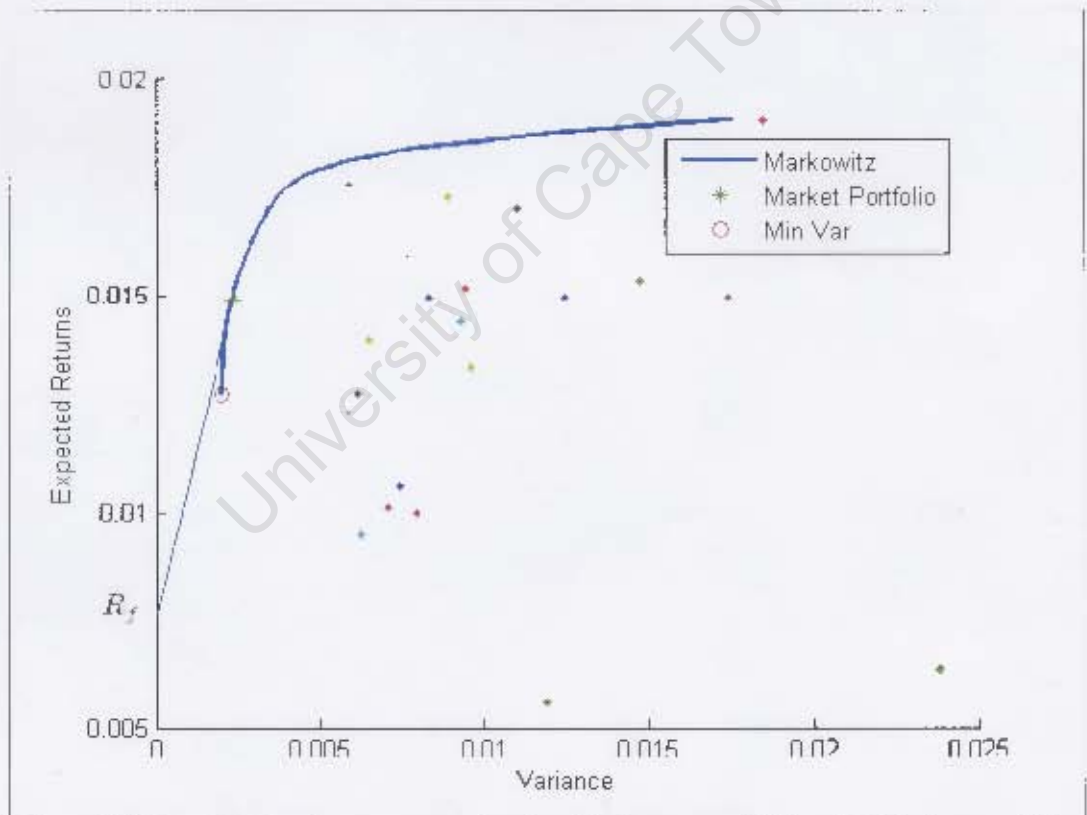


Figure 2.1 The CML is shown connecting the risk free rate of 9% to the Market portfolio. Each dot represents one of 20 stocks in the portfolio.

In figure 2.1 each point in the plot shows variance and expected return for one of the 20 shares. The most efficient portfolio, of these shares, are found on the efficient frontier.

Each point on the efficient frontier represent a vector of portfolio weights. In Figure 2.2 we consider the weights of the Market Portfolio for our 20 share portfolio in Figure 2.1. More details on these shares are provided later. As can be seen from the figure, not all shares form part of the market portfolio and one share, Remgro Ltd, is dominant. This is typical of an unconstrained portfolio optimization. To remedy this situation, constraints can be used on each share, to keep the weights between say 5% and 20%.

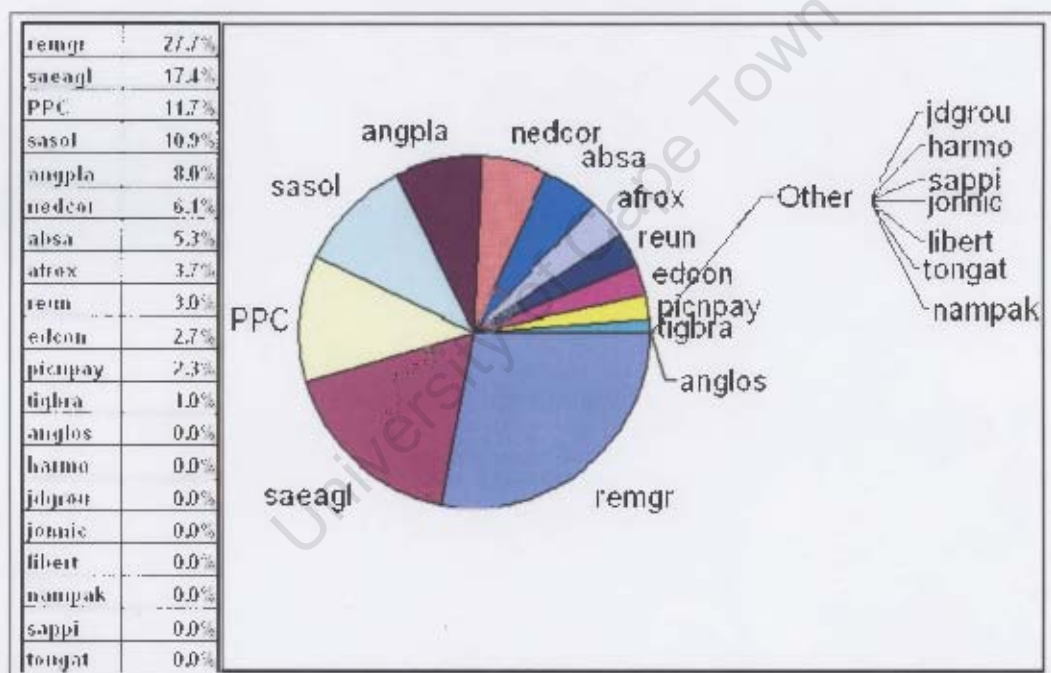


Figure 2.2 The market portfolio for the 20 shares. Note that not all the shares are part of the market portfolio and that the best performing share remgro is the major component of the portfolio.

2.3 Bivariate Normal Distribution and the Market Index

Let the log return of a particular share be

$$R_t = \log P_t - \log P_{t-1}, \quad t = 1, \dots, N,$$

where the time t is large enough for R_t to be estimated by a normal $N(\mu_r, \sigma_r^2)$ distribution.

Similarly, let

$$I_t = \log I_t - \log I_{t-1}, \quad t = 1, \dots, N$$

be the return of the market proxy (usually the JSE Overall Index or any of the other sector indices) such that it also follows a normal distribution $N(\mu_I, \sigma_I^2)$. We then have that

$$\begin{pmatrix} R_t \\ I_t \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_r \\ \mu_I \end{pmatrix}, \begin{pmatrix} \sigma_r^2 & \sigma_{rI} \\ \sigma_{Ir} & \sigma_I^2 \end{pmatrix} \right] \quad (2.6)$$

a bivariate normal distribution where $\sigma_{rI} = \sigma_{Ir}$ is the covariance between the return of share R and the market proxy I . From the properties of the bivariate normal distribution (Rice, 1995) the expectation on of the share, conditional on the index, is:

$$E(R_t|I_t) = \rho(I_t - \mu_I)\sigma_r|\sigma_I$$

or

$$E(R_t|I_t) = \alpha + \beta I_t \quad (2.7)$$

$$\text{where } \beta = \rho \frac{\sigma_r}{\sigma_I},$$

$$\alpha = \mu_I \rho \frac{\sigma_r}{\sigma_I}$$

and the variance of the share conditional on the index is

$$\text{var}(R_t|I_t) = \sigma_{r,I}^2 = \sigma_r^2(1 - \rho^2),$$

where ρ is the correlation between the share return and the market proxy. The model can thus be written as

$$R_t = \alpha + \beta I_t + e_t, \quad (2.8)$$

where we normally assume that

$$e_t \sim N(0, \sigma_e^2).$$

It is usually assumed that the residuals e_t are independently distributed over time, that is

$$E(e_t e_s) = 0, \text{ for } t \neq s.$$

The beta coefficient which is given in Equation (2.7) is an important statistical parameter for a share in finance. It measures 'elasticity' of the share to the market proxy used for the regression. A value of one means that the share follows the market perfectly (assuming a good model fit) and the share has the same systematic risk as the market. A value higher than one indicates that, historically, the share overreacts to the market ups and downs. We say this share has a high elasticity or higher systematic risk than the market. A value smaller than 1 but positive would point to a low elasticity share with lower than market systematic risk. A negative beta would only be possible if the share is negatively correlated with the market proxy. The assumption is made that the beta coefficient is fairly stable over time, especially as the length of the period under consideration increases. This implies that historical data can be used to estimate the parameters and the estimates can be used for future portfolio creation. Even if the beta coefficients do change over time, the ranking based on betas of the securities does not change. This is most relevant if the index model is used for portfolio optimization. The beta coefficients do give a fairly good measure of the systematic risk inherent in a security. Risk, in terms of variance of a share, can be divided

in two terms. From Equation (2.8):

$$\begin{aligned} \text{Var}(R_t) &= \text{Var}(\alpha + \beta I_t + e_t) \\ &= \beta^2 \sigma_I^2 + \text{Var}(e_t) \end{aligned}$$

$$\text{So that } \sigma_R^2 = \beta^2 \sigma_I^2 + \sigma_e^2.$$

The systematic risk of a share (or portfolio) explained by the market is given by $\beta^2 \sigma_I^2$ and the non-systematic, or unique risk of the share, is given by σ_e^2 and represents the variance not explained by the market.

Even if the bivariate normal assumption, specified in Equation (2.6), is not accepted the linearity assumption specified in Equation (2.7) appears to be very well satisfied (Fama, Fisher, Jensen and Roll, 1969).

Furthermore, if the assumption that the error terms e_t are normally distributed is lifted, but we still assume that $E(e_t) = 0$, $E(e_t)^2 = \sigma_e^2$ and $E(e_t e_s) = 0$, $t \neq s$, then the Gauss-Markoff Theorem guarantees that the Ordinary Least Square Estimate of β is the best linear unbiased estimate in a minimum variance sense.

2.4 The Sharpe Index Model

The Sharpe Index Model (SIM) is based on a one-factor regression model of shares returns. It is a response to the observation that shares follow the movements of the market. Therefore the main assumption of the SIM is that shares move together only because of a common co-movement with the market. Or, that the only correlation that two stocks share is ex-

plained via the share correlation with the market index. It specifically assumes that e_{it} is independence of e_{jt} for all values of i and j . This implies that $E(e_{it}e_{jt}) = E(e_{it})E(e_{jt}) = 0$.

For a portfolio made up of several shares, let for the i^{th} share be modelled as:

$$R_{it} = \alpha_i + \beta_i I_t + e_{it}, \quad i = 1, \dots, p; \quad t = 1, \dots, N, \quad (2.9)$$

where R_{it} is the return of the i^{th} share at time t and I_t is the return of the index at time t . The unknown parameters α_i and β_i are estimated by least square regression. We assume that there are p shares in the portfolio and that there are N samples in time of the return of both the shares and index are available at N consecutive times. The standard assumptions for the SIM are as follows

$$E(e_{it}^2) = \sigma_{ei}^2 \quad (2.10)$$

$$E(e_{it}e_{is}) = 0, \quad t \neq s = 1, \dots, N, \quad (2.11)$$

$$E(e_{it}I_t) = 0, \quad t = 1, \dots, N, \quad (2.12)$$

$$E(e_{it}e_{jt}) = 0, \quad t = 1, \dots, N, \quad i \neq j. \quad (2.13)$$

Equation (2.10) assumes each share has its own variance for the error term. Equation (2.11) assumes that the error terms of each share are independent over time and therefore that there is no autocorrelation in the residual series. Equation (2.12) assumes that the errors of each share and the explanatory variable I are uncorrelated which is the usual assumption in regression. Equation (2.13) assumes that the error terms of the shares are uncorrelated with each other, so the shares are only related through their mutual relationship with the index I . This assumption forces the covariance matrix of the errors to be a diagonal matrix with

only the σ_{ei}^2 terms non zero. We now let

$$E(I) = \mu_I \text{ and } \text{var}(I) = \sigma_I^2$$

be the mean and variance of the index. Then in vector notation the SIM is given by

$$\mathbf{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}I_t + \mathbf{e}_t \quad t = 1, \dots, N,$$

where

$$\mathbf{R}_t = \begin{pmatrix} R_{1t} \\ \vdots \\ R_{pt} \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \text{ and } \mathbf{e}_t = \begin{pmatrix} e_{1t} \\ \vdots \\ e_{pt} \end{pmatrix},$$

so that

$$E(\mathbf{R}) = \boldsymbol{\alpha} + \boldsymbol{\beta}\mu_I,$$

and as a consequence of Equation (2.13) it follows that the covariance matrix of the residuals is

$$\text{cov}(e) = \begin{pmatrix} \sigma_{e1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e2}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \sigma_{ep}^2 \end{pmatrix}.$$

Also the covariance between two returns is

$$\text{cov}(R_i R_j) = E[(R_i - E(R_i))(R_j - E(R_j))] = \beta_i \beta_j \sigma_I^2$$

and the variance of one return is

$$\text{var}(R_i) = \beta_i^2 \sigma_I^2 + \sigma_{ei}^2 = \sigma_i^2.$$

This implies that the covariance of the returns can be estimated from

$$\text{cov}(\mathbf{R}) = \sigma_I^2 \boldsymbol{\beta} \boldsymbol{\beta}' + \begin{pmatrix} \sigma_{e1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e2}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \sigma_{ep}^2 \end{pmatrix}. \quad (2.14)$$

We now define the matrix Ω_{diag} so as to emphasize the diagonal nature of the covariance matrix in the SIM:

$$\Omega_{diag} = \begin{pmatrix} \sigma_{e1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \sigma_{ep}^2 \end{pmatrix}. \quad (2.15)$$

Equation (2.14) can then be written as

$$Cov(\mathbf{R}) = \sigma_I^2 \beta \beta' + \Omega_{diag} = \Phi_{diag}.$$

For portfolio $P = \mathbf{W}'\mathbf{R}$, we have

$$E(P) = \mathbf{W}'(\alpha + \beta\mu_I) = \mu_p$$

and

$$\begin{aligned} var(P) &= \mathbf{W}'(\sigma_I^2 \beta \beta' + \Omega_{diag})\mathbf{W} \\ &= \mathbf{W}'\Phi_{diag}\mathbf{W} = \sigma_p^2. \end{aligned}$$

As a quadratic programming problem we then minimize the variance σ_p^2 subject to $\sum_{i=1}^p w_i = 1$ for a fixed portfolio return $\mu_p = \mu_{fixed}$. This is repeated to create the efficient frontier.

Each of the p regression equations can be solved individually to obtain estimates for the parameter β and the residuals as:

$$\begin{aligned} \hat{\beta}_i &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}_i, i = 1, \dots, p \\ \hat{e}_i &= \mathbf{Y}_i - \mathbf{X}\hat{\beta}_i, i = 1, \dots, p. \end{aligned}$$

A matrix of residuals can be created and is define $\hat{\mathbf{E}}$ as:

$$\hat{\mathbf{E}} = \begin{pmatrix} \hat{e}_{11} & \dots & \hat{e}_{1p} \\ \vdots & \vdots & \vdots \\ \hat{e}_{N1} & \dots & \hat{e}_{Np} \end{pmatrix}$$

the $(N \times p)$ residual matrix. The $(p \times p)$ moment matrix for the residuals is then $\hat{\mathbf{E}}'\hat{\mathbf{E}}$ and the sample covariance matrix is

$$\hat{\mathbf{\Omega}} = \frac{1}{N-2} \hat{\mathbf{E}}'\hat{\mathbf{E}}. \quad (2.16)$$

Note that we lose two degrees of freedom for estimating the parameters β and α and thus the denominator in Equation (2.16) is $N-2$. For the SIM we are only interested in the sample variances and therefore only $\hat{\mathbf{\Omega}}_{diag}$ is used. The estimation of the covariance matrix,

$\hat{\mathbf{\Phi}}_{SIM}$ is given by

$$\hat{\mathbf{\Phi}}_{SIM} = \hat{\sigma}_I^2 \hat{\beta} \hat{\beta}' + \hat{\mathbf{\Omega}}_{diag}. \quad (2.17)$$

It is this estimate of the covariance that is used as input for the quadratic programming algorithm to create the efficient frontier for the underlying shares based on the Sharpe Index Model.

2.5 The Improved Sharpe Index Model

The Improved Sharpe Index Model (ISIM) was introduced by Hossain, Troskie and Guo (2005b). The model is an attempt to improve on the SIM by including the covariance structure of the residuals. These were excluded by Sharpe to simplify the process of portfolio optimization with index models. The improved model has the advantage of including all the information used in the Markowitz model and is still easily extended to include ARMA and GARCH components in the model. ISIM is formulated as

$$R_{it} = \alpha_i + \beta_i I_t + e_{it}, \quad i = 1, \dots, p; \quad t = 1, \dots, N,$$

where R_{it} is the return of the i^{th} share at time t and I_t is the return of the index at time t . The parameter α_i and β_i are estimated via ordinary least squares regression with the following assumptions:

$$E(e_{it}^2) = \sigma_{ei}^2 \quad (2.18)$$

$$E(e_{it}e_{is}) = 0, \quad t \neq s = 1, \dots, N, \quad (2.19)$$

$$E(e_{it}I_t) = 0, \quad t = 1, \dots, N, \quad (2.20)$$

$$E(e_{it}e_{jt}) = \sigma_{ij}, \quad i \neq j, \quad t = 1, \dots, N. \quad (2.21)$$

$$E(e_{it}e_{jt}) = \sigma_{ei}^2, \quad i = j \quad (2.22)$$

The ISIM lifts the assumption that the correlation between the residuals term e_{it} , $i = 1, \dots, p$, $t = 1, \dots, N$, of the different shares is zero as specified in Equation (2.21). And the estimates of the covariances of the residuals are included in the estimated covariance matrix used to produce the efficient frontier for the ISIM model. Empirically, it was found that the correlations are not zero and correlations of up to 0.25 were observed in a sample of the 20 JSE listed shares used later in this mini-dissertation.

Thus Equation (2.15) becomes

$$E(\mathbf{ee}') = \mathbf{\Omega} = \begin{pmatrix} \sigma_{e1}^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{e2}^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \dots & \dots & \sigma_{ep}^2 \end{pmatrix}$$

with

$$\text{cov}(\mathbf{R}) = \sigma_I^2 \beta \beta' + \mathbf{\Omega} = \mathbf{\Phi}.$$

For portfolio $P = \mathbf{W}'\mathbf{R}$ we have

$$E(P) = \mathbf{W}'(\boldsymbol{\alpha} + \beta\mu_I) = \mu_p.$$

and

$$var(P) = \mathbf{W}'(\sigma_I^2\beta\beta' + \Omega)\mathbf{W} = \mathbf{W}'\Phi\mathbf{W} = \sigma_p^2.$$

We then use the Quadratic Programming optimization with the Φ matrix subject to $\sum_{i=1}^p w_i = 1$ and any other equality or inequality constraints and bounds.

The quantities to be estimated are

$$\mu_I, \sigma_I^2, \alpha, \beta \text{ and } \Omega.$$

The regression residuals are again used to create the matrix $\hat{\mathbf{E}}$

$$\hat{\mathbf{E}} = \begin{pmatrix} \hat{e}_{11} & \hat{e}_{12} & \dots & \hat{e}_{1N} \\ \hat{e}_{21} & \hat{e}_{22} & \dots & \hat{e}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{e}_{p1} & \dots & \dots & \hat{e}_{pN} \end{pmatrix}$$

where

$$\hat{\Omega} = \frac{1}{N-2} \hat{\mathbf{E}}\hat{\mathbf{E}}'$$

and the estimated covariance is

$$\hat{\Phi} = \hat{\sigma}_I^2\hat{\beta}\hat{\beta}' + \hat{\Omega}.$$

2.6 Empirical Study: Sharpe Index and Improved Sharpe Index Models

2.6.1 The Data

A diversified set of 20 shares from the JSE in South Africa was chosen for the empirical study. A list of the shares with their abbreviated name and description follows:

1. **Anglo American plc (anglos):** A global leader in mining and natural resources and owns a well diversified range of high quality assets.
2. **JD Group Ltd (jdgrou):** The Group carries on the business of furniture and appliance retailing as well as the provision of financial services.
3. **Pick 'n Pay Stores Ltd (picnpay):** The company is an investment holding company whose subsidiaries are active in the retail area. The Pick 'n Pay chain is one of the largest supermarket chains in South Africa.
4. **Remgro Ltd (remgr):** The company is an investment holding company. Cash income is derived mainly from dividends and interest. The group's interests consist mainly of investments in tobacco products, banking and financial services, printing and packaging, building and motor components, medical services, mining, petroleum products, food, wine and spirits and various other trade mark products.
5. **South African Eagle Insurance Company Ltd (saeagl):** A short-term insurance Company in South Africa.
6. **Sappi Ltd (sappi):** Sappi is a pulp and fine paper products group with manufacturing facilities on four continents and an international sales network that markets the group's products to over 100 countries.
7. **Sasol Ltd (sasol):** The Sasol Group of companies comprises diversified fuel, chemical and related manufacturing and marketing operations. These core operations are complemented by coal-mining operations and oil and gas exploration and production

8. **Tiger Brands Ltd (tigbra):** A management company with operations primarily in a variety of food and health related businesses. These include, inter alia, milling, baking, confectionery, rice, pasta, fruit and vegetables, wholesaling, edible oils, and health care products. Tiger Brands has interests in fishing and overseas investments in the barley, malt and food industries.
9. **Tongaat-Hulett Group Ltd (tongaat):** Tongaat is a diversified industrial group with interests in the sugar and aluminum industries.
10. **Absa Group Ltd (absa):** ABSA is one of the four big banks in South Africa that offers a comprehensive range of banking services, bank assurance and wealth management products and services.
11. **African Oxygen Ltd (afrox):** The company is an integrated, full-spectrum gas and welding products business and is the largest such business on the African continent.
12. **Anglo Platinum Ltd (angpla):** The world's largest platinum producer, it mines platinum and PGM metals in the Bushveld Complex with 5 underground mines and one open pit mine. Gold, copper, nickel and cobalt are recovered as by-products. The company has its own precious metals and base metals refinery.
13. **Harmony Gold Mining Company Ltd (harmo):** Harmony is an unhedged gold mining company with an annualised production in excess of 2 million ounces of gold. The company is engaged in gold mining in South Africa and Australasia in addition to its focused portfolio of exploration projects. Harmony has been successful in its

strategy of growth through acquisitions and has emerged as the fifth largest gold producer in the world.

14. **Johnnic Communications Ltd (johnnic):** This company has an investment holding company in the media and entertainment industries.
15. **Liberty Group Ltd (libert):** This is a financial services group focused on developing, marketing and managing a comprehensive range of investment and risk products designed to cater for all personal and corporate investment, life assurance, disability, health assurance and retirement needs.
16. **Nampak Ltd (nampac):** The group is the largest and most diversified packaging manufacturer in Africa with operations in the United Kingdom and Europe. It produces packaging products from metal, glass, paper and plastics, and is a major manufacturer and marketer of tissue products
17. **Nedbank Group Ltd (nedcor):** Nedbank is one of South Africa's large banking groups. The group provides banking, mortgage loan finance and general financial services.
18. **Reunert Ltd (reun):** Reunert manages a number of businesses focused on electronics and low-voltage electrical engineering. Reunert is included in the ALSI 40 index
19. **Pretoria Portland Cement Company Ltd (PPC):** PPC manufactures and sells cement, lime and related products.

20. **Edgars Consolidated Stores Ltd (edcon):** The Edgars group business is the retailing of clothing, footwear, accessories, home textiles and others through its stores in southern Africa. In its factories the group manufactures a broad range of family clothing which is sold through its own outlets and to the outside market

JSE Overall Index: This is the major equity index of South Africa listed shares.

An extensive time series was used of monthly log return data spanning 224 observations from 1988 to 2007. A program written in EVIEWS 3 was used to calculate the returns, perform the regression, and calculate the covariance and return estimate.

2.6.2 Study Objectives

The objective of this empirical study is to investigate the effect of using the Sharpe Index Model (SIM) and Improved Sharpe Index Model (ISIM) on the mean-variance response as represented by the efficient frontier and to use a more substantial set of shares to verify the results of Hossain et al. (2005a) for the single index model independently.

2.6.3 Methodology

A program written in EVIEWS 3 was used to calculate the returns, perform the regression, and calculate the covariance and return estimate for each of the models. The equation for the estimate of the covariance for the SIM was used:

$$\hat{\Phi}_{SIM} = \hat{\sigma}_I^2 \hat{\beta} \hat{\beta}' + \hat{\Omega}_{diag}$$

and for the ISIM the full estimated covariance matrix is used:

$$cov(\mathbf{R}) = \Phi - \sigma_i^2 \beta_i \beta_i' + \Omega.$$

These outputs were then exported to MATLAB 7 and the quadratic programming problem was solved producing the efficient frontiers and associate portfolios as discussed in section on quadratic programming. Both sample programs can be found in the Appendices A and B respectively.

2.6.4 Primary Findings

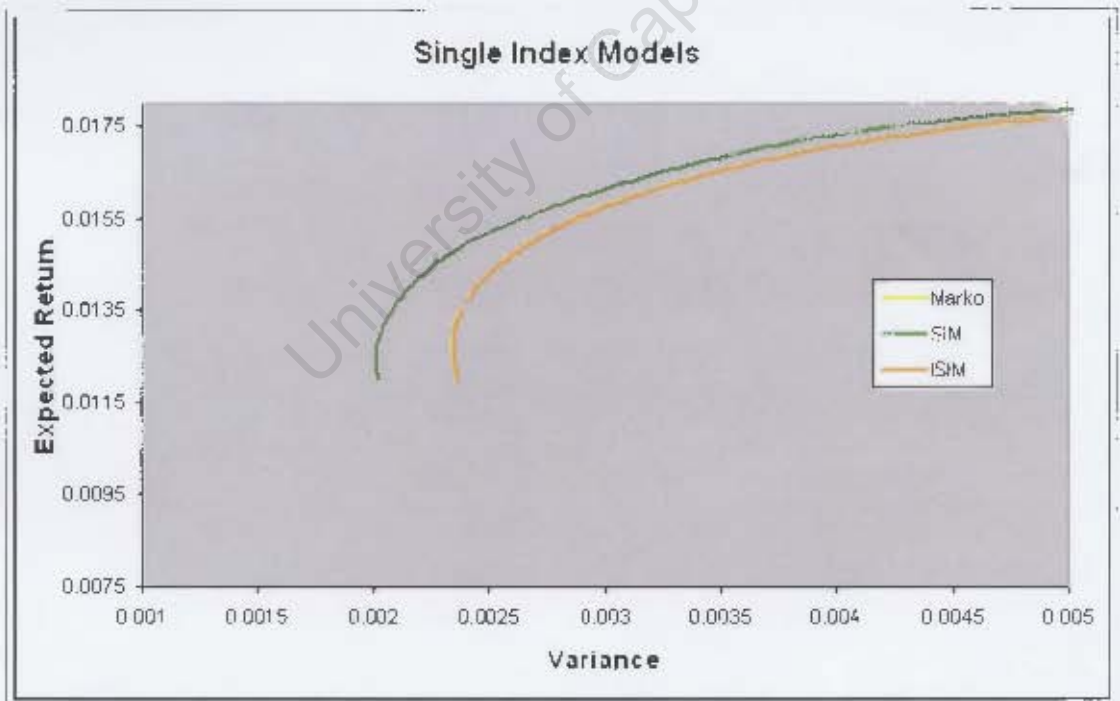


Figure 2.3 The efficient frontier for Markowitz, SIM and ISIM. From left to right SIM, ISIM and Markowitz. The ISIM and Markowitz are basically on top of each other.

Three efficient frontiers are shown in the Figure 2.3. From left to right they correspond to the SIM, ISIM and Markowitz models. The ISIM and Markowitz frontiers are

basically on top of each other. These results show that SIM under-estimates the risk of the portfolio compared to the original Markowitz model. This is due to the fact that less of the covariance structure information of the constituents is used in the construction of the SIM. The value of the ISIM is that the removed covariance structure, of the SIM, is replaced by the estimated covariance structure of the constituents' residuals. This will explain why the ISIM frontier lies 'very close to' the Markowitz frontier.

2.6.5 Conclusion

The SIM covariance produces an efficient frontier to the left of that of the historic covariance of Markowitz. However, the ISIM produces an efficient frontier 'very close to' Markowitz efficient frontier. We conclude that the residual covariances that were not included in the construction of the Sharpe Index Model are predominately positively correlated and therefore the model underestimate the risk in the portfolio compared to the ISIM and Markowitz models for this selection of shares.

The ISIM incorporates the residual covariance, resulting in an efficient frontier similar to the Markowitz. The ISIM therefore uses all the available information to create its covariance estimate. The ISIM has the advantage over the traditional Markowitz models that it distinguish between systematic and non-systematic risks. The model is also extendable using econometrics techniques including ARMA and GARCH.

Chapter 3

The Sharpe Multi Index Model

Sharpe Multi Index Models (SMIM) are an attempt to capture some of the non-market influences that cause securities to move together. The idea is to use indices that are not too correlated with the market index to capture additional information relevant to the shares that was not contained in the market index. There are other uses of multi-index models besides predicting the correlation structure of a portfolio. Multi-index models can be used to study the impact of events, or as a method for tailoring the return distribution of a portfolio to the specific needs of an investor. It can also be used to analyze the cause of good or bad performance on a portfolio according to Elton, Gruber, Brown and Goetzmann (2003).

