# Covariance matrix estimation methods for constrained portfolio optimization in a South African setting

**Jaison Pedzisai Madume**



**A mini-dissertation submitted to the faculty of Commerce of the University of Cape Town in partial fulfillment of the requirements for the Degree of Master of Philosophy in Mathematical Finance**

**School of Economics**

**· February 2010 ·**

## Plagiarism declaration form

This research is my original work, produced with normal supervisory assistance from my supervisors. All the relevant sources of knowledge that I have used during the course of writing this mini-dissertation have been fully credited using the Harvard convention for citation and referencing. Also, this mini-dissertation has not been submitted for any academic or examination purpose at any other university.

```
---------------------------                    -----------------
Jaison Pedzisai Madume                         Date
```

# Acknowledgements

My gratitude and appreciation goes to the following people: My supervisors Prof. David Bradfield and Mr. Brian Munro who were tireless in giving me guidance throughout this project, the School of Economics staff for all their assistance during the course of my studies, my family and friends for their support and love.

**Abstract**

One of the major topics of concern in Modern Portfolio Theory is portfolio optimization which is centred on the mean-variance framework. In order for this framework to be implemented, estimated parameters (covariance matrix for the constrained portfolio) are required. The problem with these estimated parameters is that they have to be extracted from historical data based on certain assumptions. Because of the different estimation methods that can be used the parameters thus obtained will suffer either from estimation error or specification error. In order to obtain results that are realistic in the optimization, one needs then to establish covariance matrix estimators that are as good as possible.

This paper explores the various covariance matrix estimation methods in a South African setting focusing on the constrained portfolio. The empirical results show that the Ledoit shrinkage to a constant correlation method, the Principal Component Analysis method and the Portfolio of estimators method all perform as good as the Sample covariance matrix in the Ex-ante period but improve on it slightly in the Ex-post period. However, the improvement is of a small magnitude, as a result the sample covariance matrix can be used in the constrained portfolio optimization in a South African setting.