The Sharp Multi Index model can be written as

$$R_{it} = \alpha_i + \beta_{i1}I_1 + \beta_{i2}I_2 + \dots + \beta_{iM}I_M + e_{it},$$

$$i = 1, \dots, p, t = 1, \dots, N,$$

with the following assumptions

$$E(e_{it}^2) = \sigma_{ei}^2 \tag{3.23}$$

$$E(e_{it}e_{is}) = 0, t \neq s = 1, \dots, N, \tag{3.24}$$

$$E(e_{it}I_{jt}) = 0, j = 1, \dots, M, t = 1, \dots, N, \tag{3.25}$$

$$E(e_{it}e_{jt}) = 0, t = 1, \dots, N, \tag{3.26}$$

$$Cov(I_{jt}I_{kt}) = c_{jk}, j, k = 1, \dots, M. \tag{3.27}$$

These assumptions are identical to the Single Index Model where in Equation (3.25) we now also assume that the disturbance term e_{it} is also independent of the indices I_j , $j = 1, \dots, M$. This again is a normal assumption in regression. We further assume that the indices are dependent with covariances given by c_{jk} in Equation (3.27). Also note that this equation introduces a further $M(M + 1)/2$ covariances with covariance matrix

$$\mathbf{C} = [c_{jk}], \quad j, k = 1, \dots, M$$

between the M indices. Let

$$\begin{aligned} E_i &= E(R_i) = \alpha_i + \beta_{i1}\mu_1 + \dots + \beta_{iM}\mu_M, \\ i &= 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{\alpha} &= \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \dots & \beta_{1M} \\ \vdots & \vdots & \vdots \\ \beta_{p1} & \dots & \beta_{pM} \end{pmatrix}, \quad \mathbf{e}_t = \begin{pmatrix} e_{1t} \\ \vdots \\ e_{pt} \end{pmatrix} \\ R_t &= \begin{pmatrix} R_{1t} \\ \vdots \\ R_{pt} \end{pmatrix}, \quad \mathbf{I}_t = \begin{pmatrix} I_{1t} \\ \vdots \\ I_{Mt} \end{pmatrix}, \quad t = 1, \dots, N. \end{aligned}$$

Then

$$\mathbf{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{I}_t + \mathbf{e}_t$$

with

$$\mathbf{E}(\mathbf{R}_t) = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{\mu}_I$$

and

$$\boldsymbol{\mu}_I = \mathbf{E}(\mathbf{I}_t) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_M \end{pmatrix}.$$

The covariance matrix of the p returns \mathbf{R}_t is then

$$\begin{aligned}
 cov(\mathbf{R}_t) &= E[\mathbf{R}_t - E(\mathbf{R}_t)][\mathbf{R}_t - E(\mathbf{R}_t)]' \\
 &= E[\beta(\mathbf{I}_t - \boldsymbol{\mu}_1) + \mathbf{e}_t][\beta(\mathbf{I}_t - \boldsymbol{\mu}_1) + \mathbf{e}_t]' \\
 &= \beta E(\mathbf{I}_t - \boldsymbol{\mu}_1)(\mathbf{I}_t - \boldsymbol{\mu}_1)' \beta' + E(\mathbf{e}_t \mathbf{e}_t') \quad \text{since } E(\mathbf{I}_t \mathbf{e}_t') = 0 \\
 &= \beta \mathbf{C} \beta' + \boldsymbol{\Omega} \quad \text{since } E(\mathbf{e}_t \mathbf{e}_t') = \boldsymbol{\Omega} \text{ and } \mathbf{C} = cov(\mathbf{I}_t) \\
 &= \boldsymbol{\Phi}.
 \end{aligned} \tag{3.28}$$

For portfolio $P = \mathbf{W}'\mathbf{R} = \sum_{i=1}^p w_i R_i$ we have that

$$E(P) = \mathbf{W}'(\boldsymbol{\alpha} + \beta \boldsymbol{\mu}_1) = \boldsymbol{\mu}_p$$

and

$$\sigma_p^2 = var(P) = \mathbf{W}'\boldsymbol{\Phi}\mathbf{W}. \tag{3.29}$$

The objective function for our quadratic optimization problem then becomes Equation (3.29) with the constraints set to $\sum_{i=1}^p w_i = 1$ and any other constraints or bounds relevant to the portfolio creation

Estimation of $\boldsymbol{\Phi}$ and related parameters

Our estimates for the return of the portfolio would be

$$\hat{E}(P) = \mathbf{W}'\hat{\mathbf{R}}$$

where $\hat{\mathbf{R}}$ is the vector of estimates for R_i

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{R}_1 \\ \vdots \\ \hat{R}_p \end{pmatrix}$$

where $\mathbf{I}_t^{m.adj}$ is the mean adjusted matrix of the indices and $\hat{\Omega}$ is estimated by

$$\hat{\Omega} = \frac{1}{N - M - 1} \hat{\mathbf{E}} \hat{\mathbf{E}}'$$

and again, as in the case of the SIM, $\hat{\Omega}_{diag} = diag(\hat{\Omega})$.

We have lost $M + 1$ degrees of freedom via estimation in the regressions. $\hat{\mathbf{E}}$ is again the matrix of regression residuals :

$$\hat{\mathbf{E}} = \begin{pmatrix} \hat{e}_{11} & \hat{e}_{12} & \dots & \hat{e}_{1N} \\ \hat{e}_{21} & \hat{e}_{22} & \dots & \hat{e}_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{e}_{p1} & \dots & \dots & \hat{e}_{pN} \end{pmatrix}.$$

3.1 Improved Sharpe Multiple Index Model.

The Improved Sharpe Multiple Index Model (ISMIM) can again be written as

$$\begin{aligned} R_{it} &= \alpha_i + \beta_{i1} I_{1t} + \beta_{i2} I_{2t} + \dots + \beta_{iM} I_{Mt} + e_{it}, \\ i &= 1, \dots, p, \quad t = 1, \dots, N, \end{aligned}$$

with the following assumptions

$$E(e_{it}^2) = \sigma_{ei}^2 \quad (3.30)$$

$$E(e_{it} e_{jt}) = \sigma_{ij}, \quad i \neq j, \quad t = 1, \dots, N. \quad (3.31)$$

$$E(e_{it} e_{is}) = 0, \quad t \neq s = 1, \dots, N, \quad (3.32)$$

$$E(e_{it} I_{jt}) = 0, \quad j = 1, \dots, M, \quad t = 1, \dots, N, \quad (3.33)$$

$$Cov(I_{jt} I_{kt}) = c_{jk}, \quad j, k = 1, \dots, M \quad (3.34)$$

Similar to the Improved Sharpe Index Model, the covariances between the residuals are included in the model as is seen in Equation (3.31). The residuals are still independent of each of the indices, but the covariances between the index and residuals are now included in this model c_{jk} .

The model is then again

$$\mathbf{R}_t = \alpha + \beta \mathbf{I}_t + \mathbf{e}_t.$$

The covariance matrix of \mathbf{R}_t is the same as Equation (3.28),

$$\begin{aligned} \text{cov}(\mathbf{R}_t) &= E[\mathbf{R}_t - E(\mathbf{R}_t)][\mathbf{R}_t - E(\mathbf{R}_t)]' \\ &= \beta \mathbf{C} \beta' + \Omega = \Phi. \end{aligned}$$

For the estimate of σ_p^2 we have

$$\hat{\sigma}_p^2 = \mathbf{W}' \hat{\Phi} \mathbf{W}$$

with

$$\hat{\Phi} = \hat{\beta} \hat{\mathbf{C}} \hat{\beta}' + \hat{\Omega}. \quad (3.35)$$

The matrix $\hat{\mathbf{C}}$ is the estimated covariance matrix of the indices,

$$\hat{\mathbf{C}} = \frac{1}{N-1} (\mathbf{I}_t^{m.adj.}) (\mathbf{I}_t^{m.adj.})' \quad (3.36)$$

where $\mathbf{I}_t^{m.adj.}$ is the matrix of M mean adjusted indices (the mean subtracted from each element in the time series), each of length N and $\hat{\Omega}$ is estimated by

$$\hat{\Omega} = \frac{1}{N-M-1} \hat{\mathbf{E}} \hat{\mathbf{E}}' \quad (3.37)$$

with

$$\hat{\mathbf{E}} = \begin{pmatrix} \hat{e}_{11} & \hat{e}_{12} & \dots & \hat{e}_{1N} \\ \hat{e}_{21} & \hat{e}_{22} & \dots & \hat{e}_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{e}_{p1} & \dots & \dots & \hat{e}_{pN} \end{pmatrix}. \quad (3.38)$$

The estimated return of the portfolio $\hat{\mathbf{E}}(P) = \mathbf{W}'\hat{\mathbf{R}}$ is calculated in the same manner as in the previous section.

3.2 Empirical Study: Multiple Index Model

3.2.1 The Data

The same 20 portfolio constituents were used as in the Single Index empirical study. For the indices a range of 10 possible candidate indices or factors were considered.:

- The **JSE Overall Index** (JSE);
- **Anglo Gold** share price (anggol) as a proxy for a gold index;
- The **Dow Jones Transport** index;
- The **Gold Price** in ZAR;
- **Impala Platinum Holdings Ltd** share price as a proxy for a platinum index;
- **Richemont Securities** as a proxy for international share markets;
- **Palabora Mining Company Ltd** (palam) as a proxy for the copper index;
- The **FTSE100** Index;

- The price of **Crude Oil** per Barrel
- The **Rand Dollar** (dollar) exchange rate in Rands.

It was unfortunate that we were forced to use the shares as indices. However only the JSE Overall Index remained unchanged over the period from August 1988 to February 2007. Most of the current FTSE/JSE indices were only launched in the last 5 years and therefore did not have enough data to be used. It was decided to limit the multi index models to only four indices. Therefore a variable selection procedure was applied to select four indices from the set of ten.

3.2.1 Variable Selection

The process of selecting the four indices was as follows:

1. Run multi index regression models against all indices for all of the 20 shares.
2. The t-statistics of the 10 coefficients were tabulated. The number of absolute t-statistics greater than 2 (representing a greater than 95% significance) was calculated.
3. The two indices with the lowest number of absolute t-statistics greater than 2 were removed. The Dow Jones Transport and Gold Price in Rands were removed for the set. The procedure was repeated.
4. After the second round of elimination the FTSE100 and Richemont were removed.
5. After the third round of evaluating the t-statistics for the coefficients of the index for the twenty models, the Crude Oil price and Impala Platinum were removed.

3.2.2 Schwarz Criterion

The Schwarz Criterion, as used in EVIEWS, for the twenty regression models is shown in Figure 3.1 below. We compare the 10-index model with the reduced 4-index model for each regression. The Schwarz Criterion is calculated as follows:

$$SC = \frac{-2l}{N} + \frac{M \log(N)}{N}$$

where l is the log likelihood (assuming normally distributed errors), N is the number of observations and M is the number of indices. The Schwarz Criterion is an alternative to the Akaike Information Criterion that discriminates against a higher number of independent variables. A smaller value is preferred. Discussion on both the Schwarz and Akaike Information Criterion can be found in the EVIEWS3 help files.

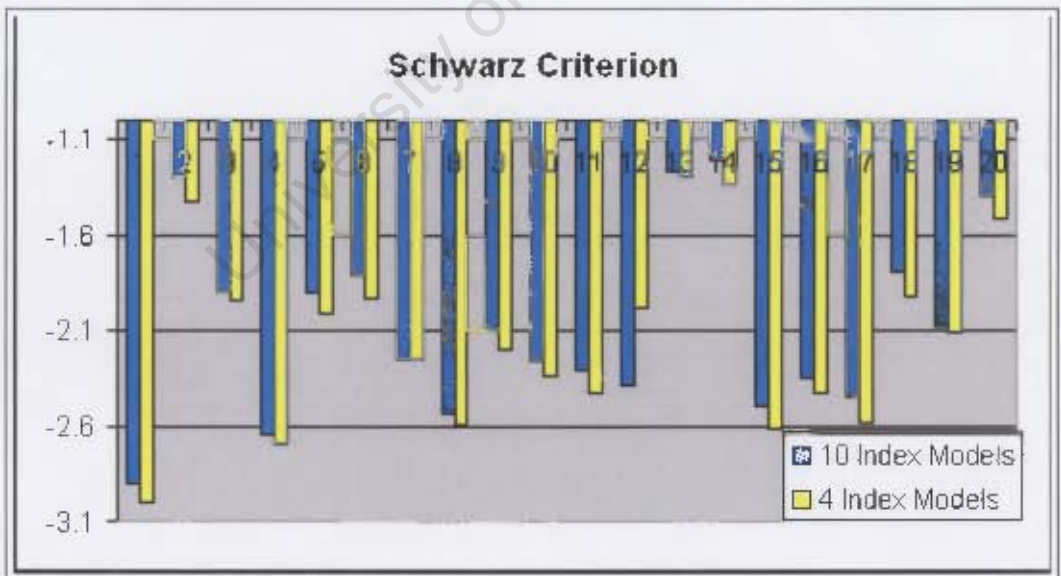


Fig. 3.1. Schwarz Criterion for the two sets on twenty regression models. We compare the 10 index model with the reduced four index model for each regression. The more negative Schwarz Criterion is preferred.

In each case, except share number 12, namely Anglo Platinum Ltd, the Schwarz Criterion improved with the reduced model. One reason why the model for Anglo Platinum was negatively affected by the reduced indices was that Impala Platinum, one of the candidate indices, was removed in the final step of variable selection. Anglo Platinum had a very strong beta coefficient with a t-statistic of 12.0 for the Impala Platinum independent variable, as can be expected, and removing this index had a detrimental impact on the model for Anglo Platinum.

The four indices used are as follows: JSE Overall Index, Anglo Gold, Palabora Mining Company Ltd, and the dollar price in Rands. The correlation matrix of the four indices, in that order, is shown below:

$$\begin{bmatrix} 1 & 0.505 & 0.2716 & -0.0039 \\ 0.505 & 1 & 0.0371 & -0.0123 \\ 0.2716 & 0.0371 & 1 & 0.03734 \\ -0.0039 & -0.0123 & 0.03734 & 1 \end{bmatrix}$$

As can be seen from the correlations above, only the JSE overall and Anglo Gold are significantly correlated (0.505). The rest of the indices are relatively uncorrelated.

3.2.2 Study Objectives

In this study we extend the single index model by introducing additional indices. The objective is to see the effect of introducing new indices on the risk return efficient frontiers of the different risk models. Also, we would like to see if the SMIM and ISMIM model frontiers have the same relative positions as the SIM and ISIM efficient frontiers. Also, we create baselines to measure the impact of additional models that still need to be discussed.

3.2.3 Methodology

After the variable selection process was completed, the reduced regression models were computed using EVIEWS3. Covariances of the indices were estimated using Equation (3.36). The \hat{E} matrix was calculated as in Equation (3.38). The $\hat{\Omega}$ and $\hat{\Theta}$ covariance matrixes were created as in Equations (3.37) and (3.35). These 20x20 matrixes were exported to MATLAB where the quadratic programming routine was applied to calculate the efficient frontiers and market portfolios. A VBA program in Excel was used to plot the results.

3.2.4 Primary Findings

After following the above methodology, the Markowitz, Sharpe Multi Index Model and Improved Sharpe Multi Index Model efficient frontiers were created and are displayed in Figure 3.2. The same profile between the Markowitz, SMIM and ISMIM was found as in the single index case. Again the efficient frontiers of the ISMIM and Markowitz models are basically on top of each other. In Figure 3.3 we compare the impact of adding additional indices to the location of the efficient frontiers of both the Sharpe and Improved Sharpe models.

In Figure 3.3 it is shown that the risk return profiles of the ISIM and ISMIM are very similar. Again this is predictable based on the fact that both are close to the same Markowitz frontier. The effect of using the SMIM compared with the SIM produces a less risky profile closer to the Markowitz frontier.

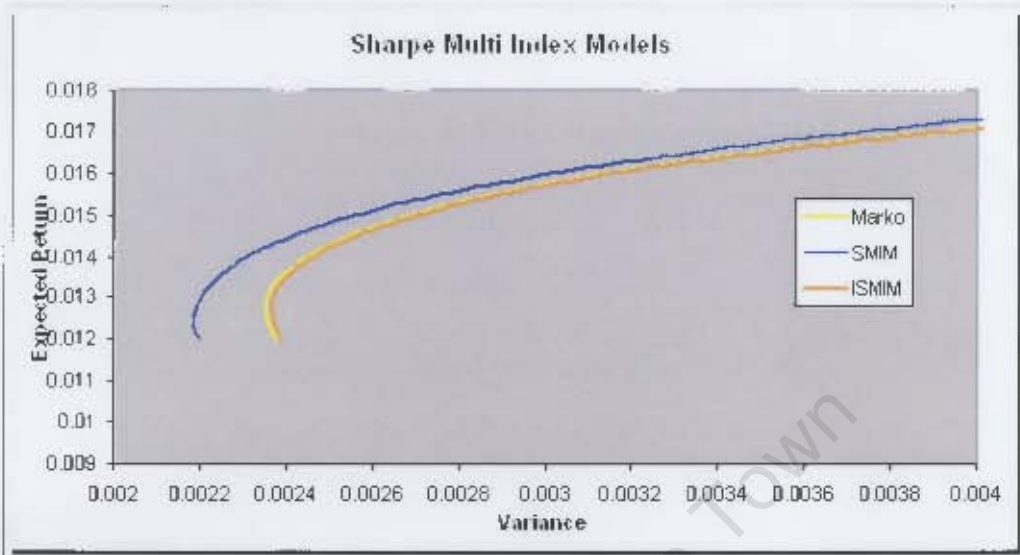


Fig. 3.2. The efficient frontiers for the SMIM and ISMIM compared with the Markowitz frontier.

3.2.5 Conclusion

We conclude that using a multi index regression model had minimal impact on the position of the ISMIM efficient frontier compared to that of the ISIM. It had, on the other hand, a significant impact on the resulting efficient frontier of the SMIM after the additional indices had been added. The effect was a right shift of the frontier, creating a model that looks more risky than in the single index case. This could be understood as follows: The SMIM incorporates more market information, using four indices rather than the one of the SIM, so less information is lost by disregarding the cross correlations in the residuals in the SMIM compared with the ISMIM. Because this information is included in both the ISMIM and ISIM models, the resulting frontiers are practically the same.

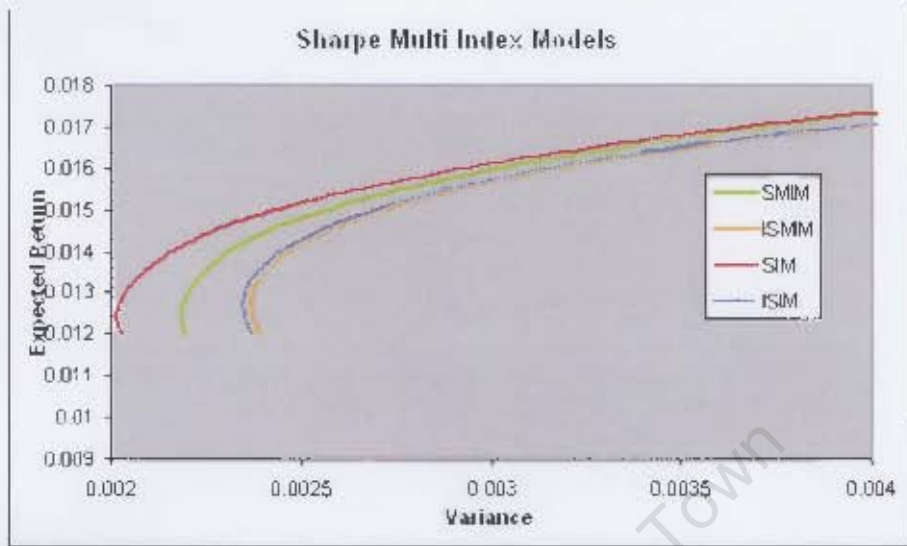


Fig. 3.3. Comparing efficient frontiers for the Multi Index and Single Index models.

3.3 Conclusion

In this chapter we introduced the Sharpe Multi Index Model and the improved version of the model that includes the residual cross correlation terms. In the empirical study we then demonstrated the effect of using the multiple indices on the mean-variance response of the efficient frontiers. The multi index model, that incorporates more information in the additional indices, reduces the left shift effect of the SMIM on the frontier compared with the SIM case. However, all the improved versions of the efficient frontiers of the models lie on top of each other, incorporating all available information in the model.

The ISMIM is therefore a more accurate model compared to the SMIM, confirming the observation of Hossain (2006). But this 'improvement' needs to be understood in the context that the ISMIM model results in the same efficient frontier as that of the original Markowitz model based on the historic covariance and therefore does not add any addi-

tional value for portfolio building. The one benefit these regression models have over the Markowitz model is their ability to be extended with ARMA and GARCH components. This is the topic of the next chapter.

University of Cape Town

Chapter 4

ARMA and GARCH Regression Models

In this chapter we extend our SIM and ISIM regression models of the shares by including ARMA terms that model serial correlation and GARCH extensions used to model conditional variance in the time series data. After introducing these models in section 4.1 and 4.2, we apply the methodology the JSE shares, first with a 9 share portfolio in sections 4.3 and then with the full set of 20 shares in section 4.4, and investigate the effect on the efficient frontier and the resulting market portfolios. We then investigate and evaluate which of these methodologies created better models.

4.1 Linear Regression Models with Autocorrelated Errors.

ARMA models (Box and Jenkins, 1970) are used to model the conditional expectation of the current observation, given the past observations. An ARMA model achieves this by modelling the current observation as a linear function of the past observation and error terms combined with a new error term. If a satisfactory ARMA is found, it can be used to forecast the next data point in the series conditional on all the previous data points. But, unfortunately, forecasting will not be possible in the context of a regression model. In the regression model the ARMA terms are fitted to the residuals of the regression and are therefore dependant on the independent variable (index in our case). So even if you have a perfect model to calculate the forecasted return R_{t+1} data point, you will need I_{t+1} data to do the calculation.

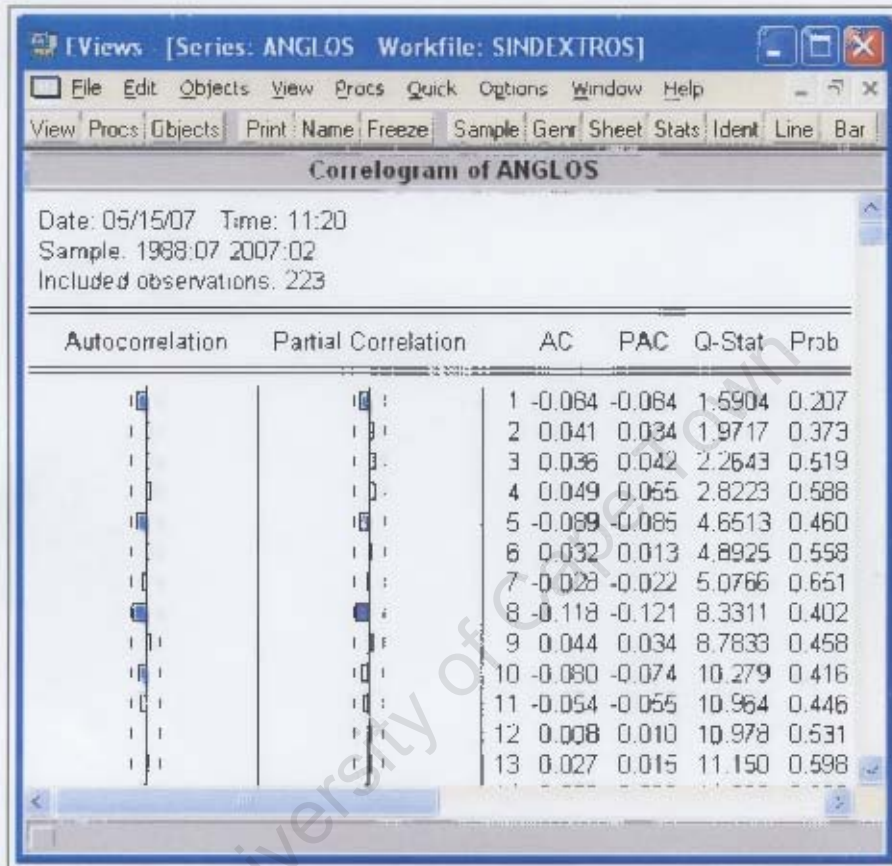


Fig. 4.4. Correlogram of Anglos returns.

between the share and the index. The Akaike information criterion (AIC) for this model was -2.876. Next we consider the correlogram of the index model. Comparing the correlograms in Figures 4.4 and 4.5, shows the serial correlation induced in the residuals. For example, on the first 6 months lag, the Ljung-Box Q-statistic shows autocorrelation within 10% significance (a probability of 0.074). This weak serial dependence in the residuals can be modelled by using linear regression with time series errors. Based on the correlogram in Table 4.2, we chose an $ARMA(2, 2)$ model in the residuals and modify the regression

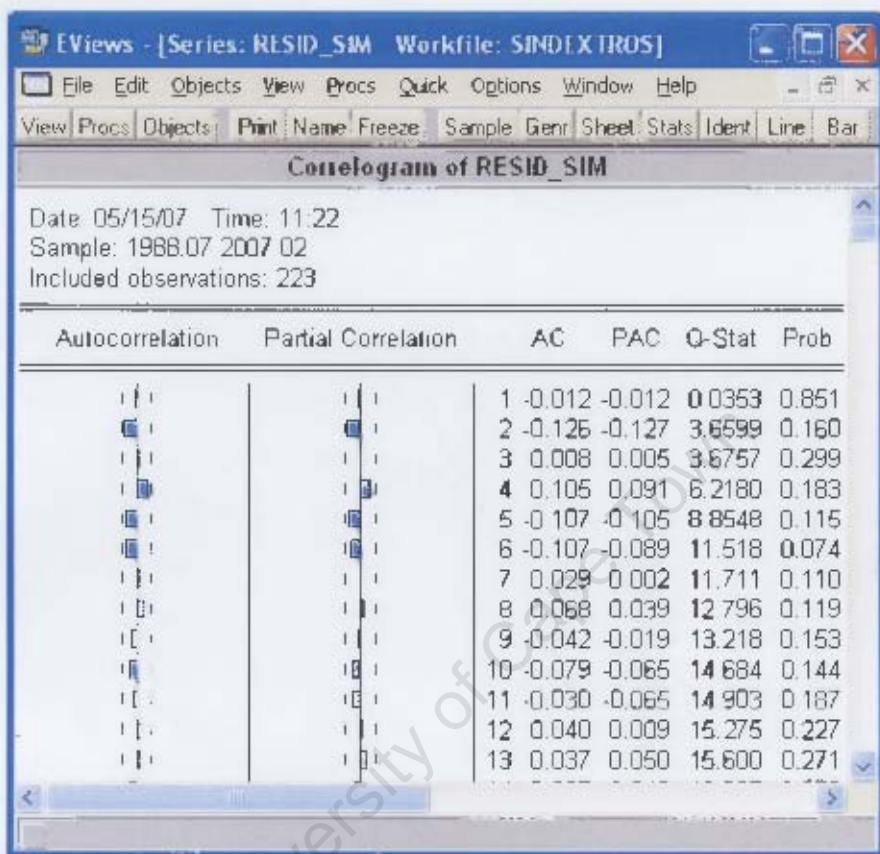


Fig. 4.5. Correlogram of Anglos Residuals

model to:

$$R_t = \alpha + \beta I_t + e_t \quad (4.39)$$

$$e_t = \phi_2 e_{t-2} + \theta_2 w_{t-2} + w_t \quad (4.40)$$

where w_t is white noise and both ϕ_1 and θ_1 are assumed to be zero. The model was estimated using EVIEW3, resulting in:

$$R_t = -0.00238 + 1.401179449I_t + e_t$$

$$e_t = -0.881e_{t-2} + 0.784w_{t-2} + w_t \quad (4.41)$$

$$\hat{\sigma}_w = 0.0537, \quad (4.42)$$

with $R^2 = 0.695$ and standard errors on the parameters for the index I_t , $AR(2)$ and $MA(2)$ of 0.0629, 0.102 and 0.134 respectively. They were all highly significant with the AIC equal to -2.991. Figure 4.6 shows the correlogram of the residuals of the model. By adding the ARMA terms in the residuals, we are able to remove the induced time series behaviour of the residuals and create a more accurate model with a slightly higher R^2 . This improved the AIC from -2.876 to -2.991, gave a smaller standard error on the beta estimation of the index and resulted in no autocorrelation in the residuals.

The equations (4.39) and (4.40) can be generalized to include SIM (or SMIM) with any $ARMA(p, q)$ error terms as:

$$R_t = \alpha + \beta I_t + e_t$$

$$e_t = \phi_1 e_{t-1} + \dots + \phi_p e_{t-p} + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q} + w_t \quad (4.43)$$

$$w_t \sim rv(0, \sigma_w^2), \text{ some random variable independent in } t.$$

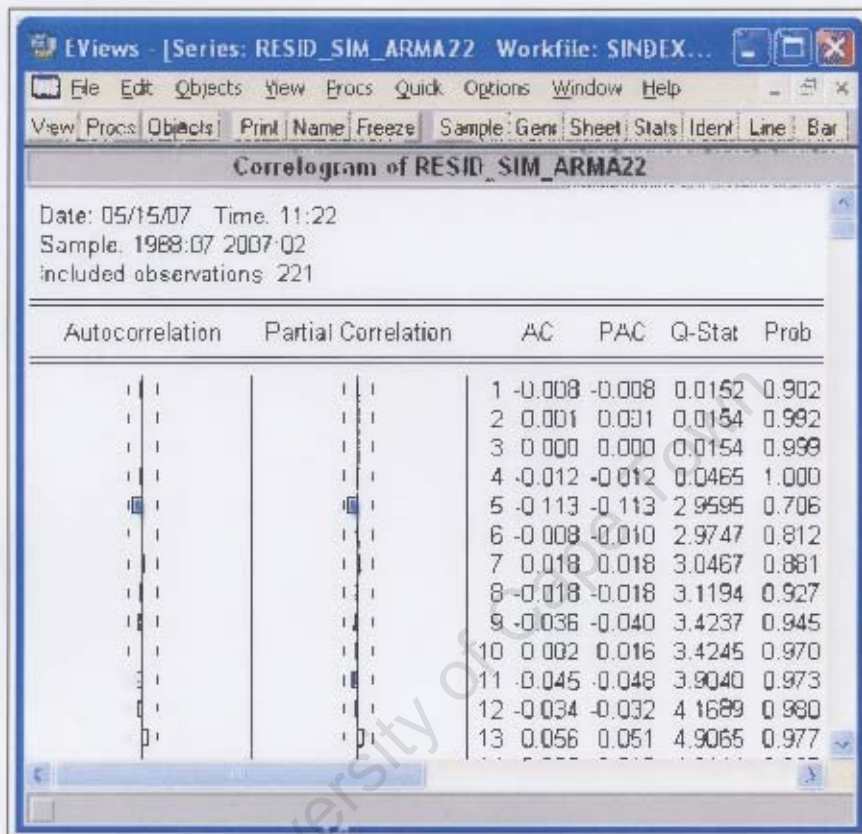


Fig. 4.6. The Residuals of the SIM for an ARMA(2,2) model.

4.2 GARCH Models

There are several reasons why one might want to model volatility. In finance you might need to analyze the risk in a portfolio of assets or to evaluate an option. ARMA models on their own are unsatisfactory for modelling of returns of financial assets with changing volatility. The reason for this is that ARMA models assume constant conditional variance and are not capable of modelling changes in volatility observed in the return of financial assets (Ruppert, 2004, p363).

Autoregressive Conditional Heteroscedasticity (ARCH) models were introduced by Engle in 1982 and are specifically designed to model and forecast conditional variance or volatility. In the model, the variance of the dependant variable is modelled as a function of the past values of dependant and independent variables.

The ARCH model was generalized by Bollerslev (1986) and Taylor (1986) as the Generalized *ARCH* (*GARCH*) model. In a *GARCH* process the variance is modelled as a function of past values of the dependant variable, independent variable and past variance values.

These models are widely used in financial time series analysis. The reason for this is that standard regression type models assume that the residuals from the regression model are homoscedastic. When this assumption is violated, as is the case in many financial time series data, the results should be adjusted in order to compensate for heteroscedastic errors. So in using *GARCH* type models, where the heteroscedasticity in the errors is handled properly, more efficient estimators can be obtained.

It is also commonly known that share returns have more extreme values than can be expected from the standard assumptions of the normal distribution. One approach is to incorporate these outliers in the model by assuming that the conditional variance is not constant but heteroscedastic. In such models, outliers occur naturally when the variance is large. Therefore *GARCH* processes are categorized by changing variances as well as heavy tails. This makes *GARCH* models especially useful in modelling financial data.

A general $GARCH(p, q)$ model in the residuals of a regression model is given by the following equations:

$$R_t = \alpha + \beta I_t + e_t \quad (4.44)$$

$$e_t = \sigma_t z_t \quad \text{where } z_t \sim rv(0, 1) \quad (4.45)$$

$$\sigma_t^2 = \gamma_0 + \sum_{j=1}^q \gamma_j e_{t-j}^2 + \sum_{j=1}^p \omega_j \sigma_{t-j}^2 \text{ for } t = 1, \dots, N. \quad (4.46)$$

Here e_t are the residuals of the regression for each time step t and z_{it} represents a random error with mean zero and variance one. The parameter p is the degree of the $GARCH$ component and q is the degree of the $ARCH$ process. Since the equations above express the dependence of the variance of returns in the current period on historic data (i.e. the values of the variables e_{t-j}^2 and σ_{t-j}^2) from previous periods, it is a conditional variance. The most basic and widespread version of this model, and also the one used in this mini-dissertation, is the $GARCH(1,1)$ that can be expressed as:

$$R_t = \alpha + \beta I_t + e_t \quad (4.47)$$

$$e_t = \sigma_t z_t, \quad z_t \sim rv(0, 1) \quad (4.48)$$

$$\sigma_t^2 = \gamma_0 + \gamma_1 e_{t-1}^2 + \omega_1 \sigma_{t-1}^2 \text{ for } t = 1, \dots, N, \quad (4.49)$$

where z_{it} represents a random error with mean zero and variance one. The variance is always positive, so we expect that the regression coefficients γ_0 , γ_1 and ω_1 will also be positive. The conditional variability of the errors defined in Equation (4.47) is determined by the constant part, γ_0 , the $ARCH$ component $\gamma_1 e_{t-1}^2$ in the previous error and the $GARCH$ component given by the predicted variability from the previous period and expressed by $\omega_1 \sigma_{t-1}^2$.

The sum of the regression coefficients $(\gamma_1 + \omega_1)$ expresses the influence of the variability of variables from the previous period on the current value of the variability. This value is usually close to 1.0, which is a sign of increased inertia in the effects of shocks on the variability of returns on financial assets. This value is required to be smaller than 1 for a stable $GARCH(1, 1)$ process (Tsay, 2002, p.94)

4.2.1 Properties of the GARCH models

If one sets $v_t = e_t^2 - \sigma_t^2$ in Equation (4.49), it follows that:

$$\begin{aligned}\sigma_t^2 &= e_t^2 - v_t, \\ \omega_1 \sigma_{t-1}^2 &= \omega_1 (e_t^2 - v_t) \quad \text{and} \\ e_t^2 &= \gamma_0 + (\gamma_1 + \omega_1) e_{t-1}^2 + v_t - \omega_1 v_{t-1}.\end{aligned}\tag{4.50}$$

So, the $GARCH(1,1)$ model can be regarded, based on Equation (4.50), as an $ARMA(1,1)$ process in the squared error e_t^2 series. Furthermore, from the properties of the $ARMA$ model it follows that the unconditional variance of e_t is:

$$Var(e) = E[e^2] = \frac{\gamma_0}{1 - \gamma_1 - \omega_1},\tag{4.51}$$

provided $\gamma_1 + \omega_1 < 1$.

As is evident in Equation (4.49), a large value of e_{t-1} or σ_{t-1} produces a large value of σ_t^2 . So, large values of e_t^2 tend to be followed by large values of e_{t+1}^2 producing the volatility clustering observed in financial time series.

A $GARCH$ series produces heavier tails than the normal distribution. This can be observed if one analyses the unconditional kurtosis of e_t . It can be shown under the

assumption $1 - 2\gamma_1^2 - (\gamma_1 + \omega_1)^2 > 0$, that the kurtosis is given by (Tsay 2002, p.94)

$$\gamma_2 = \frac{E(e_t^4)}{[E(e_t^2)]^2} = \frac{3(1 - (\gamma_1 + \omega_1)^2)}{1 - (\gamma_1 + \omega_1)^2 - 2\gamma_1^2} > 3$$

and therefore the distribution has fatter tails than can be expected from the normal distribution, again confirming the link between volatility clustering and fat tails. Another interesting property of the GARCH(1,1) model is that the unconditional multi-step ahead forecast converges to the unconditional variance given in Equation (4.51), as the forecast horizon increases, that is for:

$$\sigma_{t+l}^2 \rightarrow \frac{\gamma_0}{1 - \gamma_1 - \omega_1} \text{ as } l \rightarrow \infty.$$

under the condition that $\gamma_1 + \omega_1 < 1$ and that $\text{var}(e_t)$ exists.

4.3 Empirical Study: Comparing a 9 share and a 20 share portfolio.

In this study we build efficient frontiers using the first 9 shares presented in Chapter 3 of the set and compare them to the efficient frontiers using the full set of 20 shares. We demonstrate consistency in the results when using either a set of 9 shares, as was used in the study of Hossain (2006), or a set of 20 shares as is used in this dissertation.

4.3.1 The Data

The same data set was used as in the previous empirical studies. It is a set of 20 shares and 9 selected from them, including the JSE All Share Index. The data are monthly returns spanning the period from September 1988 to February 2007.

4.3.2 Study Objectives

The objective of this study is to compare the relative positioning of the efficient frontier generated by the different models using sets of 9 and 20 shares.

4.3.3 Methodology

The following models are considered:

- Marko is the model based on the historic covariance.
- SIM is the Sharp Index Model of Chapter 2.
- ISIM is the Improved Sharp Index Model used in Chapters 2 and 3.
- ARG is an autoregressive with GARCH(1,1) time series model introduced earlier in this chapter. In this case unique AR components were fitted to each time series of residuals to minimize the autocorrelation. The specific model fitted for each share is summarized in Table 4.1. The ARG model was used with both SIM and ISIM.

4.3.4 Primary Findings

The following results were calculated and compared with the results for a portfolio of 9 shares from the Ph.D. of Hossain (2006) and results for the 20 shares used in this dissertation.

- Comparison of the SIM, ISIM and Markowitz efficient frontier positioning.
- Comparison of the ARG, SIM and SIM models efficient frontier positioning.

- Comparison of the ARGA, ISIM and ISIM models efficient frontier positioning.

The relative positions of the efficient frontier of the above mentioned models are shown in Figures 4.1 and 4.2 and were found to be consistent, irrespective of using sets of 9 or 20 shares. The results was also found to be consistent with Hossain's results (2006)

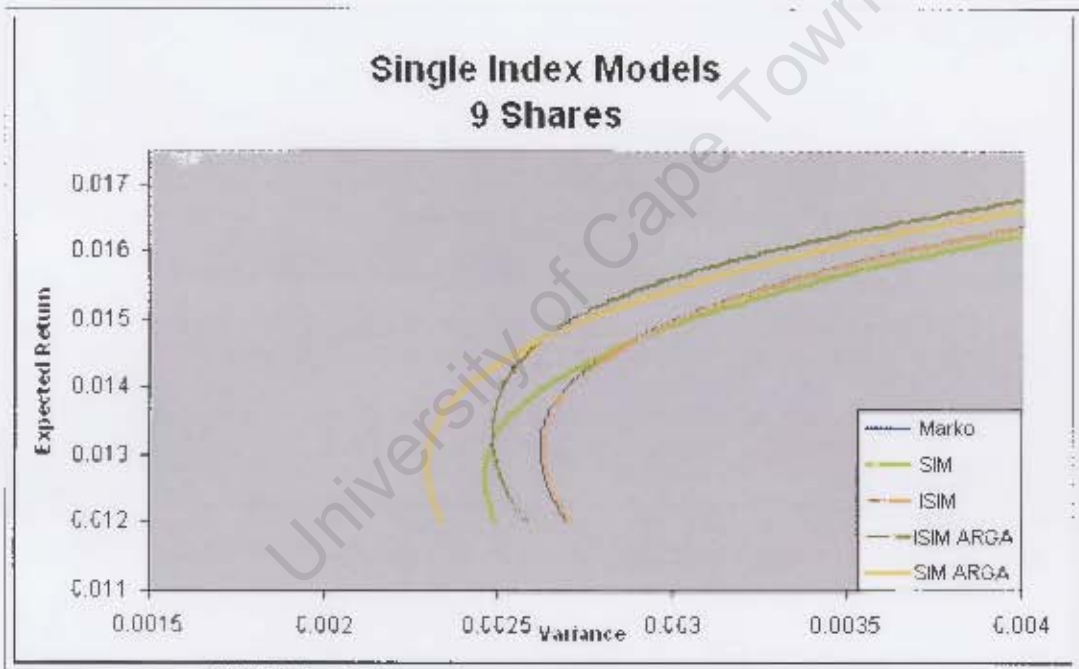


Fig. 4.7. Single Index Models for the set of 9 shares.

Comparing SIM with ISIM

In all cases presented in Figure 4.7 and Figure 4.8 the efficient frontier of the Sharpe index model is positioned to the left of the Markowitz and Improved Sharpe Index models. This was found to be the case when positively correlated residuals are dominant. In Chapter 6 of this mini-dissertation we introduce the measurement of the average residual correla-

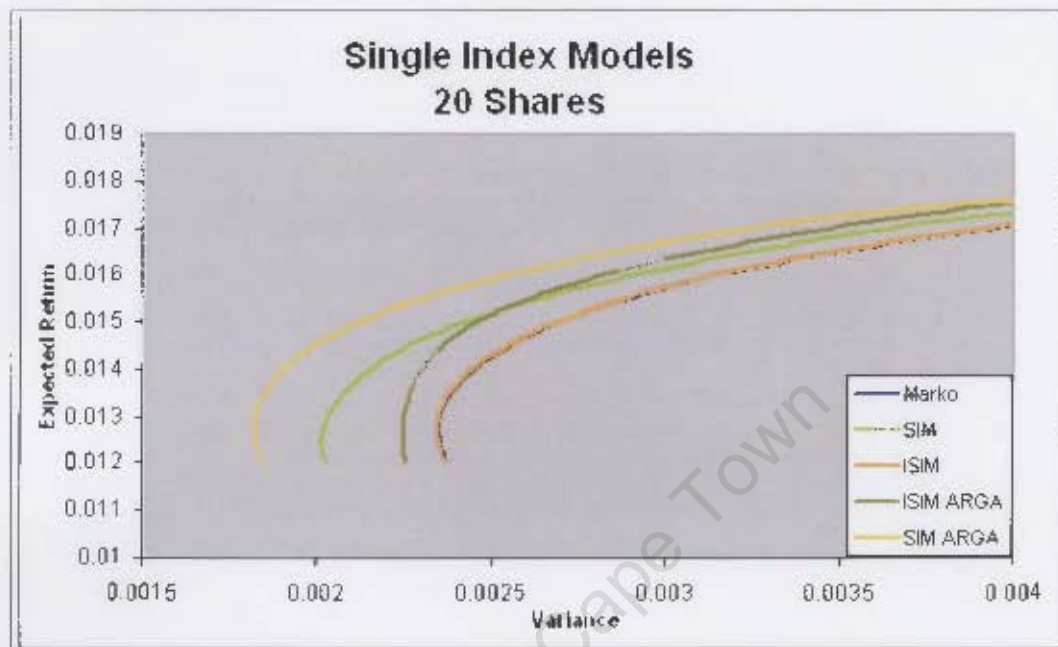


Fig. 4.8. Single Index Models for the set of 20 shares

tion and show that the relative position of the SIM efficient frontiers can be predicted by comparing the average residual correlation. Indeed we found a case where average residual correlation is negative and the efficient frontier of the SIM is to right of the corresponding ISIM efficient frontier.

These results show that SIM underestimates the risk of the portfolio compared to the ISIM model. This is due to the fact that less of the covariance structure information of the constituents is used in the construction of the SIM. The improvement of the ISIM is that the unused covariance structure of the SIM is replaced by the estimated covariance structure of the residuals of the constituents.

Comparing The ARGA model extensions

The efficient frontiers of the AR with GARCH(1,1) extension to the SIM and ISIM models presented in Figure 4.7 and Figure 4.8 show the left shift effect. Again this result was consistent irrespective of whether we use 9 or 20 shares. This extension did create slightly better models, but results varied as is shown in the next empirical study.

4.3.5 Conclusion

Based on these results we conclude that consistent results are obtained compared to previous studies using a set of only 9 shares. In the next study we look in more detail at the regression model with ARMA and GARCH extensions using the set of 20 shares.

4.4 Empirical Study: Regression models with ARMA and GARCH

In this empirical study we will extend our index regression models by including ARMA and GARCH terms in our model. We will investigate whether these extensions produce improved models in terms of improved fit of the model, more accurate estimators of the coefficients of the equation and less correlated and more independent residuals than the simplified models.

Then we will compare the risk-return behaviour of the portfolio created by these models as illustrated by the efficient frontiers of the models applied to our sample set of shares from the JSE. This will illustrate the effect of including ARMA and GARCH components into the models. We will demonstrate the shifting effect that they have on the resulting ef-

ficient frontiers. We then consider the portfolio members and their corresponding weights in the resulting market portfolios.

4.4.1 The Data

The same data set was used as in the previous empirical studies. It is a set of twenty shares and the JSE All Share Index. The data are monthly returns spanning the period from September 1988 to February 2007.

4.4.2 Study Objectives

Our objective is to establish the effect of using ARMA and GARCH(1,1) models on the residuals for both the SIM and ISIM models. We also want to compare the behaviour of these extensions for the SMIM and ISMIM models.

4.4.3 Methodology

It was decided to limit the study to the simple GARCH(1,1) extended models. The reason for this is that it has been shown by Hossain, Troskie and Guo (2005a) that the GARCH(1,1) model exhibits “the most superior mean-variance frontiers when compared to other GARCH extensions in the single index setting.” We will therefore limit ourselves to:

- **SIM**, the Sharpe Index Model
- **ISIM**, the Improved Sharpe Index Model
- **SIM ARMA**, a Sharpe Index Model with ARMA time series components.

- **ISIM ARMA**, an Improved Sharpe Index Model with ARMA time series components applied to each share regressed against the index.
- **SIM GARCH**, with GARCH(1,1) applied to the regressions.
- **ISIM GARCH**, with GARCH (1,1) applied to the regressions.
- **SIM ARGA**, with ARMA and GARCH(1,1) applied to the regressions.
- **ISIM ARGA**, the Improved Sharpe Index Model with ARMA and GARCH(1,1).

Also the study was repeated with the Sharpe Multi Index Models (SMIM) with the same set of chosen indices of the previous chapter. The above eight models were repeated for the multi-index case, with the model names stipulated below:

- **SMIM**, the Sharpe Multi Index Model
- **ISMIM**, the Improved Sharpe Multi Index Model
- **SIM ARMA**, a Sharpe Multi Index Model with ARMA time series components applied to each share regressed against the index.
- **ISMIM ARMA**, an Improved Sharpe Multi Index Model with ARMA time series components applied to each share regressed against the index.
- **SMIM GARCH**, with GARCH(1,1) applied to the regressions.
- **ISMIM GARCH**, with GARCH (1,1) applied to the regressions.
- **SMIM ARGA**, with ARMA and GARCH(1,1) applied to the regressions.

- ISMIM ARGA, the Improved Sharpe Index Model with ARMA and GARCH(1,1) applied to the regressions.

ARMA Identification

The ARMA time series model was identified for each of the shares using the residuals for both the single and multi index models. The partial autocorrelation function was used to identify the specific lag orders. ARMA models were only applied if they produced a significant reduction in the autocorrelation of the residuals compared to no ARMA terms. AR terms were preferred over MA terms and were found to be adequate to model the time series dependencies of the regression residuals. A table of the resulting ARMA terms for each share is shown in Table 4.1.

Eviews Code	Name	Code	ARMA Single Index	ARMA Multiple Index
r1	Ang	AGL	AR(1), AR(2)	AR(2)
r2	JDG	JDG	None	None
r3	Pick	PIK	AR(1)	AR(1)
r4	Remgro	REM	AR(1)	AR(1)
r5	SAEagle	SAE	None	None
r6	Sappi	SAP	None	None
r7	sasol	SOL	None	None
r8	Tigbran	TBS	AR(6)	AR(1) AR(6)
r9	longaat	INI	AR(2)	AR(2)
r10	absa	ASA	None	AR(1)
r11	afrox	AFX	AR(1)	AR(1)
r12	angpla	AMS	None	None
r13	harmo	HAR	None	None
r14	jonnice	JNC	AR(1), AR(2)	AR(1)
r15	libert	LBL	AR(7)	AR(1) AR(7)
r16	nampak	NPK	AR(3) AR(4)	AR(1) AR(4)
r17	nedcor	NED	None	None
r18	reun	RLO	None	None
r19	PPC	PPC	AR(2)	AR(2)
r20	edcon	ECO	AR(1)	AR(2)

Table 4.1 ARMA residual identification.

University of Cape Town

The ARMA models we are using are a combination of $ARMA(p, q)$ models to remove the serial correlation and $GARCH(1, 1)$ models to model the volatility of the residuals. The full regression model can be summarized as follows:

$$R_{it} = \alpha_i + \beta I_t + e_{it} \quad (4.52)$$

$$e_{it} = \phi_{i1} e_{i,t-1} + \theta_{i1} a_{i,t-1} + a_{it} \quad \text{assuming } ARMA(1, 1) \quad (4.53)$$

$$a_{it} = \sigma_{it} z_{it} \quad (4.54)$$

$$\sigma_{it}^2 = \gamma_{i0} + \gamma_{i1} a_{i,t-1}^2 + \omega_{i1} \sigma_{i,t-1}^2 \quad \text{assuming } GARCH(1, 1) \quad (4.55)$$

$$i = 1, \dots, p, \quad t = 1, \dots, N, \quad (4.56)$$

where z_{it} has mean zero and variance one. The standard assumption is that $z_{it} \sim N(0, 1)$ and is the one we will follow. The $ARMA$ only model excludes the terms in Equation (4.55) and the $GARCH$ only model excludes the terms in Equation (4.53)

EViews 3 uses a nonlinear regression technique to estimate the ARMA models. More detail can be found in the EViews 3 help files.

ARCH-GARCH type models are estimated in EViews 3 by the method of maximum likelihood, under the assumption that the errors are conditionally normally distributed. Because the variance appears in a non-linear way in the likelihood function, the likelihood function needs to be maximized using iterative algorithms.

An EViews 3 program was created to calculate the returns, perform the regressions, and to produce the covariance and return estimates for each of the models. The equations

from Chapter 2 were used to estimate the covariances for the SIM and ISIM:

$$\hat{\Phi}_{SIM} = \hat{\sigma}_I^2 \hat{\beta} \hat{\beta}' + \hat{\Omega}_{diag} \tag{4.57}$$

$$\hat{\Phi} = \sigma_I^2 \hat{\beta} \hat{\beta}' + \hat{\Omega}. \tag{4.58}$$

These 20×20 matrices were then exported to MATLAB 7 and the quadratic programming problem was solved producing the efficient frontiers and associated portfolios. The programs can be found in the Appendix A and B respectively.

4.4.4 Primary Findings

Model Building

From a model building point of view, the added complexity of the ARMA and GARCH models has a beneficial effect on the estimation of beta value for the shares. This can be seen, from the higher value of the average t-statistic for the shares in Table 4.2 below.

Average	SIM	SIM ARMA	SIM GARCH	SIM ARGA
Beta T Statistic	9.326	9.446	10.246	10.813
adjusted R ²	0.2765	0.2913	0.2733	0.2866
Standard Error	0.0852	0.0846	0.0860	0.0854

Table 4.2 Comparing model statistics

In Figure 4.9 we display the t-statistic for each of the 20 shares for the SIM, SIM ARMA and SIM GARCH and SIM ARGA, labeled as SIM, ARMA, SIM GARCH ONLY, SIM and ARMA GARCH respectively, and in that order in the figure. We notice that the effect of using GARCH in the model had varied results between the shares. It has a strong positive improvement on equities numbered 1, 2, 12 and 14. The last value, called "share

21", is the average for the 20 shares. The average value shows slight improvement in the estimation of the betas. If one considered the adjusted R^2 of each of the regressions for

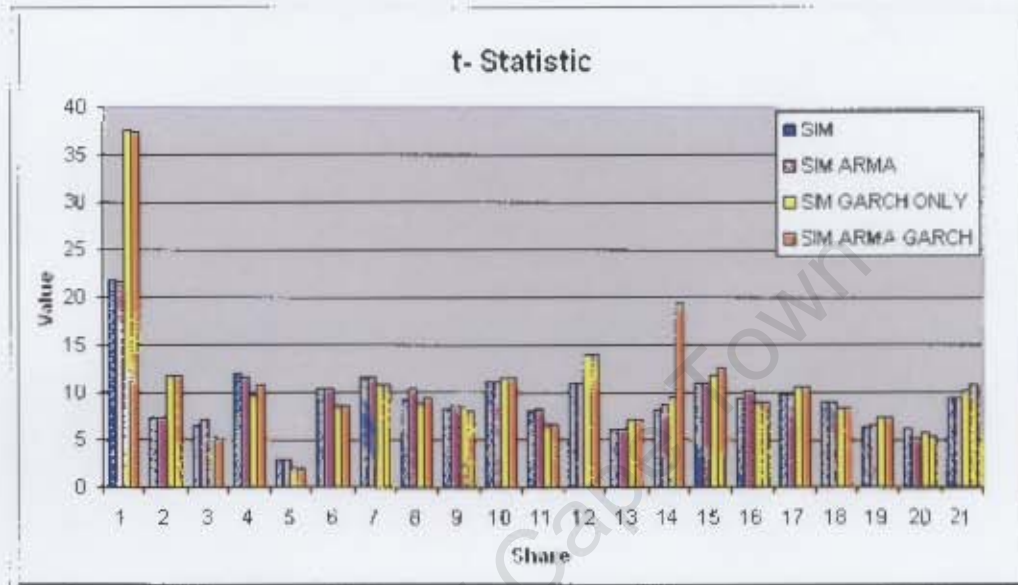


Fig. 4.9. The t -statistic of the β 's in the four models (from left to right): SIM, SIM ARMA, SIM GARCH and with ARMA and GARCH(1,1)

the different models in Table 4.2 above and the in Figure 4.10 below, we notice that both the 'ARMA' and 'ARGA' extension had a minimal improvement on the model fit. ARMA produced a better fit than the 'ARGA' extension and GARCH ONLY combination. This behavior was consistent for all the equities. Table 4.2. above also presents the average standard errors of the shares for the three different models. This standard error is the standard deviation of the residuals for each regression. As expected, the ARMA extension of the model reduced the average standard error of the regression models. The ARMA-

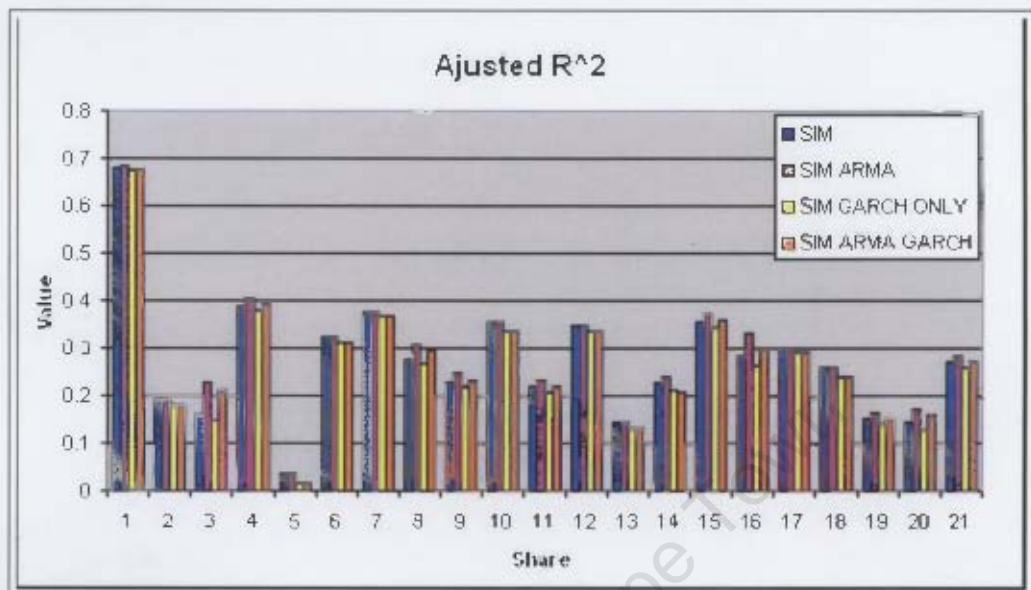


Fig. 4.10. Adjusted R^2 from left to right, SIM, SIM ARMA, SIM GARCH (1,1) and SIM ARGA. 21 refers to the average value of the 20 regressions models.

GARCH extension had an increasing effect on the observed average error. These changes are small and in all cases resulted in less than 1% change from the original average value.

Mean Variance Response

In this section we consider the mean variance response of 16 different models. First we consider the effect of five different estimation methods: Historic covariance (called Marko); OLS; ARMA; GARCH, ARMA and GARCH called ARGA as introduced in this chapter. These estimation methods are applied to both the SIM and ISIM methods to construct the covariance matrices, as introduced in Chapter 2. We then repeat the exercise in the multi index framework to create the SMIM and ISMIM's as introduced in Chapter 3.

We start by looking at the impact of changing the estimation methods on the SIM for our portfolio of shares. Figure 4.11 shows the resulting efficient frontiers from using

different estimation methods. From right to left the frontiers are: the Markowitz frontier with no regression model applied, the SIM and the SIM ARMA close to each other. The SIM GARCH and SIM ARGA together on the far left. This pattern is repeated for the multi

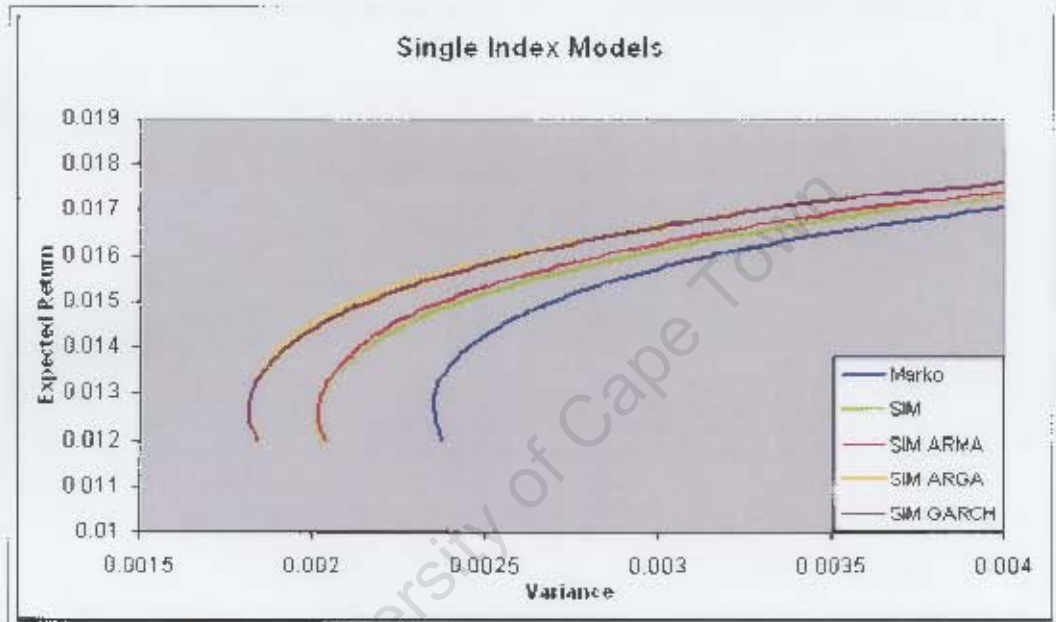


Fig. 4.11. From right to left the frontiers are: the Markowitz frontier with no regression model applied; the SIM and SIM ARMA close to each other. The SIM GARCH and SIM ARGA together on the far left.

index case in Figure 4.12. Again from right to left the frontiers are: the same Markowitz frontier as before; followed by the SMIM and SMIM ARMA on top of each other with only SIM ARMA visible in the figure, the SMIM GARCH and to the far left the SMIM ARGA. Again, as was seen in Chapter 3, all the resulting frontiers lie further to the right, closer to the Markowitz frontier, compared to the single index models, and for the same reason as was discussed in the conclusion of Chapter 3. The Sharpe Multi Index Models incorporate

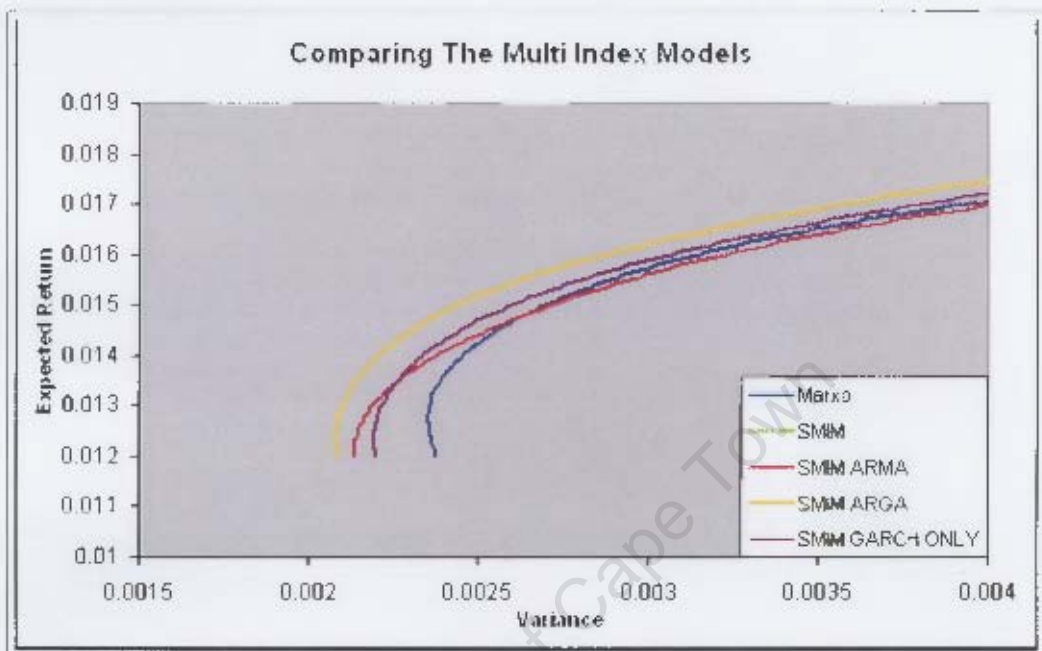


Fig. 4.12. The SMIM efficient frontiers From left to right: SMIM ARGA, SMIM ARMA, SMIM GARCH(1,1), Sim and Marko extensions for the 20 stock portfolio. SMIM is underneath the SMIM GARCH frontier.

more market information in the covariance matrix than the SIM and are therefore moved close to the ISIM and Markowitz frontiers.

In the SIM case, the GARCH(1,1) and the combined ARMA GARCH(1,1), called ARGA, has a left-shift effect on the SIM frontiers. For the SMIM, the GARCH(1,1) had no effect, and only the combined ARGA model has some left shift effect on the frontier as can be seen in Figure 4.12. This lesser effect of multi index models to residual model extensions than the single index model could be explained as follows. The residuals from a multi index model are on aggregate smaller than the single index model. This can be seen from the standard errors of regression. They are on average, over the twenty shares, 4.17%

smaller for the multi index case compared with the single index models. Modelling the residuals of multi index models has less effect than on single index models.

We now repeat the investigation for the Improved Sharpe Models (ISIM). In these models the full covariance structure of the residuals is used to construct the covariance matrices. In Figure 4.13 four efficient frontiers are shown, namely from right to left, ISIM ARMA, ISIM, ISIM ARGA and ISIM GARCH. The first two are grouped on the right and the second grouped to the left. Adding the ARMA extension had the same minimal effect on both the ISIM and ISIM GARCH models, as can be seen from Figure 4.13. In Figure

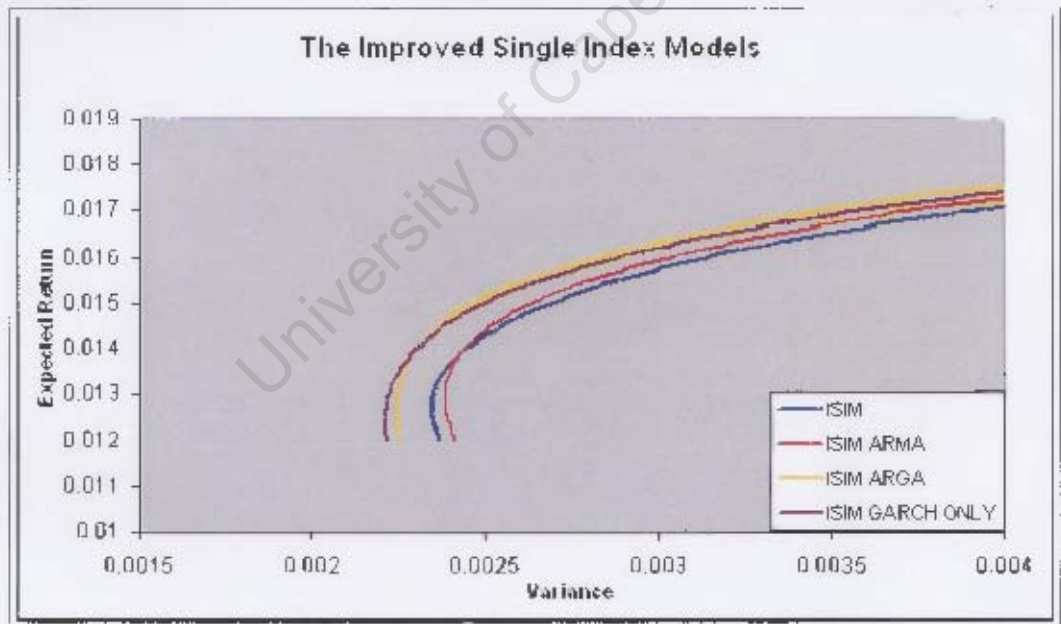


Fig. 4.13. Four frontiers are shown, from right to left, ISIM ARMA, ISIM, ISIM ARGA and ISIM GARCH.

4.14 we display the results of the multi index case. As was the case in Figure 4.12, by using the multi index regression the resulting frontiers are compressed closer to the original

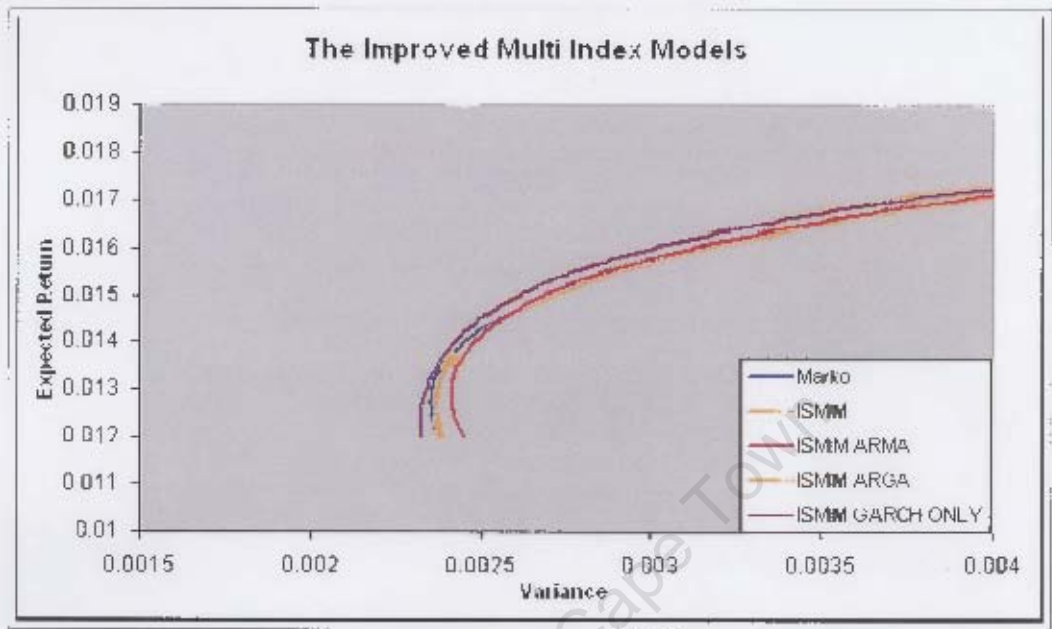


Fig. 4.14. The Improved Multi Index Models: ISMIM ARMA is on the right and ISMIM GARCH is on the left and Marko, ISMIM, ISMIM ARGA in the center.

Markowitz frontiers. The ISMIM ARMA is on the right and ISMIM GARCH is on the left and Marko, ISMIM, ISMIM ARGA in the center. Using multi index models seems to make the effect of modelling the residuals less potent, keeping the resulting frontiers closer together.

The market portfolios

Even though different models have different shifting effects on the resulting efficient frontier, it does not necessarily affect the resulting market portfolios. In this section we examine the resulting portfolios of the models presented. The market portfolios were constructed assuming an annual risk free rate of 9%. If we examine the annualized Sharpe ratios of the portfolios in Figure 4.7, focusing on the single index models, we see that the

highest Sharpe ratio of 0.56 for the SIM ARGA portfolio corresponds to the SIM ARGA efficient frontier position to the far left in Figure 4.11. The SIM GARCH Sharpe ratio

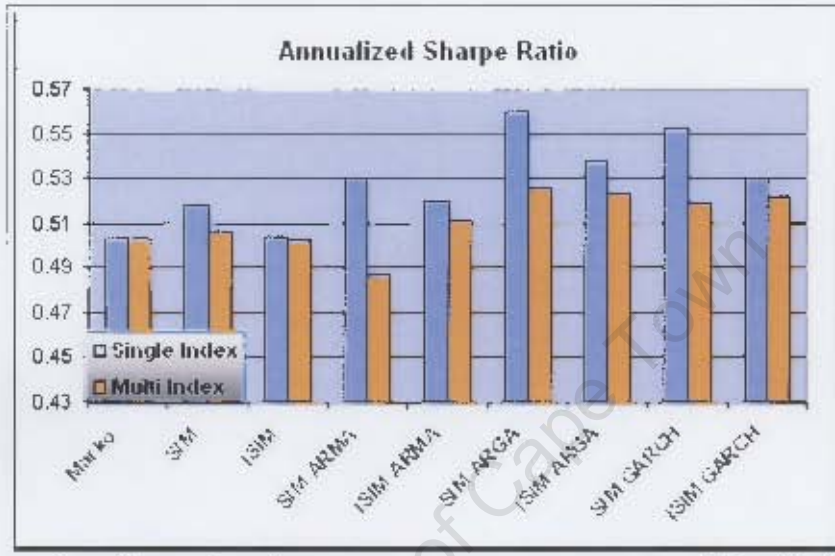


Fig. 4.15. Sharpe Ratios of the Market portfolios of the single and multi-index type models assuming an annual risk free rate of 9%.

is very close (0.552) to this value and also the corresponding frontier in Figure 4.11. The remainder of the Sharpe ratios follow in order of the positioning of the corresponding frontiers in the preceding four figures. The ISIM's Sharpe ratios are less dispersed, and again these mimic the behaviour of the frontiers in Figure 4.13. The ISIM and ISMIM Sharpe ratios are very close to the original Marko portfolio. The Sharpe ratios of the multi index models are also observed to be lower than the corresponding single index models. This can be verified by comparing Figures 4.11 through Figure 4.14.

In Figures 4.16 and 4.17 we show the expected returns and standard deviations of the same portfolios. Figure 4.16 shows us that the expected returns of the portfolio do not al-

ter much, with only a 1.03% spread in the returns. The small variation can be explained by the variation in the resulting portfolios. In Figure 4.17 we see that the standard deviation

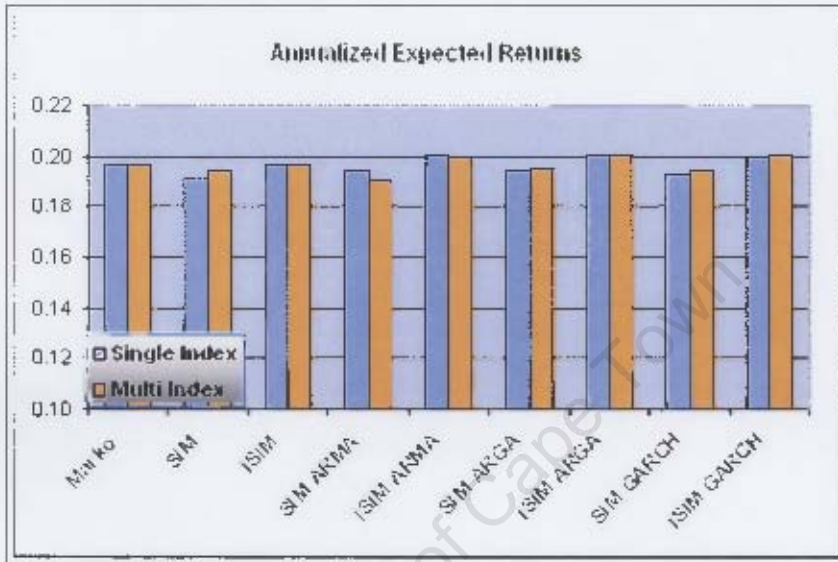


Fig. 4.16. The Annualized Expected Returns of the portfolios.

of the portfolio inversely corresponds with Sharpe ratios and the positioning of the frontiers from left to right. This makes sense if one considers the formula for the Sharpe ratio and the minimal variation in the expected returns. Now we consider the construction of the portfolios. Only about half of the shares available were included in the portfolios. This is typical of unconstrained market portfolios using quadratic programming. In Figure 4.18 and 4.19 we display portfolio members for SIM and ISIM cases. In all the portfolios remgro is the most prominent share. The reason for this is that remgro has the highest Sharpe ratio (0.453) of the available shares. Interestingly enough it does not have the highest historic return. The share Jonnie has the highest return of 0.0190 per month compared with

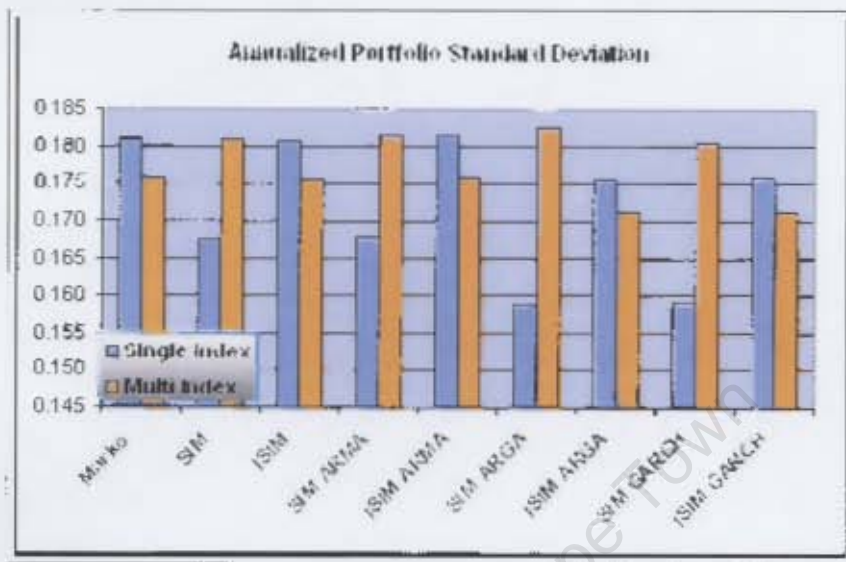


Fig. 4.17. The Annualized Standard Deviation.

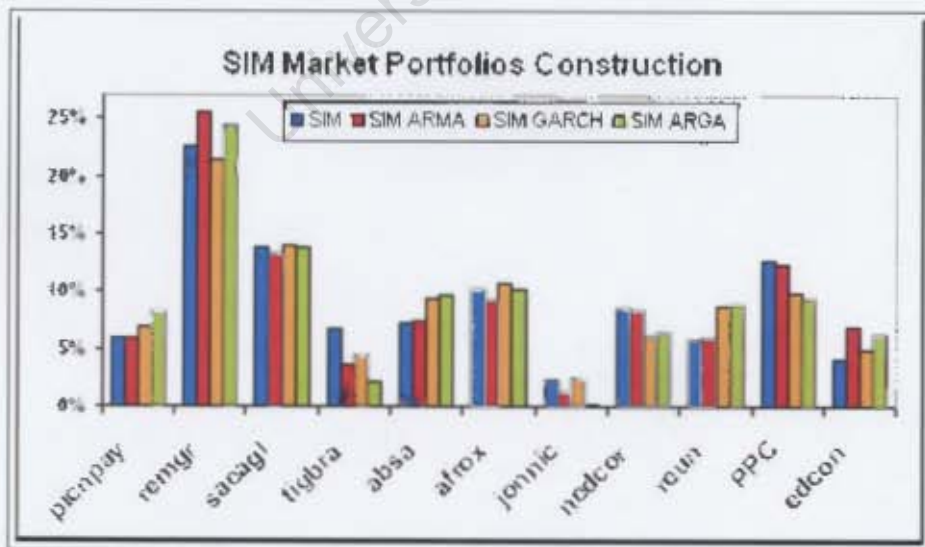


Fig. 4.18. The SIM Portfolios

remgo 0.0175. Jonnic is included in the portfolio, but as one of the minor components. From Figure 4.19 and Figure 4.20 we see that the different models produce quite similar

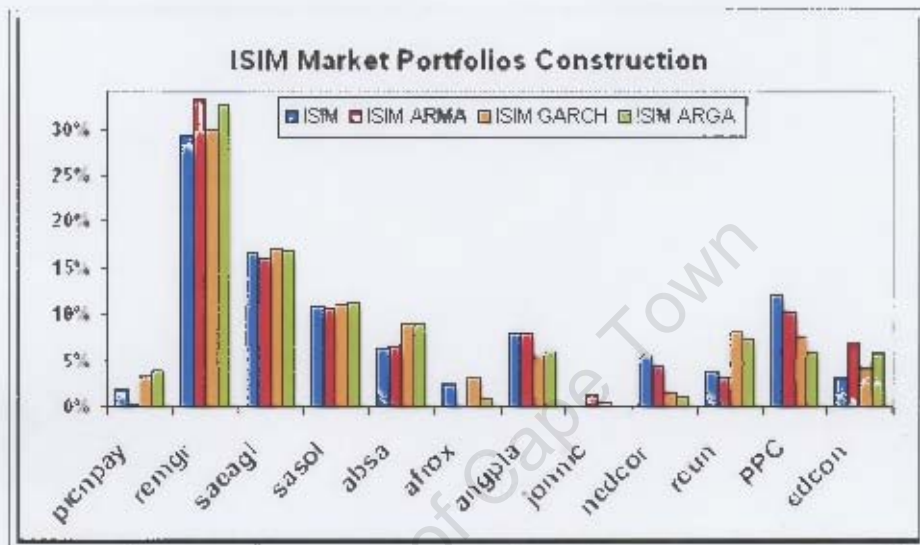


Fig. 4.19. The ISIM Portfolios

portfolios if the full 17-year data set and unconstrained optimization is used. The variations between the portfolios, though small, might have a significant effect on the resulting performance of the portfolios. This statement and the relative performance of the models will be tested in Chapter 7.

4.5 Conclusion

Using ARMA and GARCH did improve beta estimation and R^2 in the regression models. The effect on the residuals varied. In general, the addition of ARMA components reduced standard error and residual autocorrelation. The addition of the GARCH(1,1) extensions,

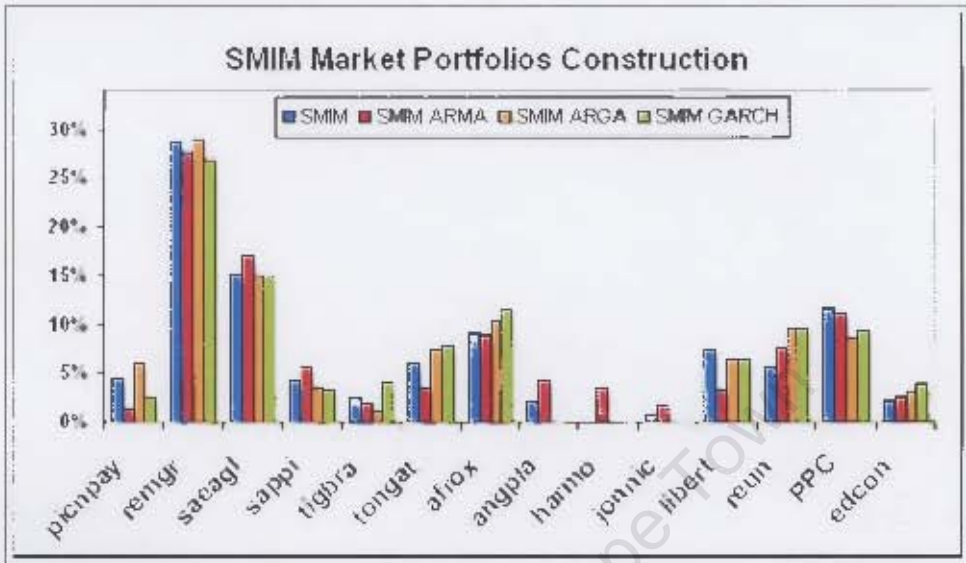


Fig. 4.20. The SMIM Portfolios



Fig. 4.21. The ISMIM Portfolios

on the other hand, seem to have a slight negative effect on the resulting the residuals . These effects where less pronounced in the multi index case compared with the single index.

As shown in Figures 4.11 to 4.14 the ARGA (ARMA and GARCH(1,1)) extensions move the frontiers to the left, discounting the risk in the portfolio. This effect is most pronounced in the SIM type models in Figure 4.11. In the SMIM models the effects are still visible but to a lesser extent and we find a general tighter grouping of the frontiers closer to the Markowitz frontier. For the ISIM and ISMIM models, the concentration closer to the Markowitz frontier is even more pronounced. The ISIM model is by its nature closer to the Markowitz, as is explained in Chapter 2. The multi index models are less sensitive to residual modelling. The reason for this is that the additional indices incorporate more of the variances in the share prices than can be incorporated by just the one index, resulting in smaller residuals.

The resulting models did produce different portfolios, but the biggest changes were produced by using improved Sharpe models or multi index models rather than an ARMA or GARCH extension. It is also noted that the SIM type models are less concentrated than the other portfolios.

In conclusion, the ARMA and GARCH extensions of the index models have a minimal effect on the resulting portfolio construction compared to effect of other model assumptions.

Chapter 5

Bounded Influence Regression

5.1 Introduction

Managing the effect of outliers in statistical data is an extensive and ongoing research area in Statistics. In financial research there is an ongoing debate on the management of outliers in modelling. In this chapter we follow one approach called Bounded Influence Regression. This comprises of down weighting excessive log returns, in both the independent and dependant variables, before fitting a regression model. In the next three sections we introduce three weighting functions. We then use two empirical studies to investigate the effect of these weighting functions on the efficient frontiers and the resulting market portfolios of the models.

5.2 The Flat Weights Function

One possible elementary weighting function is to set the high and low extreme return values to zero. Formally, if $x_{(1)}, \dots, x_{(n)}$ are the n ordered of observations from x_1, \dots, x_n we can choose some real number $\tau \in [0, 0.5)$ and let $L = [\tau n] + 1$ be a τ -lower index and $U = n + 1 - L$ the τ -upper index of the ordered series, where $[a]$ is the integer part of real

a. Define r_1, \dots, r_n as the ordered ranks of x_1, \dots, x_n . Then the weights will simply be:

$$\begin{aligned} W_i &= 0, & r_i < L, \\ &= 0, & r_i > U, \\ &= 1, & L \leq r_i \leq U \end{aligned}$$

for $i = 1, \dots, n$.

For example in Figure 5.22 the flat weights function is applied to the log returns of the JSE Overall Index with a τ value of 0.15. Note that the top and bottom 15% points are set to zero. This is a rather crude way of dealing with fat tails and potential outliers in the returns

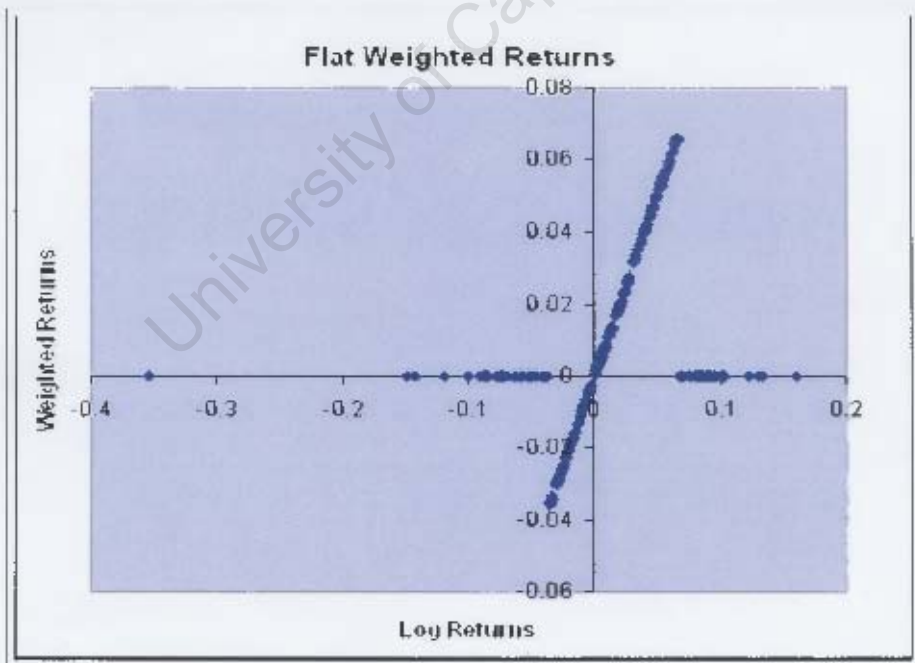


Fig. 5.22. A plot of the log returns vs. the weighted log returns for the so called flat weighting function.

and can be improved upon. We now introduce two more sophisticated weighting functions.

5.3 Mallows Weight Function

Here we will follow a similar approach to de Jong, de Wet, and Welsh (1988) in which they used the Mallows weights in bounded influence regression. The following weighting function was introduced by Mallows (1973). We follow the set up as in Section 5.2. but the weights associated with the x_i are now:

$$\begin{aligned} W_i &= \frac{x(L) - x(U)}{D_i}, \quad r_i < L, \\ &= \frac{x(U) - x(L)}{D_i}, \quad r_i > U, \end{aligned} \quad (5.59)$$

$$= 1, \quad L \leq r_i \leq U$$

$$\text{for } i = 1, \dots, n, \quad (5.60)$$

where $D_i = 2x_i - x(U) - x(L)$. As de Jong et al.(1988) note, the weights are chosen so that the outliers in the independent variable space are given less weight according to their distance from the center of that space. de Jong et al (1988) chose $\tau = 0.15$ which they showed to work well in practice. The profile of the Mallows weighting function on the JSE Overall Index is illustrated in Figure 5.23.

Immediately, one notices the non-symmetrical behaviour of the weighting function on extreme positive and negative returns. Also the returns higher than the order statistic U are weighted down lower than U , so that the resulting extreme positive returns are lower than the highest unweighted returns close to U .

Initially it was suspected that this asymmetry was caused by the data not being centered. To investigate this, the returns were mean-adjusted and the Mallows weights recalculated. The resulting weighted profile is shown in Figure 5.24. As can be seen in

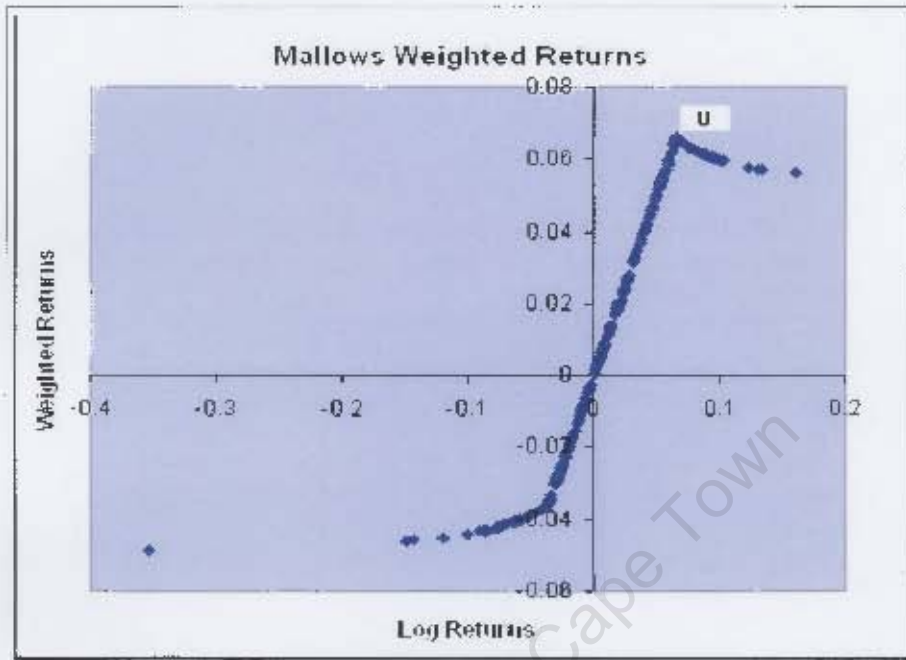


Fig. 5.23. A plot of the returns of the JSE Overall Index vs the Mallows' weighted returns. The Mallows weights profile is not symmetrical and does not increase monotonically.

that figure, the effect of using mean-adjusted returns gives a flatter, less smooth weight function. However the asymmetry, though less pronounced, still remains. The weights assigned to the positive extreme values decrease returns below the upper point U , with a sharp bend in the profile, causing a non-smooth profile in the returns. The reason why the resulting weighted returns have been reduced below the point U can be found by examining L and U order statistics for the JSE Overall Index and $\tau = 0.15$. Here we use the normal non-mean adjusted returns,

$\tau = 0.15$ $n = 223$	Value	r_i	Avg Returns below L /above U	Avg Weights below L /above U	Avg Weighted Return below L /above U
L	-0.047	34	-0.0887	0.611	-0.0542
U	0.053	190	0.0768	0.687	0.0528

Table 5.1 Statistics related to L and U

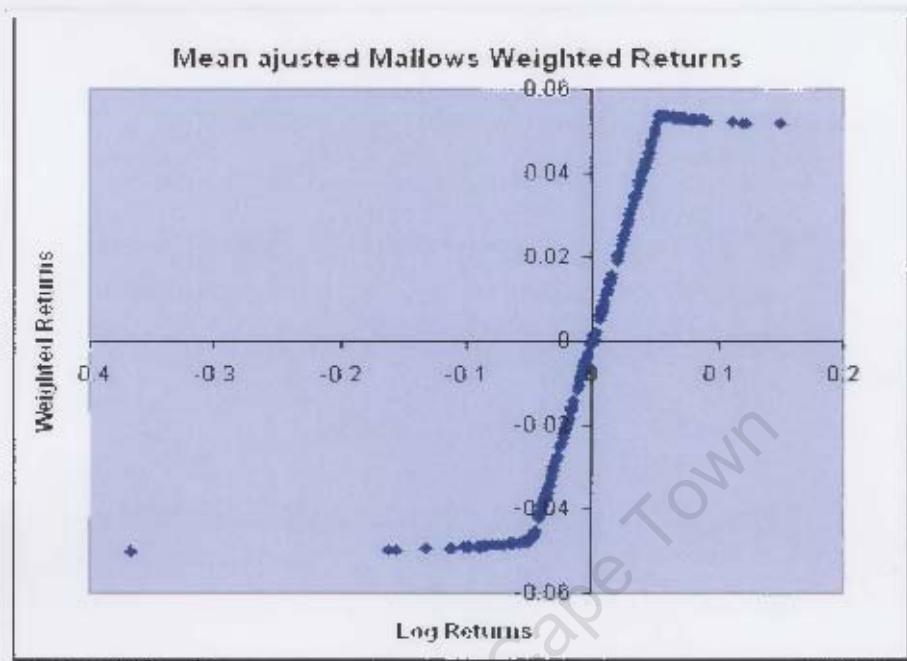


Fig. 5.24. Returns vs. Mean adjusted Mallows weighted returns for the JSE Overall Index. A flatter less smooth profile is observed.

From Table 5.1 the top 15% point U lies further away from zero than the bottom 15% point L . However, the average value of the returns above the point U is 11% lower than that of the average negative values below the point L on an absolute basis. So, on average, the extreme positive returns of 0.0768 are lower than the negative extreme returns of -0.0887 on an absolute basis. If we consider the frequency distribution of the returns in Figure 5.25 we see that the median of the returns, 0.0177, is noticeable higher than the mean of 0.0119. This is consistent with the skewness of the returns (-1.258) being negative, implying a long left tail (see Figure 5.25). The extreme negative values are pulling the mean to the left of the median. This explains why the average negative extreme returns are greater than the average positive extreme values on an absolute basis.

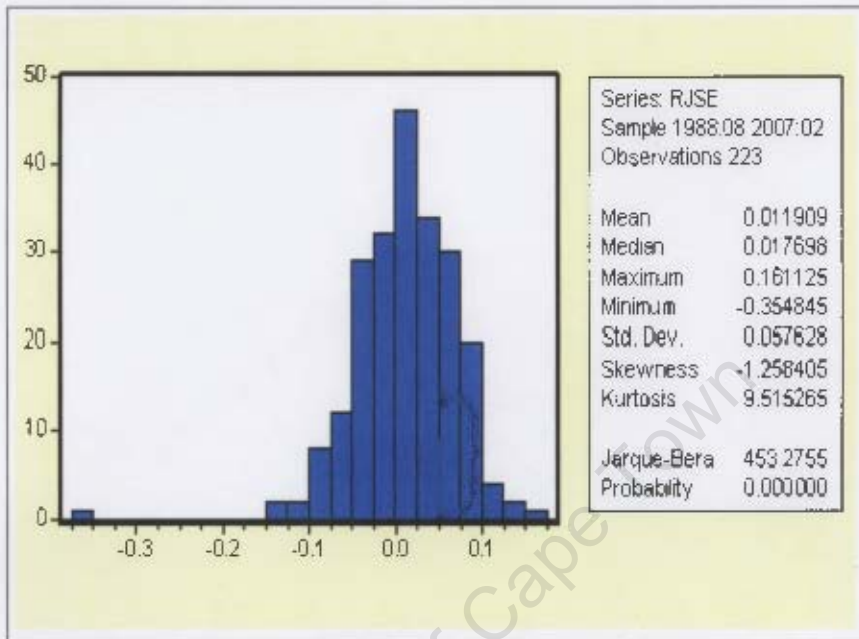


Fig. 5.25. Frequency distribution of the (non-mean adjusted) returns of the JSE Overall Index

The asymmetry between the weighted positive and weighted negative values is because the product of the average weights above U and the average returns above U , (called the Avg Weighted Return in Table 5.1) is just lower than U itself. On the negative side, the product of the average weights below L and returns below L is more negative than L itself. This explains why the resulting profile is increasing on the negative side below L and decreasing on the positive side above U . This behaviour of the function was found for all equities and index returns tested and seems to be a result of the general negative skewness of the return series investigated.

The sample skewness of the JSE Overall Index returns was -1.258. Interestingly, if the skewness of the returns inside the L - U boundaries is calculated, a positive value of

0.453 is found. So most of the negative skewness can be explained by the extreme values. In fact there are more small-to-medium up days than there are small-to-medium down days in equity market returns. This is consistent with saying that most of the big extreme values are negative in equity markets.

An interesting observation on skewness can be found if we analyze the negative and positive returns separately as is shown in Figure 5.26 and Figure 5.27. We use the mean adjusted returns.

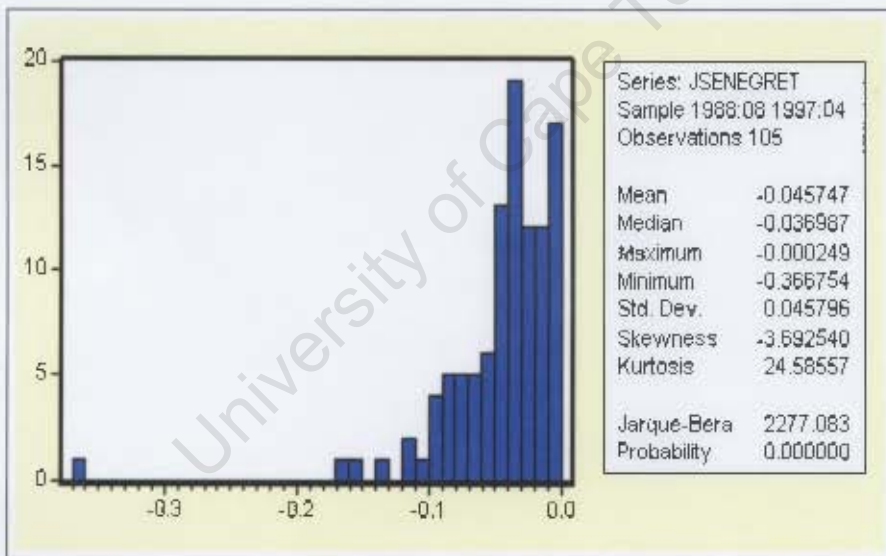


Fig. 5.26. Histogram of the mean adjusted negative returns

The skewness, is now around the centre of the new series and has a value of -3.692 for the negative and 0.878 for the positive returns. The kurtosis is quite different between the two series being 24.59 for the negative series compared to 3.71 of the positive series. This illustrates the difference between the positive and negative return series.

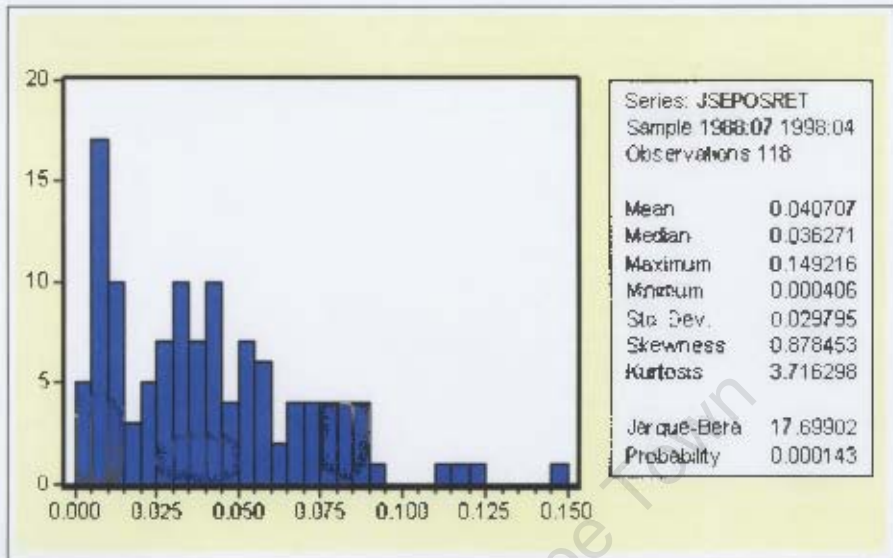


Fig. 5.27. Histogram of the mean adjusted positive returns

5.4 The ArcTan weighting function

After the rather lengthy analysis of the working of the Mallows weight function on equity return data, we would like to introduce an alternative method of weighting shares and index returns that avoids the main problems associated with the Mallows weights used on financial data:

- The asymmetry in the weighting functions.
- The non smoothness of the resulting weighted - unweighted profile.
- The reductions of the extreme positive returns below that of order statistic U .

The features of the Mallows weighting algorithm that we wanted to maintain are as follows:

- A function based on order statistics.

- A function that is independent of any maximum or minimum extreme point but rather uses the top/bottom τ percentage.
- A function that is parameter driven.

In this dissertation a weighting function was designed, using an arctan profile to meet these criteria. The function is called the ArcTan weighting function and is designed in this study based on the arctan function.

We denote the n observations of a variable as x_1, \dots, x_n , the ordered n observation as $x_{(1)}, \dots, x_{(n)}$, and define r_1, \dots, r_n as the vector of ranks from small to large of x, \dots, x_n . Let $L = [\tau n] + 1$ be a τ -lower index and $U = n + 1 - L$ the τ -upper index of the ordered series. Additionally we introduce the parameter λ , a stretching factor above the U order statistic (typically λ is between 10% and 200%). The highest return value never exceeds $(1 + \lambda)U$ and the lower value is never less than $(1 + \lambda)L$. The ArcTan weighting function weight associated with the observation x_i is then:

$$\begin{aligned}
 W_i &= \frac{x_{(L)}}{x_i} \left[1 + \lambda \frac{\pi}{2} \arctan\left(\frac{1}{\lambda} \frac{x_i - x_{(L)}}{x_{(L)}}\right) \right], \quad r_i < L, \\
 W_i &= \frac{x_{(U)}}{x_i} \left[1 + \lambda \frac{\pi}{2} \arctan\left(\frac{1}{\lambda} \frac{x_i - x_{(U)}}{x_{(U)}}\right) \right], \quad r_i > U, \\
 &= 1, \quad L \leq r_i \leq U
 \end{aligned} \tag{5.61}$$

for $i = 1, \dots, n$. The reciprocal of the stretching factor λ is used in the argument of the arctan function to scale up the relative returns before arctan is applied. This function is parameterized by τ and λ . The parameter τ determines the upper and lower $\tau\%$ points to be weighted and λ how far they should stretch above (below) the point U (L) in terms of U (L). The profile of the ArcTan function, with $\tau = 0.15$ and $\lambda = 0.3$ again applied to the

JSE Overall index, is shown in Figure 5.28. The function is guaranteed to be monotone,

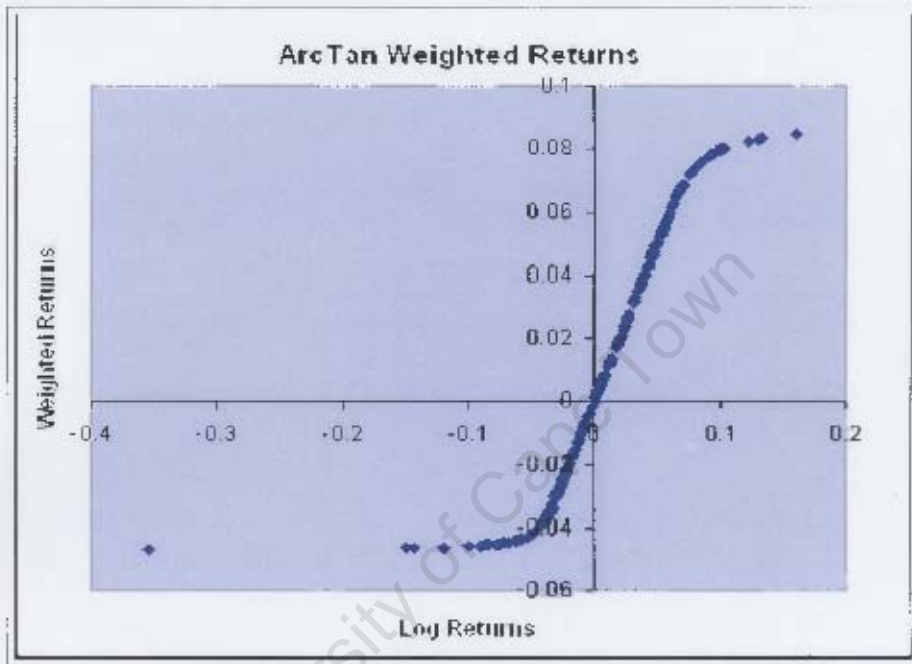


Fig. 5.28. A plot of the ArcTan Weighting function applied to the JSE Overall Index.

increasing and smooth for all returns series. To illustrate the versatility of the parameterized ArcTan function we display the profile for λ varying from 10% to 200% for $\tau = 0.15$ in Figure 5.29. A λ value of 10% allows the extreme returns to extend only 10% above (below) U (L). A λ value of 30% will limit the extreme value to 30% above (below) U (L). The higher the stretch variable the less the extreme returns are reduced and the straighter the return profile. The VBA code for the ArcTan weighting function is given in Appendix C.

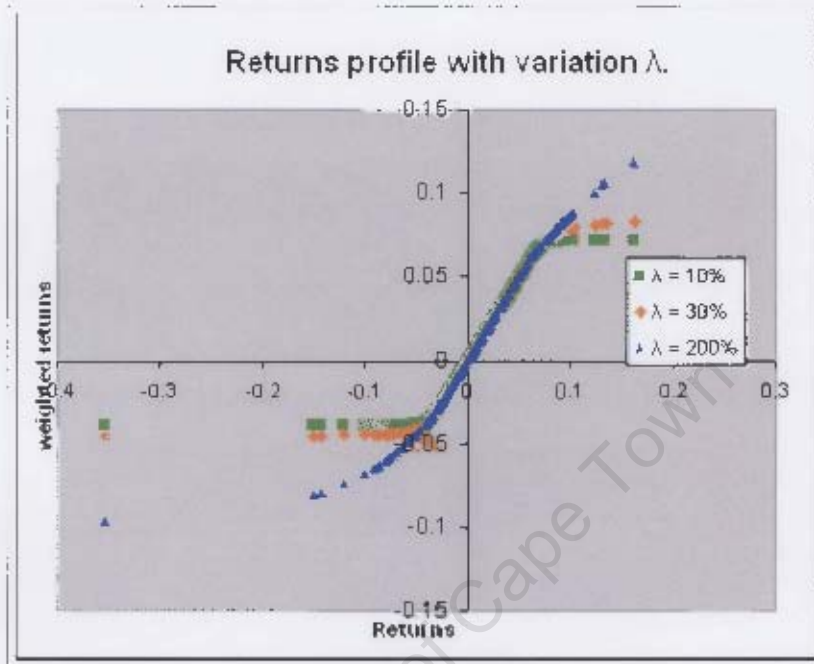


Fig. 5.29. The ArcTan function Return profile for different λ .

5.5 Empirical Study: The ArcTan weighting functions

In this section we investigate the effect of changing λ , the stretching factor in the ArcTan weighting function, on the mean-variance efficient frontier while keeping all other parameters constant.

5.5.1 The Data

The standard data set was used as in the previous chapters. It is a set of twenty shares and the JSE Over All Index. The data is monthly time intervals spanning from September 1988 to February 2007.

5.5.2 Study Objectives

The objective of this study is to investigate the influence of changing the stretching parameter λ in the ArcTan weighting function on the resulting efficient frontiers in the mean variance space using the Improved Sharpe Index Model (ISIM).

5.5.3 Methodology

The ArcTan function as described in Equation (5.61) was applied to the set of 20 share log returns and the JSE Overall Index. The τ - parameter for the ArcTan function was kept at 15% so only the top 15% and bottom 15% extreme returns are weighted. This was repeated for the λ parameter set to 0.1, 0.3 and 2.0. The weighting functions were implemented with VBA in Excel and the resulting weights were applied, in EVIEW3, to the 21 time series and the historic mean and ISIM covariance were calculated. The mean-variance optimisation was done using a quadratic programming algorithm in MATLAB.

5.5.4 Primary Findings

The effect of λ on the resulting mean-variance frontiers is displayed in Figure 5.30. First, it is observed that decreasing λ causes the efficient frontiers to shift from right to left, reducing the apparent risk in the portfolio. As λ decreases the individual variances of the weighted shares are decreased. This is consistent with the notion that the more one downweights the shares the less risky they appear to be.

Secondly, it was observed, as is displayed in Table 5.2, that the expected return does not change significantly when we use different weighting functions. This can be expected

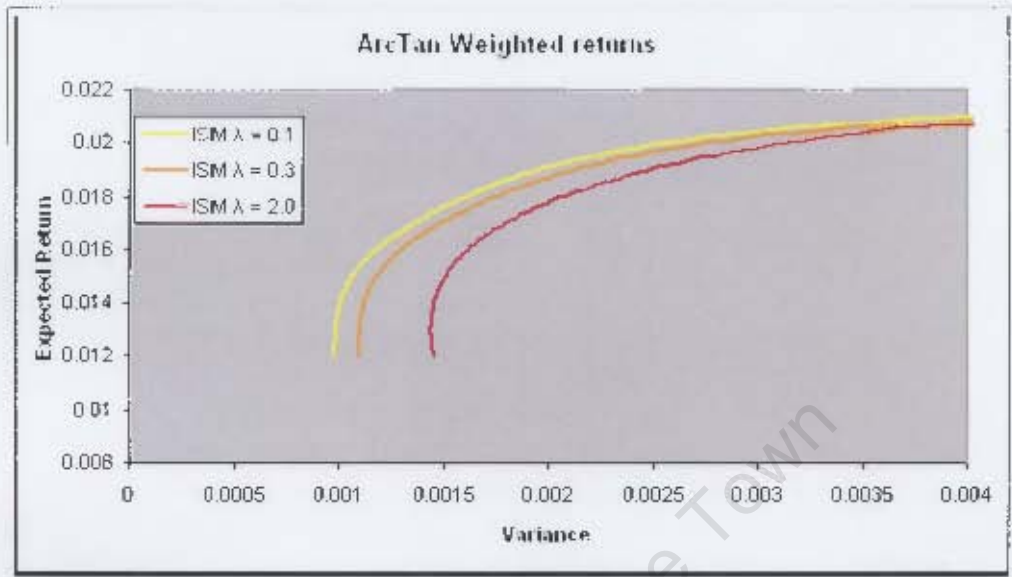


Fig. 5.30. The efficient frontiers for different λ in the ArcTan weighting functions. From left to right $\lambda = 0.1$, $\lambda = 0.3$ and $\lambda = 2$.

intuitively, based on the fact that we reduce the extreme values on both the up and down side of the returns. Therefore the frontiers did not move vertically.

λ	Mean of the Mean Returns
0.1	0.01330
0.3	0.01345
2.0	0.01343

Table 5.2 Mean Returns

5.6 Empirical Study: Comparing the weighting functions

5.6.1 The Data

The standard data set of twenty shares and only one index, the JSE Overall Index, were used. The data are based on monthly time intervals spanning from September 1988 to Feb-

ruary 2007. The log returns were weighted prior to regression using the different weighting functions.

5.6.2 Study Objectives

The objective of this study is to investigate the impact of using different weighting functions on the 20 share returns on the mean-variance profile and the resulting market portfolios created. We also investigate what effect weighting the data will have on the resulting market portfolio compared to using non-weighted share returns data.

In addition, we will compare the impact of different weighting functions on the resulting market portfolio, and, in particular, we will investigate whether using the GARCH extension has the same impact on weighted efficient frontiers as it had on unweighted efficient frontiers in Chapter 4.

5.6.3 Methodology

The original return data for the twenty shares was transformed using the Flat, Mallows, and ArcTan weighting functions. For all three of these functions the τ parameter was set to 15% and for the ArcTan function we used a stretching parameter λ of 0.3. After the data had been transformed, the regression models were built and the covariance estimate was obtained using the techniques described in the previous chapters. The relevant estimated returns and covariance matrix were used to create efficient frontiers using the quadratic programming algorithm in MATLAB. The results were compared using VBA driven plots in Excel.

5.6.4 Primary Findings and Results

The following models are compared in this empirical study:

- **Marko**: Mallows weighted historic covariance and mean are used to model the efficient frontiers.
- **WSIM**: Mallows weighted Sharpe Index Model.
- **WISIM**: Mallows weighted Improved Sharpe Index Model.
- **WSIM GARCH O**: Mallows weighted Sharpe Index Model with a GARCH(1,1) extension.
- **WISIM GARCH O**: Mallows weighted Improved Sharpe Index Model with a GARCH(1,1) extension.
- **ISIM**: Unweighted Improved Sharpe Index Model.
- **FLATISIM**: Flat weighted Improved Sharpe Index Model.
- **ArcTanISIM**: ArcTan weighted Improved Sharpe Index Model.
- **MallowsISIM**: Mallows weighted Improved Sharpe Index Model (same as WSIM).

Comparison of the Efficient Frontiers

First we look at the mean-variance response of the Mallows weighted portfolio and some of its extensions as discussed in the previous chapters. What we find initially is the same pattern as observed in previous chapters. The SIM frontiers are to the left of

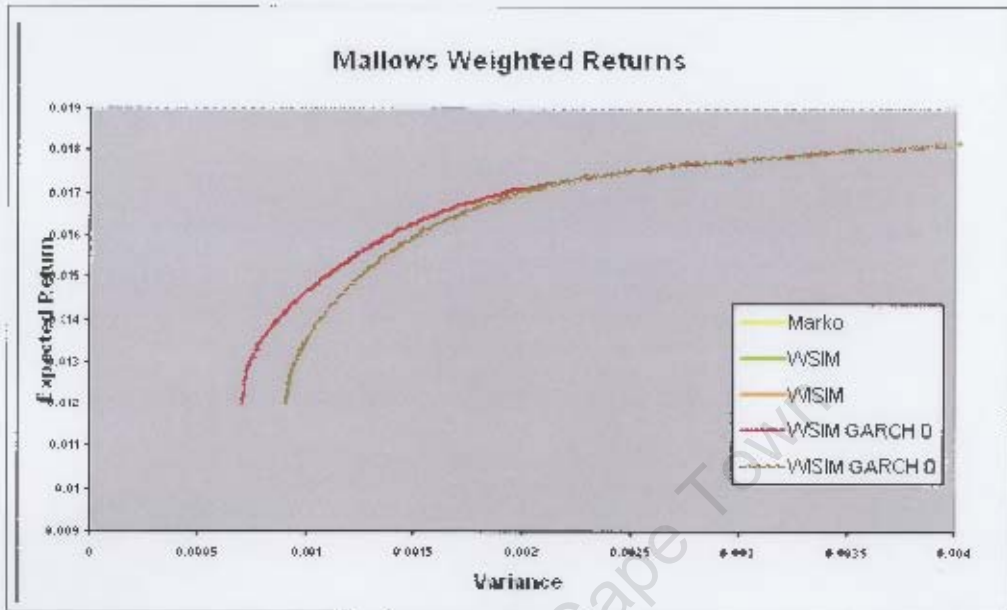


Fig. 5.31. This figure has five efficient frontiers in two groups. The top group contains the Weighted Sharpe Index Model (WSIM) and the WSIM GARCH O Only extension. The second group contains Mallows weighted Markowitz (the historic covariance), the Improved Sharpe Index model with Mallows weights applied (WISIM) and WISIM GARCH O.

the ISIM frontiers. The reason for this is that the Sharpe index model underestimates the non-systematic risk inherent in the portfolio which is incorporated in the ISIM models.

Also, the WISIM again has the same mean-variance response as the weighted Markowitz covariance structure. These two frontiers are on top of each other, together with the WSIM GARCH O. The argument here again is that all the available residual covariance has been included in the calculation of the covariance matrix and therefore the same efficient frontier is produced.

The interesting point of these graphs is the non-effect of the GARCH(1,1) extensions of the regression models. For both WSIM, and WISIM, the GARCH(1,1) extension had no real effect on the mean-variance response of the portfolio. This is quite different from the

response to normal unweighted returns. As was seen in the previous chapter, this had the effect of moving the efficient frontiers to the left for unweighted return data. By weighting the returns we reduce or eliminate the effect of GARCH(1,1) modelling on the efficient frontiers of the data.

Next we examine the efficient frontiers produced by the different weighting functions.

In Figure 5.32 we show the three mean-variance responses of the three weighted ISIM

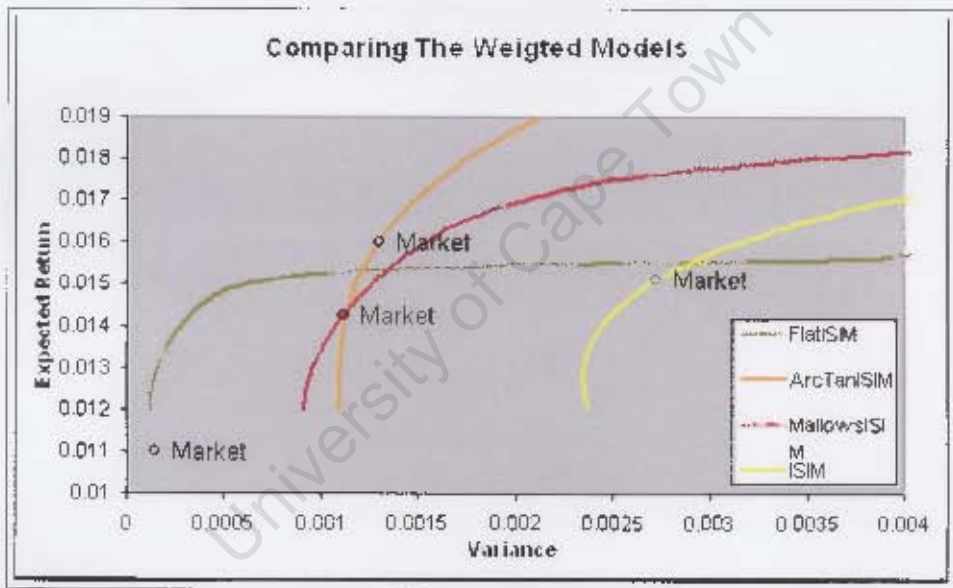


Fig. 5.32. The three mean variance response of the three weighted ISIM models: From left to right: Flat ISIM, Mallows ISIM, ArcTan ISIM, and un-weighted ISIM.

models: from left to right: FlatISIM, MallowsISIM, ArcTanISIM, and un-weighted ISIM. First we observe that all the weighted return models understate the risk compared with the original ISIM model. This can be expected, since there are less extreme events in the weighted returns. The flat efficient frontier is significantly 'lower' than the others. This is also apparent from the market portfolios of the different models. This could be explained

by examining the mean of the means of the returns (the MoMoR). The MoMoR of the flat weights returns is 0.00974 and is notably lower than 0.01345 or 0.01322, the MoMoR for the ArcTan and un-weighted returns respectively.

Next we see that the Mallows and ArcTan efficient frontiers are close to each other in the region of the market portfolios. An ArcTan function with $\lambda = 0.3$ was again used. Both frontiers are positioned as more attractive in terms of the risk and returns profiles compared to the ISIM model. As was shown earlier, the efficient frontier of the ArcTan function can be moved left and right by adjusting the λ parameter. Next we consider the interesting question of how different the resulting market portfolios of these models are and compare them to the market portfolio of the un-weighted returns.

Comparison of the Market Portfolios

The positions of the four market portfolios corresponding to the four weighted models are shown in the Figure 5.32. The market portfolio was calculated for an interest rate of 9% throughout. Table 5.3 shows the annualized Sharpe ratio for the portfolios. As can be seen from Table 5.3 and Figure 5.32, the portfolio with the highest Sharpe ratio is the FlatISIM follows by MallowsISIM, ArcTanISIM and then the un-weighted ISIM. This is consistent with the positioning of the mean-variance frontiers for the different models. If one examines the standard deviations and expected returns in Table 5.3 for the ArcTanISIM and MallowsISIM, it is possible to see why the Sharpe ratio is higher for the ArcTanISIM. As was discussed earlier, the Mallows weights overly discriminate against large positive returns. The ArcTan weighting function was designed not to do this, but rather penalize large

positive returns less and to produce a more smooth and symmetrical profile. Therefore the expected return for the ArcTanISIM portfolio at 21.03% is higher than the MallowISIM portfolio at 18.52%. Also, because the standard deviations differ only by 0.9%, the Sharpe Ratio for the ArcTanISIM is higher than the MallowsISIM.

	ISIM	FlatISIM	MallowISIM	ArcTanISIM
Annualized Sharpe Ratio	0.50	1.00	0.70	0.82
Annualized Standard Deviation	18.1%	4.2%	11.6%	12.5%
Annualized Expected Return	19.72%	14.03%	18.52%	21.03%

Table 5.3 Portfolio Measures

It is important to note that the portfolio performance figures calculated in this section are based on the weighted returns of the constituents of the portfolios and not the real un-weighted return data of the shares in the portfolio. In Chapter 7 we will construct portfolios and run them over time and recalculate performance measures based on real market returns. We now look at the resulting portfolios. The composition of our four portfolios is shown in Figure 5.33. As can be seen from the figure, there are large changes in the portfolio construction for the different models. The normal un-weighted returns market portfolio has the highest concentration in shares. The more aggressive weighting is applied, the better the resulting portfolios balanced between assets. This can be seen by counting the portfolio weights above 5%.

Portfolio weights above 5%				
	ISIM	FlatISIM	MallowISIM	ArcTanISIM
Number of shares with more than 5% weights	7	11	9	9

The quadratic programming algorithm used to create these portfolios has no maximum holding constraints on the assets. It is therefore interesting to note that using the flat weights has a similar effect to setting the maximum holding in one asset to approximately

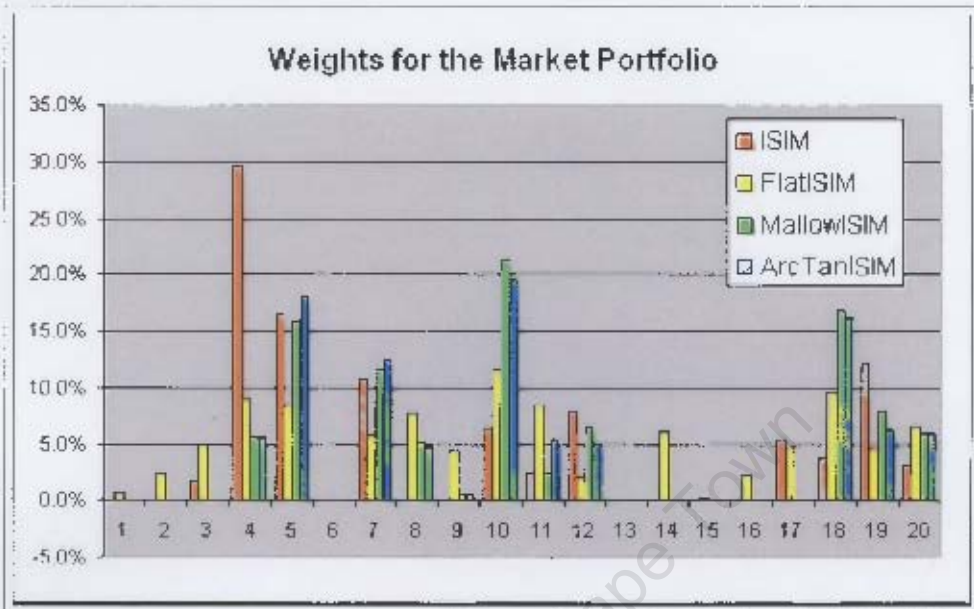


Fig. 5.33. The weights of the portfolio constituents for the four market portfolios under consideration.

10%. The question arises as to whether 10% constrained quadratic programming will produce the same or similar share portfolio than in this portfolio. This point is addressed in Figure 5.34. The un-weighted ISIM covariance was used but the quadratic programming was updated to use a 10% upper constraint. The resulting portfolio is shown next to the flat weighted portfolio with $\tau = 0.15$ in Figure 5.34. Every asset that is included in the capped ISIM10% portfolio is included in the FlatSIM portfolio and there are some major similarities between the two portfolios. The FlatSIM portfolio is, however, more diverse and includes 4 additional assets (1, 2, 9, and 16) not in the 10% capped portfolio.

It is therefore not clear as to whether the FlatSIM is a proxy for a capped portfolio. However it is interesting that each major investment above 6% that the FlatSIM model recommends is included in the 10% capped ISIM portfolio.



Fig. 5.34. FlatSIM market portfolio and the ISIM with a hard 10% upper limit in any asset.

5.7 Conclusion

In this chapter we investigated some existing methods of weighting returns and introduced a new method of weighting outliers using a parameterized extension of the arctan function. We showed that weighting had a significant effect on the mean-variance response of a portfolio of shares. Using the λ parameter of the Arc Tan function, the efficient frontier can be moved to the left or to the right.

We also showed that the use of different weighting functions has drastic effects on the resulting market portfolios. Also, it was shown that GARCH extensions had a minimal effect on the resulting efficient frontier for weighted returns. It was also shown that aggressive weighting, as in the FlatSIM, produces a more equally weighted portfolio.

Chapter 6

Principal Components Regression Models

In this Chapter we investigate the use of Principal Component Analysis (PCA) in creating alternative indices to be used in our Sharpe and improved Sharpe Multi Index Models (SMIM and ISMIM respectively). In the first two sections we introduce the PCA theory and the method of finding the eigenvalues and eigenvectors of the covariance matrix using Singular Value Decomposition (SVD). In the empirical study four sets of principal component (PC) indices are constructed using both the dependant and independent variables of the standard regression models (Aflleck-Graves, Money and Troskie, 1979). The model building performance and mean-variance response of these models are evaluated. The results demonstrate the importance of residual correlation in the relative positioning of the SMIM frontier in relation to the ISMIM. We also find a rare example of negative correlated residuals and show how this ‘moves’ the SMIM frontier to the right of the ISMIM frontier.

6.1 Theory of PCA

In this introduction we rely heavily on the class notes of Troskie (2000). Let the p -component random vector \mathbf{x} have $E(\mathbf{x}) = \mathbf{0}$ and $E(\mathbf{x}\mathbf{x}') = \Sigma$. Then there exists an orthogonal linear transformation

$$\mathbf{u} = \mathbf{B}'\mathbf{x} \tag{6.62}$$

such that the covariance matrix of \mathbf{u} is a diagonal matrix

$$\begin{aligned} E(\mathbf{u}\mathbf{u}') &= \Lambda \\ &= \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{pmatrix} \end{aligned}$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ are the eigenvalues of the covariance matrix Σ . The eigenvalues are the roots of the characteristic polynomial

$$|\Sigma - \lambda\mathbf{I}| = 0$$

and \mathbf{B} is the corresponding matrix of eigenvectors that satisfies:

$$\mathbf{B}'\mathbf{B} = \mathbf{I}, \quad \mathbf{B}'\Sigma\mathbf{B} = \Lambda.$$

The r^{th} column of \mathbf{B} , $\mathbf{B}^{(r)}$, satisfies

$$(\Sigma - \lambda_r\mathbf{I})\mathbf{B}^{(r)} = 0.$$

Since the covariance matrix Σ is always positive semi-definite, its eigenvalues are real and nonnegative. Let $(\lambda_1, \mathbf{B}^{(1)}), \dots, (\lambda_p, \mathbf{B}^{(p)})$ be the eigenvalue-eigenvector pairs of Σ , where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. We then have the following statistical result.

Result: The r^{th} principal component of \mathbf{x} is $\mathbf{u}_r = \mathbf{B}^{(r)'}\mathbf{x}$ and has maximum variance of all normalized linear combinations of \mathbf{x} uncorrelated with $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{r-1}$. Thus,

$$\text{Var}(\mathbf{u}_r) = \mathbf{B}^{(r)'}\Sigma\mathbf{B}^{(r)} = \lambda_r \quad r = 1, \dots, p$$

$$\text{Cov}(\mathbf{u}_r, \mathbf{u}_s) = \mathbf{B}^{(r)'}\Sigma\mathbf{B}^{(s)} = 0 \quad r \neq s.$$

In addition, we have that the total variance is

$$\begin{aligned} \sum_{i=1}^p \text{Var}(\mathbf{x}_i) &= \text{tr}(\boldsymbol{\Sigma}) = \text{tr}(\mathbf{B}\boldsymbol{\Lambda}\mathbf{B}') \\ &= \text{tr}(\boldsymbol{\Lambda}) = \sum_{i=1}^p \lambda_i = \sum_{i=1}^p \text{Var}(\mathbf{u}_i). \end{aligned} \quad (6.63)$$

As a result of Equation (6.63), we can calculate the portion of variance explained by a principal component. The proportions of variance are also cumulative. So, for example, the percentage of the total variation explained by the first four components is given by

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}{\sum_{i=1}^p \lambda_i} 100\%.$$

It is therefore possible to quantify the amount of variation explained by a reduced set of principal components.

6.2 Sample Principal Components

The results developed in the last section are only directly applicable if the covariance matrix of the return data is known. Even though the real (and future) covariances of return data are unknown, they can be estimated consistently by sample covariances. To answer questions regarding principal components, distributional assumptions of the observed returns have to be made. Here we follow the approach in Press (1972) and present the following theorem.

Theorem 1

Let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be $N(> p)$ observations from $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a matrix with p distinct eigenvalues. Then a set of maximum likelihood estimators of the distinct eigenval-

ues $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$ and eigenvectors $\mathbf{B}^{(1)}, \dots, \mathbf{B}^{(p)}$ are the roots

$$k_1 > k_2 > \dots > k_p$$

of

$$|\hat{\Sigma} - k\mathbf{I}| = 0$$

and the set of corresponding vectors $\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(p)}$ satisfying

$$(\hat{\Sigma} - k_r \mathbf{I})\mathbf{g}^{(r)} = 0$$

and

$$\mathbf{g}^{(r)'} \mathbf{g}^{(r)} = 1,$$

where $\hat{\Sigma}$ is the maximum likelihood estimators of Σ .

The theorem was originally proven by Girshick (1936) and later streamlined by Anderson (1958). The implication of the theorem is as follows: the sample eigenvalues and eigenvectors are maximum likelihood estimators of the sample eigenvalues and eigenvectors of Σ , under the normality assumption and provided that the eigenvalues are distinct.

Sample Computation.

Since $\mathbf{x}_1, \dots, \mathbf{x}_N$ are $N > p$ (sample) observations from $N(\boldsymbol{\mu}, \Sigma)$ the maximum likelihood estimator of Σ is

$$\begin{aligned} \hat{\Sigma} &= \frac{1}{N} \sum_{\alpha=1}^N (\mathbf{x}_\alpha - \bar{\mathbf{x}})(\mathbf{x}_\alpha - \bar{\mathbf{x}})' \\ &= \frac{1}{N} \left\{ \sum_{\alpha=1}^N (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) \right\}, \\ &= [\hat{\sigma}_{ij}], \quad \text{for } i, j = 1, \dots, p. \end{aligned}$$

where $\bar{\mathbf{x}}$ is the sample vector mean.

The unbiased estimator is then

$$\mathbf{S} = \frac{N}{N-1} \hat{\Sigma}.$$

The latter is preferred and is used in the empirical study. From Theorem 1, if \mathbf{G} is the matrix of eigenvectors, then:

$$\mathbf{G}'\hat{\Sigma}\mathbf{G} = \mathbf{K} \text{ or} \tag{6.64}$$

$$\hat{\Sigma} = \mathbf{G}\mathbf{K}\mathbf{G}' \text{ where}$$

$$\mathbf{G}'\mathbf{G} = \mathbf{I},$$

that is \mathbf{G} is orthogonal, and

$$\mathbf{K} = \begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & k_p \end{pmatrix}$$

where k_i are the sample maximum likelihood estimates of the eigenvalues.

6.3 Empirical Study: Regression with Principal Component Indices.

In this empirical study we use principal components (PC) to construct indices to be used in multi index regression models. Four sets of principal components are considered. We use the original set of ten indices in Chapter 3 to construct ten principal components. The first model considers all of the ten PC's (called PCA10). The second model uses the first four PC's (called PCA4). The third considers the four PC that produced the best fitting regression model (called PCA4fit). These four PC's were selected using a variable selection procedure based on the t-statistics of the PC's beta value. The fourth model is based on four PC's of the dependant variables (called PCA4shares)

6.3.1 The Data

The standard set of 20 un-weighted share returns and the full set of 10 un-weighted index returns, as described in Chapter 3, were used. The full time span of the data from 1988 to 2007 was incorporated.

6.3.2 Study Objectives

The objective of this study is to establish the effect of using different PC indices for both the SMIM and ISMIM models. We will compare the quality of the model fit based on adjusted R^2 and residual standard errors. We will then consider the impact on the risk-return response of these models, comparing them to the standard multi index models of Chapter 3.

6.3.3 Methodology

As mentioned above, four models were considered. The PCA10, PCA4 and PCA4fit are based on the 10 indices introduced in Chapter 3. The PCA4shares model was based on the first four principal components of the share returns. The eigenvector and eigenvalues of the covariance matrix were calculated using the SVD on the mean adjusted and $\frac{1}{\sqrt{n-1}}$ weighted returns, as described in Appendix D. The unbiased estimate of the covariance matrix,

$$S = \frac{N}{N-1} \hat{\Sigma}$$

is preferred and is used in this empirical study.

The product of the eigenvectors and the return series give the orthogonal set of principal components (PC) that are used as indices in the SMIM models. Using the eigenvalues, it is possible to quantify the percentage of variance explained by the PC included in each of the models. In the PCA10 model 100% of the variance of the indices is naturally explained by the 10 PCs. For the PCA4 model 80.0% of the variance is explained by the first four PC's. PCA4 does not produce the best fitting model and it was decided to select 4 PCs based on a variable selection procedure (PCA4fit).

For the best PCA4fit model we decided to use the four 'best' fitting principal components to the regression models of the 20 shares. Two methods were used to decide on the four best fitting components. First, the 20 shares were regressed against the complete set of 10 PC and the t-statistics for each coefficient was recorded.

In the first method the absolute values of the t-statistic for each coefficient were summed over the 20 shares and the four principal components with the largest total absolute t-statistic were selected for use as indices in the factor models.

In the second method, the number of significant coefficients, based on absolute t-statistics greater than 2, were compared. The four with the highest score were selected. Both methods resulted in the same four principal components (PC1, PC5, PC7, PC8), where PC1 is the first principal component and PC10 the last. This set of indices only accounted for 51% of the variance.

The first four principal components were used to create the PCA4share models and contained 59.5% of variance in the 20 shares.

6.3.4 Primary Findings

Model Building Results

In Table 6.1 we compare the four SMIM with PC indices, termed PC models, with each other and the two non PC models, the full index model using 10 indices, called SMIM full, and the standard SMIM4 using the four selected indices of Chapter 3. The table displays the average R^2 , average Schwarz Criterion, and average Standard Error taken over all twenty shares used in the portfolio. It also shows the variance explained by the PC included in the model as a percentage of total variance in that set of indices. First we note that the full PCA10 and full SMIM produce identical model building results. This is, of course, expected based on the fact that the PCA10 is just an orthogonal linear transformation of the 10 standard indices. The two sets of indices represent exactly the same information.

Average	PCA10	PCA4	PCA4fit	PCA4shares	SMIM Full	SMIM 4
Adjusted R^2	0.36	0.18	0.30	0.52	0.36	0.33
Schwarz criterion	-2.065	-1.940	2.08	2.63	2.065	-2.129
Residual SE	0.0796	0.0894	0.0834	0.0640	0.0796	0.0817
Variance Explained	100%	80%	51%	59.5%	100%	N/A

Table 6.1 Regression Model Statistics

Next, we observed that the PCA4 model has a very weak fit. It has an adjusted R^2 of 0.18 and a Schwarz criterion value greater than -2 . This is surprising if one considers that these four PCs explain 80% of the variation of the 10 indices. If one considers the PCA4fit with the best four fitting PC indices, the adjusted R^2 of 0.30 is still less than that of the standard four index model (SMIM4) of 0.33. The reduced PC index models do not improve the regression model fit compared to the original SMIM4 models.

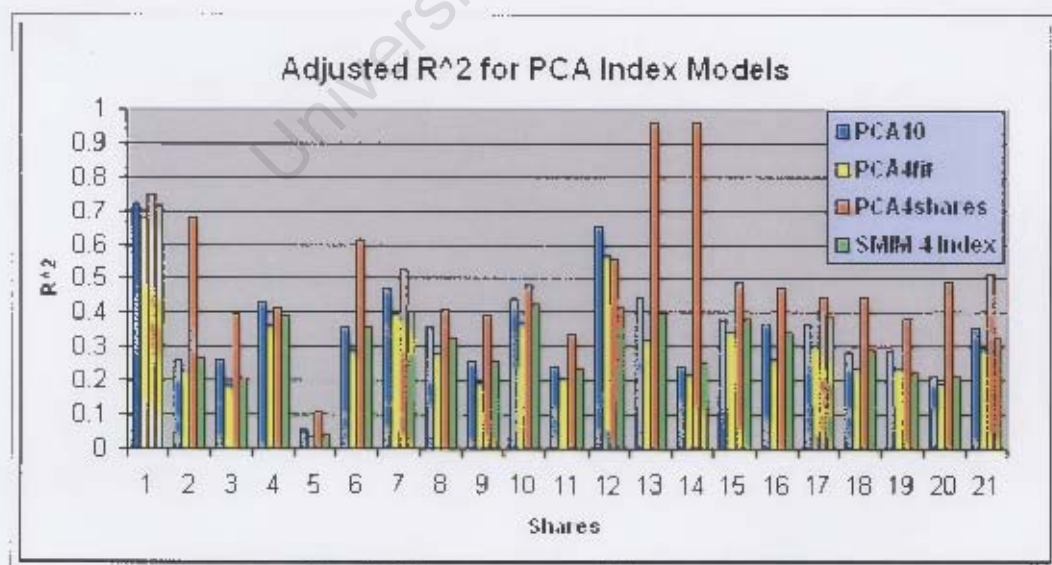


Fig. 6.35. The Adjusted R^2 for four models. From left to right: PCA10, PCA4fit, PCA4shares, and the standard SMIM with for indexes as used in Chapter 3.

Using the PC from the dependant variables has a major impact on the model fit. In Figure 6.35 we display the individual adjusted R^2 of some of the models for each share. From left to right the models are: PCA10, PCA4fit, PCA4shares, and the standard SMIM (called SMIM 4 Index) with four indices included as was used in Chapter 3. Share 21 is again the average for the 20 shares. From Figure 6.35 we see that PCA4shares model has a superior model fit for all the shares except share 12, Anglo Platinum. It is also the only share for which the PCA4fit produced a better fitting model than the SMIM 4 Index model.

Anglo Platinum aside, it should be clear that, using independent variables for regression, a linear combination of the dependent variables, as is the case with the PC indices, will produce better fitting models compared to regression on distinct market indices.

Mean-Variance Response

We begin by tabulating the average residual correlation of the four models. This gives us a good idea of the positioning of the Sharpe Index Models compared to the Improved Sharpe Index Models of Hossain et al (2005b). The average residual correlation was calculated as follows:

If we let $\hat{\mathbf{E}}$ be the matrix of p combined residuals series after the regression,

$$\hat{\mathbf{E}} = \begin{pmatrix} \hat{e}_{11} & \hat{e}_{12} & \dots & \hat{e}_{1N} \\ \hat{e}_{21} & \hat{e}_{22} & \dots & \hat{e}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{e}_{p1} & \dots & \dots & \hat{e}_{pN} \end{pmatrix}$$

then $\hat{\mathbf{C}} = \text{corr}(\hat{\mathbf{E}})$ is the correlation between the residual series, that is

$$\hat{\mathbf{C}} = \begin{pmatrix} 1 & \hat{\rho}_{12} & \dots & \hat{\rho}_{1P} \\ \hat{\rho}_{21} & 1 & \dots & \hat{\rho}_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}_{P1} & \dots & \dots & 1 \end{pmatrix}.$$

We then define the average correlation $\bar{\rho}$ between the residuals as

$$\bar{\rho} = \frac{\sum_{i=1}^p \sum_{j=1}^p (\hat{\rho}_{ij} - p)}{p(p-1)}.$$

The covariances between the residuals are ignored in the Sharpe Index Models and included in the Improved Sharpe Models of Troskie and Hossain. In the ISMIM, the correlations are modelled and for this reason all the efficient frontiers are clustered together close to the efficient frontier of Markowitz based on the historic covariance. The average correlation is a good indication of the relative position of the efficient frontiers associated with the SIM compared to associated ISIM. We demonstrate this using the average residual correlation presented in Table 6.2 and Figures 6.36, 6.37 and 6.38.

SMIM PCA10	1.98%
SMIM PCA4	18.70%
SMIM PCA4fit	6.09%
SMIM PCA4shares	-4.66%
SMIM 4	2.02%
Table 6.2 Average Residual Correlation	

We begin by considering the mean-variance response for the PCA10 and PCA4 models. In Figure 6.36 the efficient frontiers of four models are displayed, from left to right, SMIM PCA4, SMIM PCA10 and then grouped together the Markowitz historic covariance, ISIM PCA4 and ISIM PCA10. The reason the SMIM PCA4 frontier is to the far left should now be clear. It was the worst fitting model, with low R^2 and big residual standard errors. It also produced the largest positive average correlation in the residuals of 18.7%, as can be seen in Table 6.2. It therefore discounted the non-systematic risk in the portfolio. This illustrates the weakness in the Sharpe Type Index model of ignoring the regression residual correlations.

Using the full set of 10 PC's we had better models and reduced the average correlation in the residuals to 1.98%, which is still slightly positive and the SMIM PCA10 is slightly to the left of the ISMIM PCA10. Next we look at the PCA4fit model and the PCA4shares

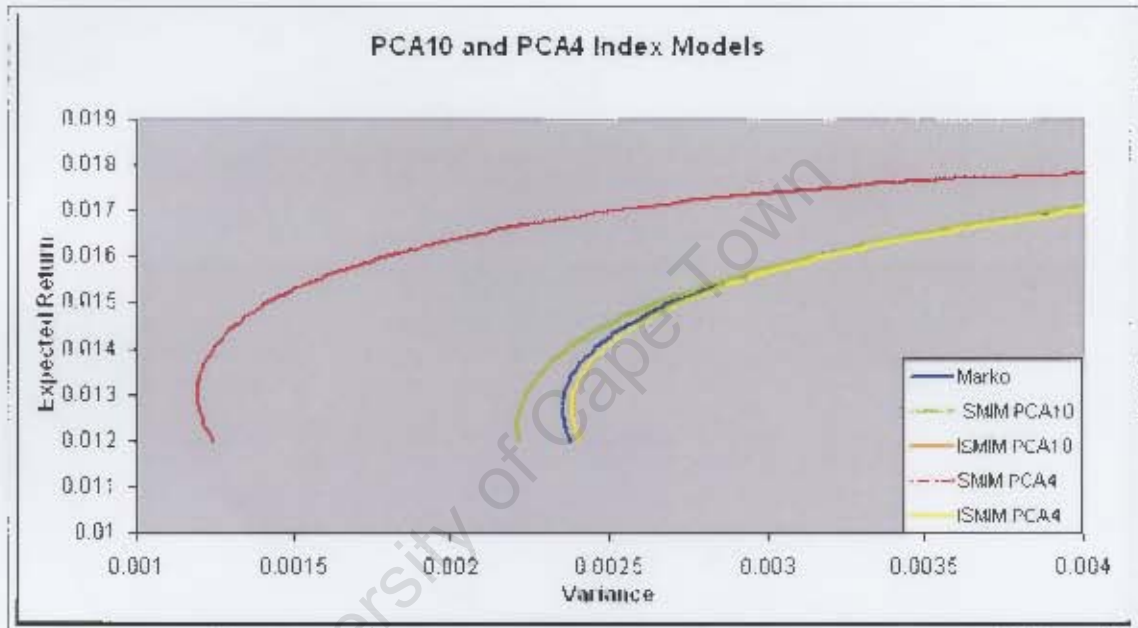


Fig. 6.36. The efficient frontiers of four models are displayed. From left to right, SMIM PCA4 SMIM PCA10 and group together: Markowitz historic covariance, ISIM PCA4 and ISIM PCA10.

models in Figure 6.38. The PCA4fit model is based on the best fitting PC from the set of indices. The PCAshares model uses PC indices created from the shares themselves, producing a very good fitting model. In Figure 6.38 we show, from left to right, SMIM PCA4fit model, ISMIM PCA4shares, ISMIM PCA4fit and SMIM PCA4shares. In Figure 6.38 the SMIM PCA4fit is furthest to the left. In actual fact the frontier is between the SMIM PCA4 and SMIM PCA10 from Figure 6.36. This is consistent with the left shift of

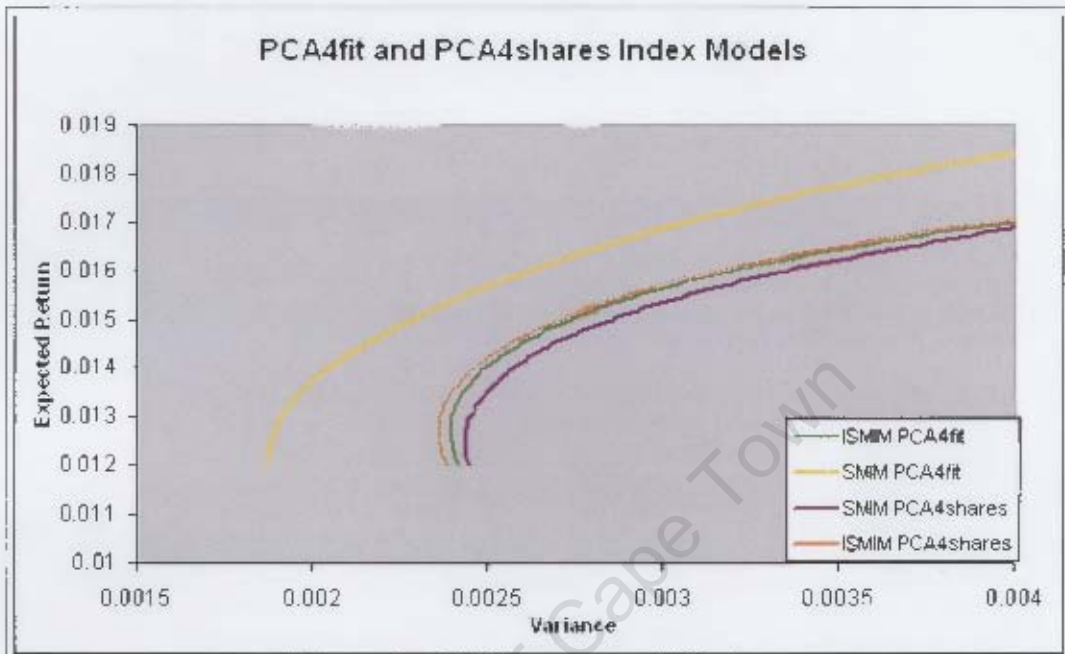


Fig. 6.37. From left to right SMIM PCA4fit model; ISMIM PCA4shares; ISMIM PCA4fit; and SMIM PCA4shares.

the SMIM for a positive average residual correlation of 6.09%. The most interesting feature of Figure 6.37 is the effect of the negative average residual correlation of the PCA4shares model. This model provides an exceptionally good fit to the shares and produces an average residual correlation of -4.66% . Furthermore, it is an example of negative average residual correlation producing a rightward shift of a SMIM compared with the ISMIM.

Finally we compare the PC approach with the original index models. In Figure 6.38 we use the PCA10 models to compare PC to the standard SMIM and ISMIM. Looking at the average residual correlations in Table 6.2 we see that the PCA10 value of 1.98% is just smaller than the SMIM value of 2.02% and this corresponds to SMIM being close to, but

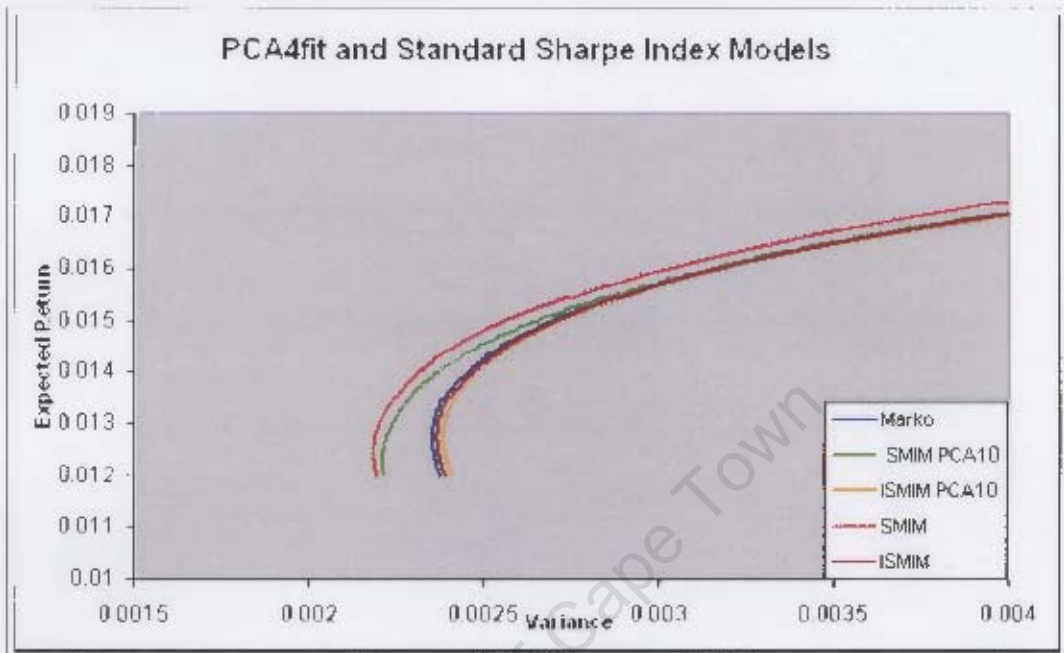


Fig. 6.38. Comparing the PC index model with the standard index models from left to right: SMIM SMIM PCA10, and then cluster together: Markowitz, ISMIM and ISMIM PCA10.

just to the left of, the SMIM PCA10 model. The two ISMIM frontiers cluster together to the right, very close to the Markowitz frontier.

6.4 Conclusion

Even though PC index models incorporate the maximum variance in the fewest number of indices, this does not guarantee meaningful model fit and good index models. Also it is possible to create better fitting models using lower order PC's than those that explain the most variation. PC indices based on the independent variables produce much better models, but do not add inherent value to the portfolio created using the SMIM approach.

What could be useful is to use the PC's of the portfolio to investigate the structure of the variance in a portfolio and the relation of the assets relative to PC coordinates. Also the advantage of the orthogonality of the PC's can be exploited to solve problems where cross correlations add unattainable complications. One example is Multivariate GARCH Models (Tsay,2002, Chapter 9).

In this Chapter we also found an unexpected bonus. Not only did we show how 'weak' regression models, with positive residual correlation, can lead to right shift of the efficient frontier of the Sharpe indices Models compared to those of the Improved Sharpe, we also saw an example of a good fitting model with slightly negative average residuals, producing a SMIM frontier to the right of the associated ISMIM. In Hossain et al (2006a) the authors were forced to use simulated data to construct negatively correlated returns to illustrate that the SMIM overestimates the risk compared with the ISMIM in the negative correlated case. It was possible to verify in this present study the results using real market data from the JSE and the corresponding PC indices.

In this chapter we also introduced and demonstrated the method of using average residual correlation to predict the relative positioning of the efficient frontiers.

Chapter 7

Portfolio Runs

7.1 Introduction

In this section we construct portfolios using different strategies and back test these portfolios over three 5 year periods: 1 January 1993 to 31 December 1997, 1 January 1998 to 31 December 2002 and 1 January 2003 to 28 February 2007. We use the same data set as in the previous chapters. The performance and risk measures of resulting portfolios are evaluated over different periods.

The strategies can be broadly divided into two groups. The first group is the index based strategies based on the previous chapters of this thesis. All these strategies use historic data of the previous 5 years to construct index models and then use mean-variance optimization to find market portfolios. From the previous chapters it is clear that the covariance structure of the ISIM is 'very close to' the covariance structure of the historic covariance and consequently produces similar efficient frontiers and market portfolios. There is therefore no need to investigate any of the permutations of the Improved Sharpe type models and we only include the base improved Sharpe index model. This strategy is also representative of the historic covariance/Markowitz market portfolio. Each of the Sharpe Index Models produced a different portfolio and so four versions: SIM (Chapter 2), SMIM (Chapter 3), ARGASIM (Chapter 4) and weighed ARCTAN ISIM (Chapter 5) are included.

The second group is a set of 5 alternative strategies including a JSE Overall Index tracker and an equally weighed share portfolio, and other strategies that are comprehensively discussed in section 7.4.

7.2 Assumption for portfolio construction

- No dividend payment. All dividend payments were excluded in return calculations for the shares.
- No transaction cost was considered when rebalancing the portfolios.
- The three portfolio reweighing dates were: 1 January 1993, 1 January 1998 and 1 January 2003.
- The portfolios were not rebalanced between the reweighing dates.
- Only information available on the reweighing dates were used to construct the portfolios.
- Alpha and Beta risk measures were calculated using the JSE Overall Index.
- The Banker Acceptance (BA) interest rates at the reweighing dates was used as the risk less rate to calculate the market portfolio.
- For all of the index type strategies the mean-variance optimization was done with an additional constraint that limits the maximum holding in any equity to 15%.

7.3 The Data Set

Table 7.1 explains how the data set was divided into periods for portfolio construction and portfolio runs.

	0th Period	1st Period	2nd Period	3rd Period
Dates	Aug 1988 - Dec 1993	Jan 1993 - Dec 1997	Jan 1998 - Dec 2002	Jan 2003 - Feb 2007
Used for portfolio construction ?	Yes	Yes	Yes	No
Portfolio run period ?	No	Yes	Yes	Yes
Table 7.1 The portfolio run periods				

7.4 Strategies for Portfolio Construction

7.4.1 Sharpe Index Model Strategy (SIM)

The mean and SIM covariance of the period are used to create market portfolios on each reweighing date. For this, and all the index type strategies that follow, the mean-variance optimization was done with an additional constraint that limits the maximum holding in any equity to 15%. This creates more realistic portfolios and avoids the tendency of mean-variance optimization to concentrate the portfolio on only a few of the available shares.

7.4.2 Improved Sharpe Index Model Strategy (ISIM)

It was shown in the previous chapters that the Improved Sharpe Index Model yields identical portfolios to the standard Markowitz historic mean-variance model. It is therefore pointless to include the additional ISIM models (ISMIM, ARGASIM, etc.) in this chapter since

they produce identical portfolios. We therefore only use the standard ISIM/Markowitz strategy.

7.4.3 Sharpe Multi Index Model Strategy (SMIM)

The SMIM uses the same methodology as the SIM strategy, but with the standard four indices discussed in Chapter 3.

7.4.4 Index Model with ARMA and GARCH(1,1) (ARGA SIM)

This strategy is similar to the SIM but with the addition of the ARMA and GARCH(1,1) terms in the regression model to estimate the covariance matrix.

7.4.5 ISIM with ArcTan weighted returns (ARCTAN ISIM)

In this strategy the ArcTan weighting function of Chapter 5 with $\tau = 0.15$ and $\lambda = 0.3$ is used on the 5 year historic data prior to the reweighing dates. A constrained mean-variance optimization is then performed and the market portfolio calculated.

The weights assigned to these portfolios are on the reweighing dates are displayed in Figure 7.39 to 7.41.

7.4.6 Equal weighted share portfolio

In this strategy we construct a portfolio of all 20 share, equally weighted. The portfolio is never reweighed and remains constant from 1 January 1993 until 28 February 2007.

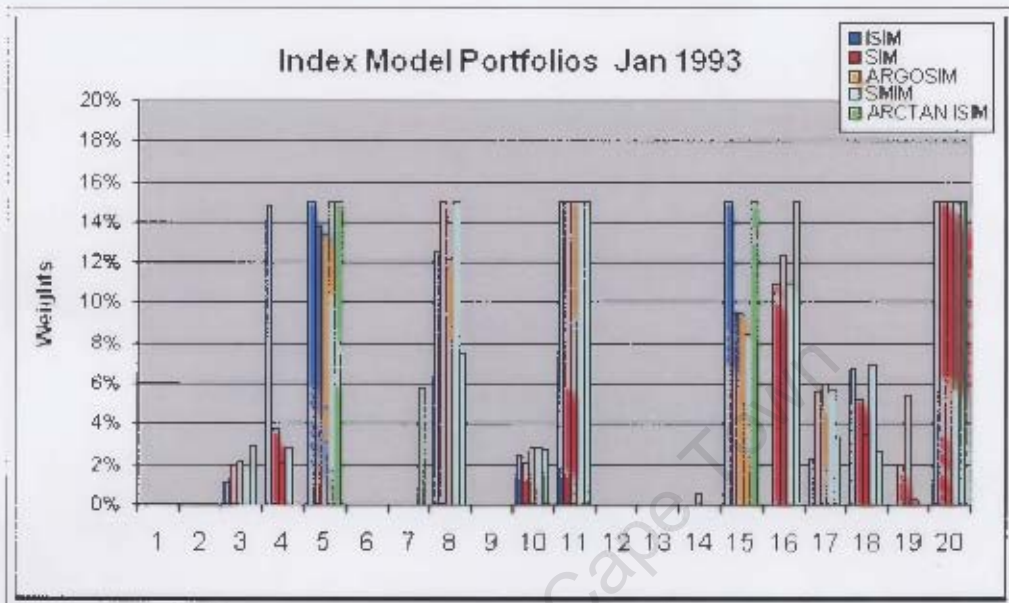


Fig. 7.39. The constituents of the index models type portfolio constructed on the first reweighing date.

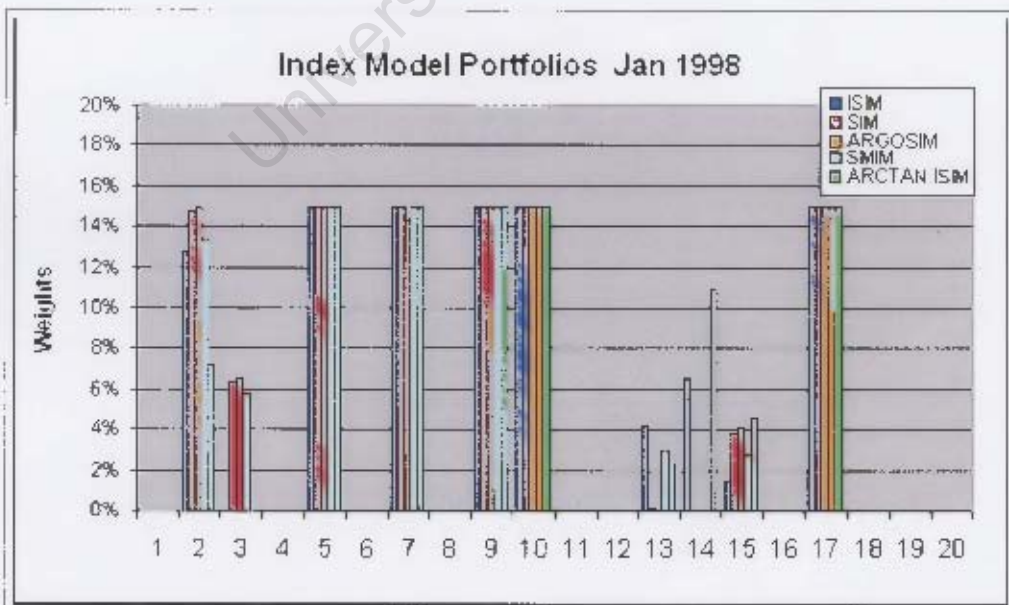


Fig. 7.40. The constituents of the index models type portfolio constructed on the second reweighing date.

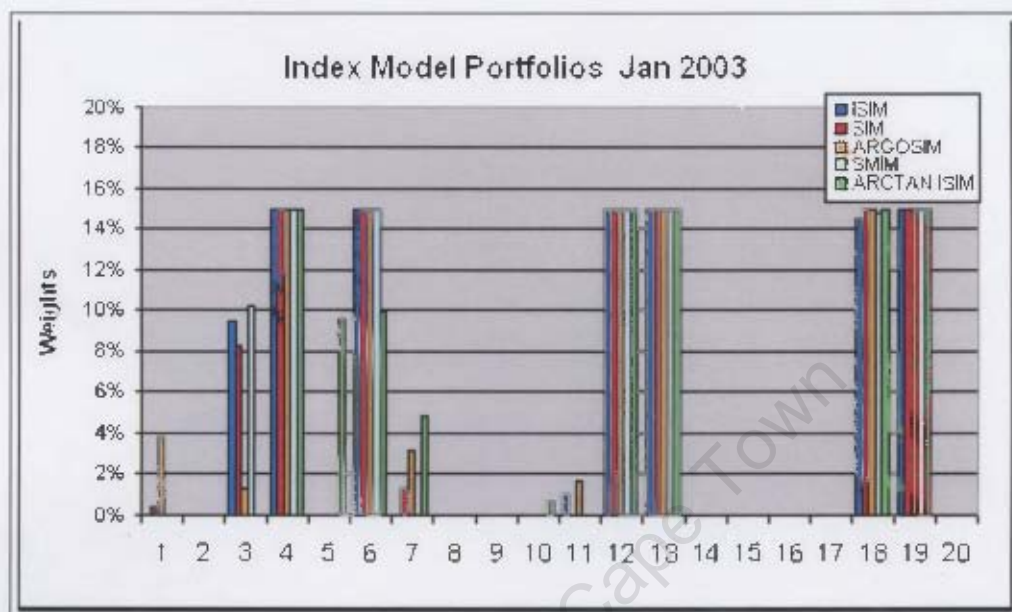


Fig. 7.41. The constituents of the index models type portfolio constructed on the last reweighing date.

7.4.7 JSE Overall Index

In this strategy the full portfolio is invested in the JSE Overall Index. The portfolio is not reweighed on the reweighing dates, but dynamically follows the index, identical to an index tracker fund.

7.4.8 "7 Best" Portfolio

The strategy works as follows: Using the set of 20 shares, choose the 7 best performing shares of the period prior to the reweighing dates. Construct a portfolio which weights these seven shares equally, i.e. $1/7$ each. The portfolio of shares is shown in Figure 7.42.

Each colour represents a portfolio constructed on a different reweighing date.

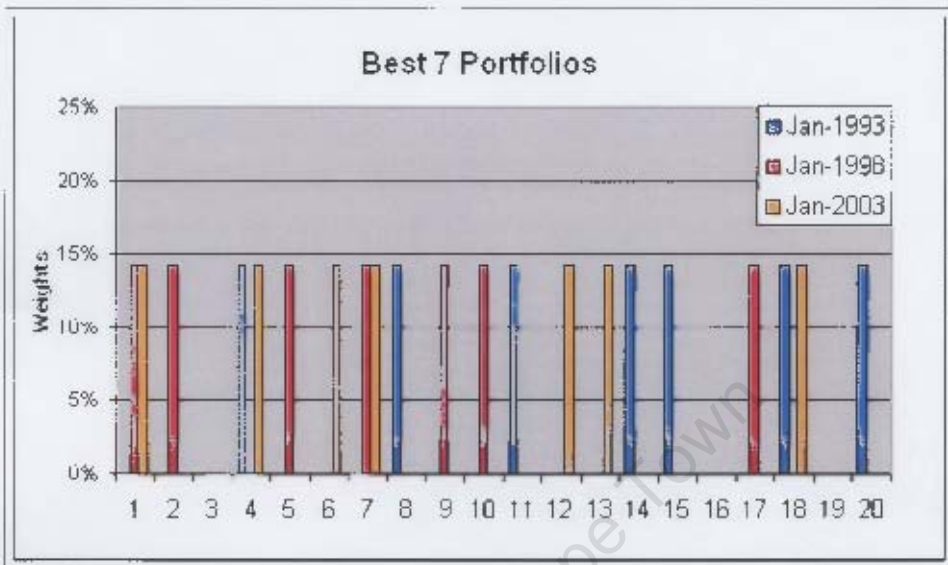


Fig. 7.42. The Best 7 strategy.

7.4.9 "7 Worst" Portfolio

Using the set of 20 shares set, choose the 7 worst performing shares of the period prior to the reweighing dates. Construct a portfolio of these 7 shares weighted $1/7$ in each. The portfolio of shares is shown in Figure 7.43. Each color represents a portfolio constructed on a different reweighing date.

7.4.10 PCA 1

The strategy is as follows: On each reweighing date, use the eigenvalues of the covariance matrix of the previous 5 years to calculate the first principal component of the 20 chosen shares. Construct portfolio weights based on the first principal component. This PCA1 portfolio is the combination of shares which produced the biggest variance over the previ-

ous period. This strategy generates a well balanced portfolio comprising positive weights in all the shares. The constituents of the PCA1 portfolio over time are shown in Figure 7.44.

7.5 The Performance of the Strategies

7.5.1 Market Overview

The performance of the JSE Overall Index from 1993 to 2007 is displayed in Figure 7.45 and quantified in Table 7.2.

The first run period (1993-1997) can be best categorized as a period of political unrest and instability in South Africa. Surprisingly, it was a period of steady growth on the market with relatively low volatility. This can be seen by considering the JSE Overall Index sigma (volatility) in the first period in Table 7.2.

The second run period (1998-2002) includes times of massive world market turmoil. From April 1998 to September 1998 there was the emerging market debt crisis. This was followed by the internet bubble and the period from September 2001 to February 2002 including 9/11 and the spike in the rand exchange rates. The results of these events can be observed in Table 7.2, if one considered the weaker performance and higher volatility of the JSE in second period compared to the first.

The third run period (2003 - February 2007) was categorized by consistent world growth, powered by China and the commodity 'super cycle'. In Table 7.2 it is observed that performance was extremely high and the volatility was historically very low. Figure 7.45 shows a plot of the JSE Overall Index over time in Figure 7.45 and indicates the three

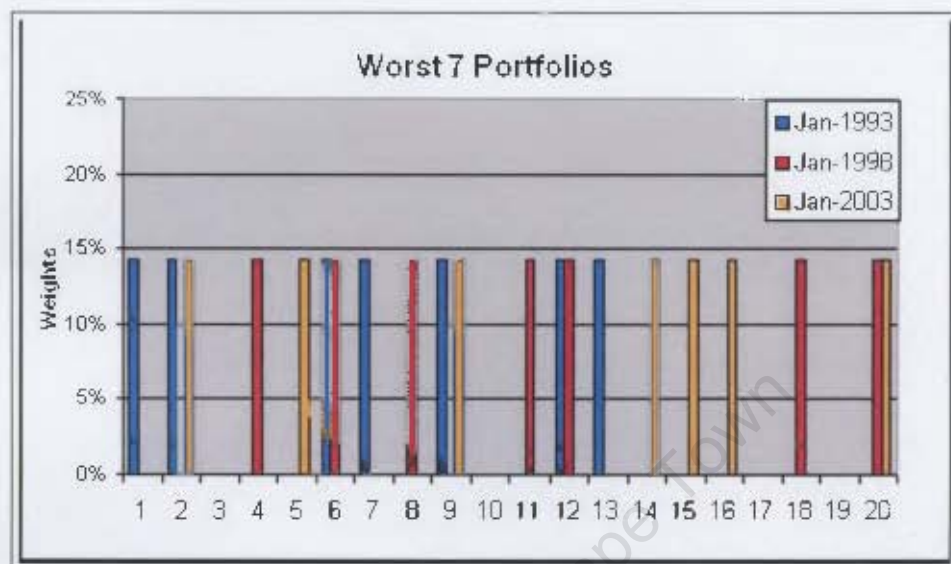


Fig. 7.43. The Worst 7 shares strategy.

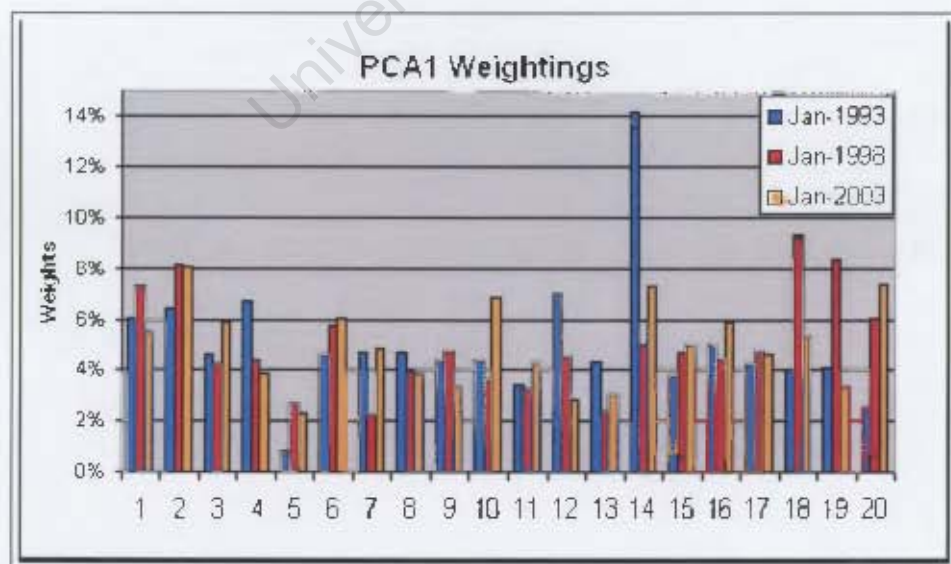


Fig. 7.44. The constituents of PCA1 portfolio.

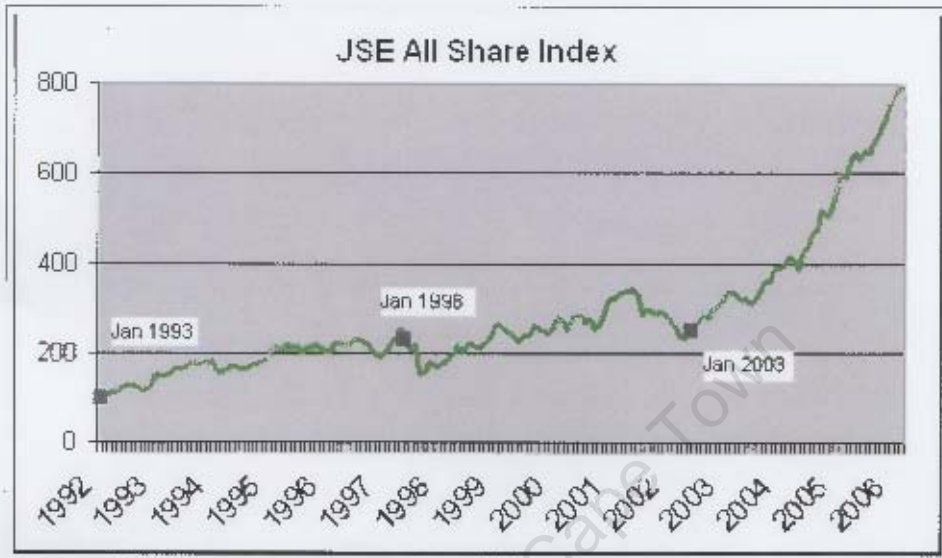


Fig. 7.45. The JSE Overall Index 1993- 2007.

reweighing dates. Prominent is the volatility in the second period and excessive growth in the third.

7.5.2 Performance of the Index Model Portfolios

In this section we consider the performance of the strategies based on the index-type models and the ArcTan weighted strategy. In Figure 7.46 and Table 7.2 we show the performances of the model relative to the JSE Over All Index. If we consider the total return in Table 7.2, we see that all the strategies underperformed the JSE Overall Index, with the exception of the ArcTan ISIM strategy. The ArcTan strategy however only outperformed the JSE Overall Index by 4% over 14 years.

1993-1997					
	Alpha	Beta	Performance	Sigma	
SIM	5.17%	0.78	114%	16.4%	
ARGOSIM	5.57%	0.79	120%	16.5%	
ISIM	4.11%	0.79	104%	16.5%	
SMIM	5.62%	0.78	119%	16.4%	
JSE Overall	0	1.00	91%	15.4%	
ARCTAN ISIM	5.57%	0.79	128%	15.6%	
1998-2002					
	Alpha	Beta	Performance	Sigma	
SIM	-2.21%	0.89	28%	28.4%	
ARGOSIM	-2.47%	0.89	26%	28.5%	
ISIM	0.01%	0.90	44%	27.8%	
SMIM	-0.24%	0.88	41%	27.8%	
JSE Overall	0	1.00	50%	27.4%	
ARCTAN ISIM	-2.47%	0.89	37%	27.4%	
2003-2007					
	Alpha	Beta	Performance	Sigma	Tot Return
SIM	-3.74%	1.02	155%	19.2%	568%
ARGOSIM	-4.54%	1.06	156%	19.7%	578%
ISIM	-3.38%	1.01	155%	19.0%	614%
SMIM	-3.33%	1.01	155%	19.0%	653%
JSE Overall	0	1.00	193%	16.2%	693%
ARCTAN ISIM	-4.54%	1.06	167%	18.9%	697%

Table 7.2 Portfolio Results. Alpha is the annualized out performance of the portfolio over the JSE Overall Index. Beta is the regression coefficient of each portfolio as regressed against the JSE Overall Index for that period. The performance value is for the full period.

From Table 7.2 we observe that all the strategies showed similar results. In the first period, all the strategies out performed the JSE Overall Index. However these strategies under performed the JSE Overall Index, both in the second period of higher volatility and the last period of higher performance. Interestingly it is clear from Table 7.2 that the Beta values tend to increase by about 10% in each subsequent period for all the strategies. The higher a given Beta of a strategy, the lower the out performance, Alpha, of that strategy.

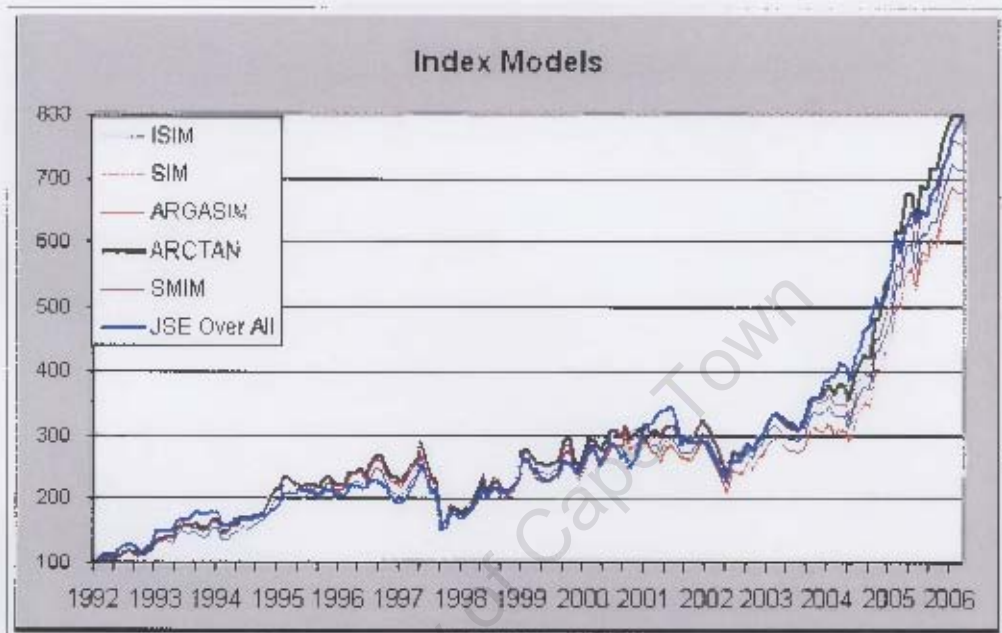


Fig. 7.46. Index Models Performance.

7.5.3 Performance of the Alternative Portfolios

The relative performances of the alternative strategies to the JSE Overall Index are shown in Figure 7.46. The JSE Overall Index is the second from the bottom. The equally weighted and PCA1 strategies performed very similarly. This is quite surprising if we compare the construction of PCA1 portfolios in Figure 7.44 to that of an equally weighted portfolio. From Table 7.3 we can see that both these strategies outperform the JSE Overall Index consistently over all three periods. The total performances for both these strategies are double that of the index.

1993-1997					
	Alpha	Beta	Performance	Sigma	
7 Best	1.39%	0.95	72%	18.2%	
JSE Overall	0.00%	1.00	91%	15.4%	
equal	6.17%	0.95	151%	16.0%	
PCA 1	4.91%	1.01	145%	16.7%	
7 Worst	6.72%	1.11	186%	21.2%	
1998-2002					
	Alpha	Beta	Performance	Sigma	
7 Best	-0.51%	0.95	43%	28.6%	
JSE Overall	0.00%	1.00	50%	27.4%	
equal	7.98%	0.87	112%	26.9%	
PCA 1	7.43%	0.90	108%	27.9%	
7 Worst	12.29%	0.87	162%	27.1%	
2003-2007					
	Alpha	Beta	Performance	Sigma	Tot Return
7 Best	-10.25%	1.21	135%	22.2%	450%
JSE Overall	0.00%	1.00	193%	16.2%	693%
equal	5.31%	0.79	186%	14.1%	1385%
PCA 1	8.25%	0.74	203%	13.8%	1432%
7 Worst	15.53%	0.54	225%	15.6%	2394%

Table 7.3 Portfolio Results. Alpha is the annualized out performance of the portfolio over the JSE Overall Index. Beta is the regression coefficient of each portfolio as regressed against the JSE Overall Index for that period. The performance value is for the full period.

Next we consider the "7 Best" and "7 Worst" strategies. Choosing a portfolio of the 7 best performing shares produces the worst performance from the group. Basing a portfolio on the 7 worst performing shares produces a very profitable strategy.

The "7 Worst" strategy produced an annual Alpha (out performance of the index) of 6.72% in the first period, 12.72% in the second and 15.53% in the third. The strategy outperformed the JSE Overall Index by 3.46 times for the full period with a total performance of 2,384%. We again notice the trend of a lower Beta portfolio associated with higher Alpha value.

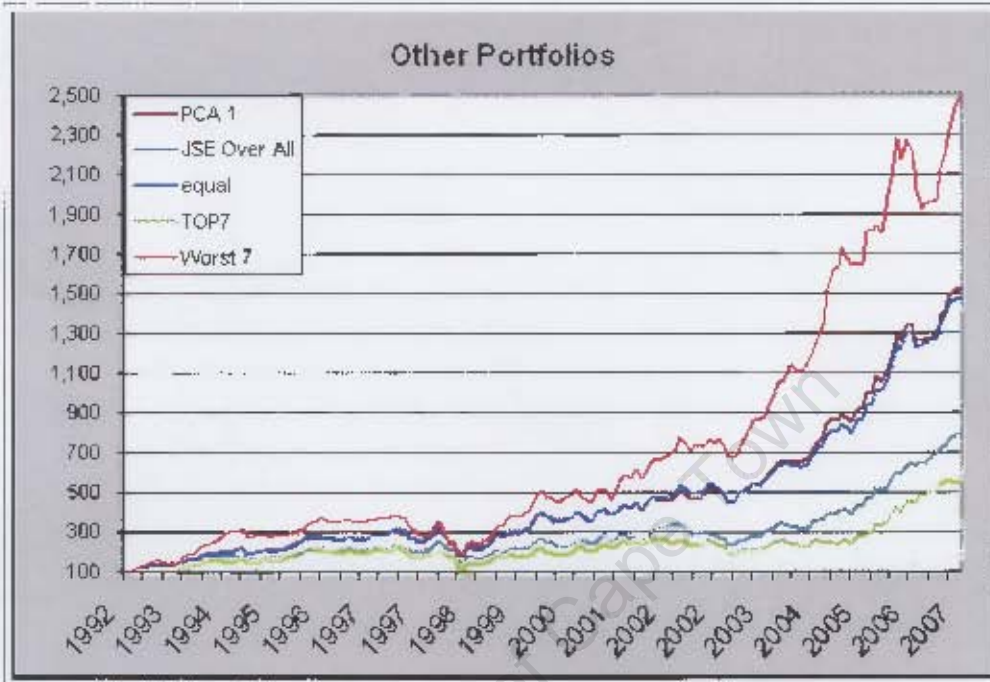


Fig. 7.47. Alternative portfolios

7.6 Conclusion

Compared to the alternative strategies, the index type models add very little value. Investing in a JSE index tracker would have outperformed most of these models. We have to note that ignoring extreme values, using the ArcTan weighed returns, and implementing mean-variance optimizing to generate market portfolios, did outperform the other market portfolio strategies and the JSE Overall Index

The PCA1 strategies were disappointing and seem to duplicate the basic equally weighted share strategy. The remarkable result is the exceptional return of the simple "7 Worst" shares strategy. The number of 7 shares was chose arbitrarily by dividing the 20 share into thirds. The 5 years between reweighing dates seem to have picked up cyclical

patterns of over and under performance of shares. This is again confirmed when considering the "7 Best" strategy performance. The question remains as to how this strategy will fair over shorter time frames and in different share universes?

University of Cape Town

patterns of over and under performance of shares. This is again confirmed when considering the "7 Best" strategy performance. The question remains as to how this strategy will fair over shorter time frames and in different share universes?

University of Cape Town

Chapter 8

Summary and Conclusions

In this mini-dissertation we attempted to appraise and assess the use of Index Models in portfolio construction. After reviewing the work of Hossain (2006) and others in this regard, we decided to repeat his empirical studies with a bigger data set of 20 shares, including Hossain's original nine, but over a longer time period of 15 years. Hossain's (2006) findings regarding the relative positioning of the efficient frontiers of the Sharpe Index Model and Improved Sharpe Index Model for positive correlated residuals were reproduced and confirmed in Chapters 2 and 3. Also, the left shift effect on the efficient frontier of the ARMA and GARCH(1,1) models, called ARGA, was reproduced and confirmed in Chapter 4. This left shift effect of the efficient frontier was most visible in the Sharpe Index Model (SIM) case, to a lesser extent in the Sharpe Multi Index Model (SMIM) and was almost negligible for the ISMIM and ISIM.

It was found that using the ISIM as opposed to the SIM to model portfolio covariance reproduced the efficient frontier of the Markowitz model. Furthermore, even though the use of the ARGA extension of the Index Model, presented in Chapter 4, does move the efficient frontier to the left, it does not mean that the model is superior or a more accurate reflection of the future risk return profile of the shares. This only means that the model discounts the residual correlation compared to the ISIM or Markowitz covariance models. This left shifting effect can be reduced by adding more indices or using the more accurate ISMIM. The ARGA models also had a minimal effect on the resulting market portfolios

and performed similarly in the portfolio runs of Chapter 7, with the ISIM outperforming the ARGASIM slightly over the 15 years considered.

In Chapter 5 of this dissertation we investigated bounded influence regression. We introduced the parameterized ArcTan function and showed how it can be used to move the efficient frontier associated with the chosen set of shares to the left or to the right. Different weighting functions have drastic effects on the resulting market portfolios. Moreover, it was shown that ARGASIM extensions had a minimal effect on the resulting efficient frontier for weighted returns. We also demonstrated that aggressive weighting produces more equally weighted portfolios. It is a similar effect to that of constraining the maximum holding in any one asset in a portfolio.

In Chapter 6 we created Principal Component (PC) indices and used them in the SMIM and ISMIM. We showed that even though the PC indices incorporate the maximum variance in the fewest number of indices, this does not guarantee meaningful model fit and good index models. Also, it is possible to create better fitting models using lower order PC's than those that explain the most variation. In this chapter we also saw an example of a good fitting model, using the PCA of the original shares, with slightly negatively correlated residuals, producing a SMIM with an efficient frontier to the right of that associated with the ISMIM. Hossain (2006) and others were forced to use simulated data to construct negatively correlated returns to illustrate the SMIM overestimating the risk compared with the ISMIM in this case. Here it was possible to verify their results using real market data from the JSE and the corresponding PC indices.

Finally in Chapter 7 we discussed the backtesting of the models over the last 15 years. We found that, compared to the other basic non-index strategies, the index type models add very little value. Investing in a JSE index tracker would have outperformed most of these index models. We also note that ignoring extreme values by using the ArcTan weighed returns to generate market portfolios did outperform the index model strategies and the JSE Overall Index. The PC based strategies were disappointing and seemed to duplicate the basic equal weighted share strategy. The remarkable result in this backtesting study is the exceptional return of the simple "7 worst" shares strategy.

Regression is an important tool in financial research and is used to appraise and to test the relation of a chosen instrument to other measurable factors in the market. If such a regression model is established, it can be used to find opportunities for which the instrument trades below or above what the model would predict, based on the assumptions of the model. Also, if assumptions are made about future values of the factors, this can be used to predict the future value of the instrument. Using regression models on historic data to create covariance matrices without forecasts or assumptions just produces the historic covariance matrix at best. It is therefore necessary to combine these techniques with assumptions and forecasts of factors to be meaningful in the creation of portfolio strategies.

- Hossain N., Troskie C.G. and Guo R. (2006a), "An Application of Principal Components Analysis to Portfolio Optimization under the Sharpe Multiple Index Framework", *Proceedings of the International Workshop of Financial Engineers*, Hong Kong, June 20-22, 2006, pp 396-401. (ISBN-13: 978-988-98671-3-3). Edited by S.I. Ao, J.A. Lee, O. Casitllo, P. Chaudhuri, D.D. Feng
- Hossain N., Troskie C.G. and Guo R. (2006b), "Comparisons of Ex Post Efficient Portfolios Generated from Sharpe and Improved Sharpe Multiple Index Models; An Empirical Study concerning the Negative Correlated Case", *Proceedings of the 11th International Conference on Industrial Engineering, Theory, Applications and Practice*, Nagoya, Japan, October 24-27, 2006.
- Mallows C.L. (1973), "Influence Functions", Unpublished Paper presented at a conference on robust regression held at Cambridge, Massachusetts and sponsored by the National Bureau of Economic Research.
- Markowitz H.M. (1952), "Portfolio Selection", *Journal of Finance*, 7, 77-91.
- Press J.S. (1972), *Applied Multivariate Analysis*, Holt, Rinehart and Winston, New York.
- Rice J.A. (1995), *Mathematical Statistics and Data Analysis*, Duxbury Press, Belmont, California.
- Ruppert D. (2004), *Statistics and Finance, An Introduction*, Springer Verlag, New York.
- Sharpe W.F. (1964), "Capital Asset Prices: a Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, 19, 425-42.
- Troskie C.G. (2000), "Principal Components Analysis" *Class notes*.
- Troskie C.G. (2004), "Financial Regression and Time Series Modelling", *Class notes*.
- Tsay R.S. (2002), *Analysis of Financial Time Series*, John Wiley & Son, New York.

Appendix A

EIEWS Code

University of Cape Town

Appendix A

EViews Code

Sample EViews3: code the calculate returns, do regression, and calculate estimator.

```
'Share Multi Index Models of 20 share Portfolio
'and 4 used indexes
'load sindextrow
scalar cnt = 223
```

```
'Define Numeric Variables with ln
```

```
ln=cnt
lk=20          'number of shares
'lp=8          'No idear what this is
lm=4          'Number of indexes
```

```
'There variables get use for Mark.SIM,ISIM and Weighted Regeression SIM,ISIM
```

```
vector(lk) alpha      'coef(1)
matrix(lk,lm) beta    'is a matrix(k shares by m index)
vector(lk) tstats     'collect tstats
vector(lk) rsquare     'to check R^2
vector(*) rsquare_adj_sub      'to check R^2 adj
vector(lk) rsquare_adj_full    'to check R^2 adj
vector(lk) schwarz_sub        'the Schwarz criterion for model building
vector(lk) schwarz_full      'the Schwarz criterion for model building
vector(lk) rsquare_w          'to check R^2 of the weighted models
vector(lk) rsquare_ar         'to check R^2 of the arma models
vector(lk) rsquare_garch_only 'to check R^2 of the garch models
vector(lk) rsquare_garch     'to check R^2 of the garch models
vector(*) rsquare_war         'to check R^2 of the weighted arma models
vector(lk) rsquare_wgarch    'to check R^2 of the weighted garch models
vector(lk) stdError         'Meture standard error
vector(lk) stdError_arma    'Meture standard error
vector(lk) stdError_garch   'Meture standard error
vector(lk) stdError_ARGA    'Meture standard error
```

```
vector(lk) evar          'estimation of variance of reseduals
vector(lk) Ereturn      'estimated returns
vector(lk) Ereturn_ar   'estimated returns based on ARMA model
vector(lk) Ereturn_gar  'estimated returns based on GARCH model
vector(lk) Ereturn_w    'Estimated Returns for Weighted securities
vector(lk) Ereturn_war  'Estimated Returns for Weighted securities
vector(lk) Ereturn_wgar 'Estimated Returns for Weighted securities
vector(lk) hisreturn    'historic returns
vector(lk) hisreturn_w  'Weighted historic returns
```

```
matrix(k,lk) covar      'historic covariance Mark
'matrix(lk,lk) covar_w   'historic covariance Mark Weighted
matrix(lk,lk) covar_ar  'historic covariance Mark ARMA ~ going to be the same ?
matrix(lk,lk) covar_gar 'historic covariance Mark GARCH ~ going to be the same ?
```

```
matrix(lm,lm) covar_index 'Covariance of the indexes
matrix(lk,lk) Ecovar      'estimated coviance sharpe single index
matrix(lk,lk) Ecovar_ar   'estimated coviance sharpe single index with ARMA(1,1)
matrix(lk,lk) Ecovar_gar  'estimated coviance sharpe single index with GARCH(1,1)
matrix(lk,lk) Ecovar_gar_only 'estimated coviance sharpe single index with GARCH(1,1) no ARMA
```

```
'matrix(lk,lk) Ecovar_wr   'estimated coviance sharpe single index with weighted regression
'matrix(lk,lk) Ecovar_war  'estimated coviance sharpe single index with WR ARMA(1,1)
'matrix(lk,lk) Ecovar_wgar 'estimated coviance sharpe single index with WRGARCH(1,1)
```

```
matrix(lk,lk) residcovar 'covariance of the residues
```

```
matrix(k,k) covmat      'correlation matrix of the residues
matrix(k,1) E_hat       'the Residuals
matrix(k,1) Omega_hat   'Estimation of Errors
                                used for all calculations
                                used for all calculations
```

```
matrix(k,1) Theta_hat   'estimate of covariance
matrix(k,1) Theta_hat_ar 'estimate of covariance
matrix(k,1) Theta_hat_gar 'estimate of covariance
matrix(k,1) Theta_hat_gar 'estimate of covariance
matrix(k,1) Theta_hat_wa 'estimate of covariance
matrix(k,1) Theta_hat_wa 'estimate of covariance
matrix(k,1) Theta_hat_wgar 'estimate of covariance
                                ISIM
                                ARMA ISIM
                                GARCH ISIM
                                GARCH ISIM ~no ARMA
                                WRISM
                                WRISM
                                WRISM
                                WRISM
```

```
group igroup            'group of returns
group igroups           'index group
group rgroup_w         NOT SURE TO USE THIS
group rgroup_wgar
```

```
group residgroup       'group for resid
group residgroup_ar
group residgroup_gar
group residgroup_garw
group residgroup_wa
group residgroup_wgar
group residgroup_wgar
```

Index returns

```
genr X0 = log(gsewew/gsewew(-1))      'used
genr X1 = log(gsewew/gsewew(-1))
genr X2 = log(djvandjvand(-1))      'used
genr X3 = log(goidr/goidr(-1))
genr X4 = log(djvandjvand(-1))
genr X5 = log(palamb/palamb(-1))     'used
genr X6 = log(rchem/rchem(-1))
genr X7 = log(R100/R100(-1))
genr X8 = log(rindex/rindex(-1))
genr X9 = log(dollar/dollar(-1))     'used
```

```
group igroup X0 X1 X5 X9
```

```
indmean(1) = @mean(x0)
indmean(2) = @mean(x1)
indmean(3) = @mean(x5)
indmean(4) = @mean(x9)
```

```
*Make X matrix (for ARMA - tweak)
smpl 1982:01 2007:02
series b1g = 1
smpl 1982:07 2007:02
group xgroup b1g X0 X1 X5 X9
matrix X = @colvec(xgroup) *223 elements
Delete b1g
```

Markovitch and SIM and ISIM

SIM & ISIM

```
counter = 1
      1      2      3      4 5 6 7 8      9      10      11      12
13 14 15 16 17 18 19 20
for %share angles idjou picnny icnrg saeap sappi nsol  tjbua kingat obsa.      afrox      angpla      hamao
      jounk libat nampak nedcor reap PPC edcon
```

```

*make returns
%rshare = 1 - @str(counter)
%residshare = %resid + @str(counter)
%var %rshare = log(%share/%share(-1))
%group add %rshare

+++++
*test Schwartz for full model
+++++
equation model is %rshare c X0 X1 X2 X3 X4 X5 X6 X7 X8 X9
schwarz_full(counter) = model.@schwarz
rsquare_adj_full(counter) = model.@rbar2
rsquare(counter) = model.@r2

*Run single index models
equation model is %rshare c X0 X1 X5 X9
equation modelSMIM(counter) is %rshare c X0 X1 X5 X9

*Save residuals series
model makeresid %residshare
residgroup add %residshare

*Save estimators
rsquare(counter) = model.@r2
schwarz_sub(counter) = model.@schwarz
rsquare_adj_sub(counter) = model.@rbar2

tlabels(counter) = model.C(2)/model.@rshare(2)
rsquare(counter) = model.@r2
str-nof(counter) = Model.@se

alpha(counter) = model.@coef(1)
beta(counter 1) = model.@coef(2)
beta(counter 2) = model.@coef(3)
beta(counter 3) = model.@coef(4)
beta(counter 4) = model.@coef(5)
envar(counter) = model.@en?
include(counter) = @mean("%share")
counter = counter + 1

next

*Test Returns and Covariance matrix marko
-----
Ereturn = alpha + beta*return

*Test covariance and index covar covar_index
covar = @cov(group)'*(n/(n-1))
covar_index = @cov(group)'*(n/(n-1))
matrix covar_index_old = @cov(group)'*(n/(n-1))

*Beta*C*Beta' + diagonal(covar) = Shape Multi single index
-----
Ecovar = beta*covar_index*@transpose(beta) + @makediagonal(covar)

*resid covariance
residcovar = @cov(residgroup)'*(n/(n-1))
residcov = @cov(residgroup)

*resid matrix (divide by n-m)
I_hat = @converf(residgroup)
Omega_hat = I/@transpose(E_hat)*E_hat/(n-f-m)

*Theta_hat = signalindex*beta*beta' + Omega_hat
-----
Theta_hat = beta*covar_index*@transpose(beta) + Omega_hat
-----

```

```

stem('residshare,TempResid')
vector(n-1) vResid
for i = 1 to la
    vResid(i)=ehat(i)
next
'no = la - la
for i = la + 1 to lres - 1
    vResid(i)*a = TempResid(i)
next
'Make X matrix (for ARMA tweak)
smpl 1963:06 2007:02
mres(vResid,%residshare)
smpl 1963:07 2007:02
else
    'Nothing
endif
endf

```

```

residgroup_ar and %residshare
rsquare(counter) = model.@se

rsquare_ar(counter) = model.@se
alpha(counter) = model.@coef(1)
beta(counter,1) = model.@coef(2)
beta(counter,2) = model.@coef(3)
beta(counter,3) = model.@coef(4)
beta(counter,4) = model.@coef(5)
covar(counter) = model.@se^2
counter = counter + 1
next

```

Build Returns and Covariance matrix's matrix

```

Freturn_ar = alpha + beta*realinear

```

'Beta'*C'*beta' + diagonal(covar) - Sharpe's full single index

```

Ecovar =beta*covar_index*transpose(beta) + @hukudiagonal(covar)

```

'resid matrix

```

E_hat = @convert(residgroup_ar)
Omega_hat = (@transpose(E_hat)*E_hat)/(n*(m+2))

```

'theta_hat = signalindex*beta*beta' + Omega_hat

```

theta_hat_ar = beta*covar_index*transpose(beta) + Omega_hat

```

'%smatler = 1

GARCH(1,1) MSM & GARCH(1,1) MSM

```

13 14 15 16 17 18 19 20 4 5 6 7 8 9 10 11 12
for %share anglog idgton picapay remgr saeanl sappi saool ligbra tongal alisa afax angpis haimo jomic
liber nainpak nedcor reur PPC edcon
%share = r^1405*(counter)

```

For each arma model MSM

```

if(counter = 1) then %armaCurExt = "ar(2)" endif
if(counter = 2) then %armaCurExt = "" endif
if(counter = 3) then %armaCurExt = "ar(1)" endif
if(counter = 4) then %armaCurExt = "ar(1)" endif
if(counter = 5) then %armaCurExt = "" endif
if(counter = 6) then %armaCurExt = "" endif
if(counter = 7) then %armaCurExt = "" endif
if(counter = 8) then %armaCurExt = "ar(1) ar(5)" endif
if(counter = 9) then %armaCurExt = "ar(2)" endif

```

```

if(counter = 10) then %armaCurExt = "ar(1)"      endif
if(counter = 11) then %armaCurExt = "ar(1)"      endif
if(counter = 12) then %armaCurExt = ""          endif
if(counter = 13) then %armaCurExt = ""          endif
if(counter = 14) then %armaCurExt = "ar(1)"      endif
if(counter = 15) then %armaCurExt = "ar(1) ar(2)"endif
if(counter = 16) then %armaCurExt = "ar(1) ar(4)"  endif
if(counter = 17) then %armaCurExt = ""          endif
if(counter = 18) then %armaCurExt = ""          endif
if(counter = 19) then %armaCurExt = "ar(2)"      endif
if(counter = 20) then %armaCurExt = "ar(2)"      endif

```

make resid:

```
%residshare = "resid" : @st(counter) : "_gar"
```

Run single index models with ARMA

```
equation model m1(d(1,1)) %share c X0 X1 X5 X9 %armaCurExt
equation model GARCH(counter) arch(1,1) %share c X0 X1 X5 X9 %armaCurExt
```

Save residual series

```
model.makesresid %residshare
```

Replace the missing residue item that was remove using the by modeling the AR terms.

```

if(counter = 1) then %a = 2      endif
if(counter = 2) then %a = 0      endif
if(counter = 3) then %a = 1      endif
if(counter = 4) then %a = 1      endif
if(counter = 5) then %a = 0      endif
if(counter = 6) then %a = 0      endif
if(counter = 7) then %a = 0      endif
if(counter = 8) then %a = 0      endif
if(counter = 9) then %a = 2      endif
if(counter = 10) then %a = 1     endif
if(counter = 11) then %a = 1     endif
if(counter = 12) then %a = 0     endif
if(counter = 13) then %a = 0     endif
if(counter = 14) then %a = 1     endif
if(counter = 15) then %a = 1     endif
if(counter = 16) then %a = 4     endif
if(counter = 17) then %a = 0     endif
if(counter = 18) then %a = 0     endif
if(counter = 19) then %a = 2     endif
if(counter = 20) then %a = 2     endif

```

if (%a > 1) then

```
vector %beta = model @beta
vector %beta = @subexract(%beta, 1, 1, 5, 1) The Alfa and beta
```

```
==> for %i=1 to %a
```

```
matrix %share, %share
```

```
vector %val = share - %i*beta
```

```
store(%residshare, TempResid)
```

```
vector(%i-1) vResid
```

```
for %j = 1 to %a
```

```
    vResid(%j) = %val(%j)
```

```
next
```

```
for %i = %a + 1 to %a - 1
```

```
    vResid(%i+1) = TempResid(%i)
```

```
next
```

```
Make X matrix (for ARMA Tweed)
```

```
smpl 1998:08 2007:02
```

```
mtos(vResid, %residshare)
```

```
smpl 1998:07 2007:02
```

else

```
Nothing
```

endif

```
resid(oup_gar) add %residshare
```

```
estimate_gar(counter) = model @est
```

```
sigma(counter) = model @sigma(1)
```

```
beta(counter, 1) = model @coef(2)
```

```

beta(counter,2) = model.@coef(3)
beta(counter,3) = model.@coef(4)
beta(counter,4) = model.@coef(5)
covar(counter) = model.@se^2
counter = counter + 1

```

next

Build Returns and Covariance matrix marks

```

Ereturn_gar = alpha + beta*evmvars

```

signalIndex="Beta"Beta" + diagonal(evar) ~Sharpe Multi Index Model

```

Ecovar_gar = beta*covar_index*@(transpose(beta)+ @makediagonal(evar))

```

resid matrix

```

E_hat = @convert(residgroup_gar)
Omega_hat = (@@transpose(E_hat)*E_hat)/(('m'+4))

```

Theta_hat = signalIndex*beta*beta' + Omega_hat

```

Theta_hat_gar = beta*covar_index*@(transpose(beta)+ Omega_hat)

```

counter = 1

GARCH(1,1) ONLY SMB & GARCH(1,1) ONLY ISMM

```

-----
13 14 15 16 17 18 19 20 3 4 5 6 7 8 9 10 11 12
for %share %rjlos plnmu nicopay tempa saeagi sangi sazol tigra tngal alsa afax angpia homo
jonic ibert nampak nedcor reut PPG edcov
%share = r' + @sin(counter)
make resid
%residshare = "resid" + @sin(counter) + "_gar"
Run single index models with ARMA
equation model.arch(1,1) %share c X0 X1 X5 X9
@qabun model.GARCH(counter) arch(1,1) %share c X0 X1 X5 X9
Save residuals series
model.makeresid %residshare
residgroup_gar add %residshare

Save estimators
%beta(counter) = model.C(2)/model.@stderrs(7)
%beta(counter) = model.@se

covars_gar_only(counter) = model.@se
alpha(counter) = model.@coef(1)
beta(counter,1) = model.@coef(2)
beta(counter,2) = model.@coef(3)
beta(counter,3) = model.@coef(4)
beta(counter,4) = model.@coef(5)
covar(counter) = model.@se^2
counter = counter + 1
next

```

signalIndex="Beta"Beta" + diagonal(evar) ~ sharpe single index

```

Ecovar_gar_only = beta*covar_index*@(transpose(beta)+ @makediagonal(evar))

```

resid matrix

```

E_hat = @convert(residgroup_gar)
Omega_hat = (@@transpose(E_hat)*E_hat)/(('m'+3))

```

Theta_hat = signalIndex*beta*beta' + Omega_hat

```

Theta_hat_gar_o = beta*covar_index*@(transpose(beta)+ Omega_hat)

```

Appendix B

MATLAB Code

University of Cape Town

Appendix C

VBA Code

University of Cape Town

Appendix C

VBA Code

VBA code to calculate the Mallows weights:

```
Public Function Mallowsw(arrRet)
'Dirk Weighting function using heap sort.

Dim i As Integer
Dim rngStart As Range
Dim arrNewRet() As Double

Dim crit As Integer, crit = UBound(arrRet)
Dim L As Integer
Dim U As Integer
Dim pt As Double, pt = Range("G1")
Dim fRange As Double, fRange = Range("F1")
Dim stretch As Double, stretch = Range("E1")

ReDim arrNewRet(crit)
ReDim w(crit)
Dim avg As Double
Dim Lsum, Usum As Double
Dim sum As Double, sum = 0#
Dim Lsum As Double, Lsum = 0#
Dim Usum As Double, Usum = 0#

For i = 1 To crit
    sum = sum + arrRet(i)
Next
avg = sum / crit
For i = 1 To crit
    arrRet(i) = arrRet(i) - avg
Next
Application.ScreenUpdating = False

'DO SORT
Dim index, index = HeapSort(arrRet)

L = pt * crit + 1
U = crit + 1 - L
Debug Print L
Debug Print arrRet(index(L))
Debug Print U
Debug Print arrRet(index(U))

For i = 1 To crit
    If i = 1 Then
        w(index(i)) = (arrRet(index(L)) - arrRet(index(U))) / (2 * arrRet(index(i)) - arrRet(index(U)) - arrRet(index(L)))
        Lsum = Lsum + w(index(i))
        Usum = Usum + 1 - w(index(i))
    Else
        w(index(i)) = (arrRet(index(U)) - arrRet(index(L))) / (2 * arrRet(index(i)) - arrRet(index(U)) - arrRet(index(L)))
        Usum = Usum + w(index(i))
    End If
    arrNewRet(index(i)) = w(index(i)) * arrRet(index(i))
Next

Lsum = Lsum / 2
Usum = Usum / 2
Debug Print Lsum
Debug Print Usum
```

```
Mallorem = arrNewRet
Application.ScreenUpdating = True
```

```
End Function
```

VBA code to calculate the ArcTan weights:

```
Option Explicit
```

```
Public Function Linkem(arrRet)
```

```
    'Link Weighing function using heap sort.
```

```
    The ARCTAN FUNCTION
```

```
    Dim i As Integer
```

```
    Dim rngStart As Range
```

```
    Dim arrNewRet() As Double
```

```
    Dim crl As Integer: crl = UBound(arrRet)
```

```
    Dim l As Integer
```

```
    Dim U As Integer
```

```
    Dim pi As Double: pi = Range("PI")
```

```
    Dim fRange As Double: fRange = Range("F")
```

```
    Dim stretch As Double: stretch = Range("E")
```

```
    ReDim arrNewRet(crl)
```

```
    Application.ScreenUpdating = False
```

```
    'NO SORT
```

```
    Dim Index: Index = HeapSort(arrRet)
```

```
    L = pi * crl + 1
```

```
    U = crl - 1 - L
```

```
    Debug.Print L
```

```
    Debug.Print arrRet(Index(L))
```

```
    Debug.Print U
```

```
    Debug.Print arrRet(Index(U))
```

```
    For i = 1 To crl
```

```
        If i < L Then
```

```
            arrNewRet(Index(i)) = arrRet(Index(L)) * (1 - fRange) * (2 / Application.WorksheetFunction.Pi()) * Ah(stretch * arrRet(Index(i)) - arrRet(Index(L))) / arrRet(Index(L))
```

```
        Else If i = U Then
```

```
            arrNewRet(Index(i)) = arrRet(Index(L)) * (1 + fRange) * (2 / WorksheetFunction.Pi()) * Ah(stretch * (arrRet(Index(i)) - arrRet(Index(L))) / arrRet(Index(U)))
```

```
        Else
```

```
            arrNewRet(Index(i)) = arrRet(Index(i))
```

```
        End If
```

```
    Next
```

```
    Dim m = arrNewRet
```

```
    Application.ScreenUpdating = True
```

```
End Function
```

```
'Heap sort routine.
```

```
'Requires a sorted Index array for the Keys array.
```

```
'Author: Christian d'Heureuse (www.source-code.biz)
```

```
Public Function HeapSort(Keys)
```

```
    Dim Base As Long: Base = LBound(Keys) 'array index base
```

```
    Dim n As Long: n = UBound(Keys) - LBound(Keys) + 1 'array size
```

```
    ReDim Index(Base To Base + n - 1) As Long 'allocate index array
```

```
    Dim i As Long, m As Long
```

```
    For i = 0 To n - 1: Index(Base + i) = Base + i: Next 'fill index array
```

```
    For i = n \ 2 - 1 To 0 Step -1 'generate ordered heap
```

```
        Heapify Keys, Index, i, n
```

```
    Next
```

```
    For m = n - 1n 2 Step -1
```

```
        Exchange Index, 0, m - 1 'move highest element to top
```

```
        Heapify Keys, Index, 0, m - 1
```

```
Next
HeapSort = Index
End Function
```

Sort Help function

Included for completeness, not my work.

```
Private Sub Heavily(Keys, Index) As Long, ByVal i) As Long, ByVal n As Long)
    'Heap order rule, a[i] <= a[2*i+1] and a[i] <= a[2*i+2]
    Dim Base As Long, Base = LBound(Index)
    Dim nDiv2 As Long, nDiv2 = n \ 2
    Dim i As Long, i = 1
    Do While i < nDiv2
        Dim k As Long, k = 2 * i + 1
        If k + 1 < n Then
            If Keys(Index(Base + k)) < Keys(Index(Base + k + 1)) Then k = k + 1
        End If
        If Keys(Index(Base + i)) >= Keys(Index(Base + k)) Then Exit Do
        Exchange Index, i, k
        i = k
    Loop
End Sub
```

```
Private Sub Exchange(a() As Long, ByVal i As Long, ByVal j As Long)
    Dim Base As Long, Base = LBound(a)
    Dim Temp As Long, Temp = a(Base + i)
    a(Base + i) = a(Base + j)
    a(Base + j) = Temp
End Sub
```

```
Public Function Array_Sort(ByVal NotSortedArray As Variant) As Variant
    'NotSortedArray must be a 1D array
```

```
Make a sorted array
Dim i As Long, i = Counter
Dim j As Long, j = Counter
Dim v() As Variant, v = Each element
```

```
For i = LBound(NotSortedArray) To UBound(NotSortedArray)
    For j = i + 1 To UBound(NotSortedArray)
        If NotSortedArray(i) > NotSortedArray(j) Then
            v() = NotSortedArray(i)
            NotSortedArray(i) = NotSortedArray(j)
            NotSortedArray(j) = v()
        End If
    Next
Next
Array_Sort = NotSortedArray
End Function
```

Appendix D

Singular Value Decomposition

Singular Value Decomposition (SVD) can be used to calculate the eigenvalues and eigenvectors of the covariance matrix. In this Appendix we follow the methods recommended by Troskie (2000)

Consider any $(n \times p)$ data matrix \mathbf{X} of n observations of p dimensions.

$$\mathbf{X}_{(n \times p)} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \vdots \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

\mathbf{X} could be the returns of a p shares over n time periods.

The SVD of \mathbf{X} is then

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}' \text{ where}$$

$$\mathbf{U}_{(n \times p)} \text{ is orthogonal}$$

$$\mathbf{D}_{(p \times p)} = \text{diag}(d_1, \dots, d_p),$$

$$d_1 \geq d_2 \geq \dots d_p > 0,$$

$$\mathbf{V}_{(p \times p)} \text{ is orthogonal}$$

$$\mathbf{V}'\mathbf{V} = \mathbf{I}.$$

(Note: The \mathbf{U} used above is a matrix and not the same as vector \mathbf{u} used in Equation (6.62).)

Then if we take the SVD of $\mathbf{X}'\mathbf{X}$

$$\mathbf{X}'\mathbf{X} = \mathbf{V}\mathbf{D}\mathbf{U}'\mathbf{U}\mathbf{D}\mathbf{V}' \quad (\text{D.1})$$

$$= \mathbf{V}\mathbf{D}^2\mathbf{V}' \quad (\text{D.2})$$

Notice the similarity between Equation(D.2) and (6.64) . We use this to compute the Principal Components of a sample.

We start off with our data matrix given by

$$\mathbf{X}_{(n \times p)} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \vdots \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

We compute the means of each returns series

$$\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij}, \quad j = 1, \dots, p.$$

We subtract from each element of \mathbf{X} its respective mean i.e.

$$\tilde{\mathbf{X}}_{(n \times p)} = \begin{pmatrix} X_{11} - \bar{X}_1 & \dots & X_{1p} - \bar{X}_p \\ X_{21} - \bar{X}_1 & \dots & X_{2p} - \bar{X}_p \\ \vdots & \vdots & \vdots \\ X_{n1} - \bar{X}_1 & \dots & X_{np} - \bar{X}_p \end{pmatrix}$$

and we call this new matrix $\tilde{\mathbf{X}}$ the *mean adjusted matrix*. We divide each element of this matrix by $\sqrt{n-1}$ assuming we are using the unbiased estimator, for covariance. Using \sqrt{n} would have produced the maximum likelihood estimates.

Thus we have

$$\tilde{\mathbf{X}}_{n-1} = \frac{1}{\sqrt{n-1}} \tilde{\mathbf{X}}$$

and

$$\begin{aligned}\hat{\Sigma} &= \left[\frac{1}{\sqrt{n-1}} \tilde{\mathbf{X}} \right]' \left[\frac{1}{\sqrt{n-1}} \tilde{\mathbf{X}} \right] \\ &= \frac{1}{n-1} \tilde{\mathbf{X}}' \tilde{\mathbf{X}} \\ &= \hat{\Sigma}.\end{aligned}$$

the unbiased estimate of the covariance matrix Σ under multivariate normal theory.

The SVD is then performed on $\tilde{\mathbf{X}}_{n-1}$

$$\tilde{\mathbf{X}}_{n-1} = \frac{1}{\sqrt{n-1}} \tilde{\mathbf{X}}$$

so that

$$\tilde{\mathbf{X}}_{n-1} = \mathbf{U} \mathbf{D} \mathbf{V}'$$

with the result that

$$\begin{aligned}\hat{\Sigma} &= \tilde{\mathbf{X}}_{n-1}' \tilde{\mathbf{X}}_{n-1} \\ &= \mathbf{V} \mathbf{D}^2 \mathbf{V}' \\ &= \mathbf{V} \mathbf{K} \mathbf{V}'\end{aligned}$$

with $\mathbf{V} = \mathbf{B}$.

The vectors of

$$\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \dots, \mathbf{V}_p)$$

are the sample (unbiased) estimates of the eigen vectors and the diagonal elements of

$$\mathbf{K} = \begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k_p \end{pmatrix}$$

are the sample (unbiased) of the eigenvalues.

. The SVD also allows us to give a geometric view of principal components. Suppose our data matrix is

$$\mathbf{X}_{(n \times p)} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ \vdots & \vdots & \vdots \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

It can be thought of as giving the coordinates of n points in a multi-dimensional space of p dimensions. It is possible to represent these n data points in fewer dimensions, minimizing the amount of information lost. And also quantify the amount of information (as variance) lost in the reduced model.

Mathematically speaking we want to find the best space of $r \leq p$ dimensions that is closest to the original space of p dimensions. Here we mean we want to minimize the squared distance from our p dimensional space to this new space r dimensions space. This is equivalent of minimizing the squared residuals from the p space, to r dimensions.

The SVD gives a solution to this problem. Let

$$\begin{aligned} \mathbf{X} &= \mathbf{U}\mathbf{D}\mathbf{V}' \\ &= \mathbf{U}_{[r]}\mathbf{D}_{[r]}\mathbf{V}'_{[r]} + \mathbf{U}_{[p-r]}\mathbf{D}_{[p-r]}\mathbf{V}'_{[p-r]} \\ &= \mathbf{X}_{[r]} + \mathbf{E} \end{aligned}$$

where \mathbf{E} , is the residual then $\mathbf{X}_{[r]}$ minimizes

$$\text{tr}[(\mathbf{X} - \mathbf{A})(\mathbf{X} - \mathbf{A})'] = \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - a_{ij})^2$$

among all other $(n \times p)$ matrices \mathbf{A} of rank r . Thus the singular value decomposition matrix $\mathbf{X}_{[r]}$ can be used as a *matrix approximation* to \mathbf{X} . But this is precisely what principal

components analysis is doing since

$$\text{Total variance} = \text{Explained variance} + \text{Residual}$$

PCA maximizes the explained variance which is equivalent to minimizing the residual.

(Troskie (2000))

University of Cape Town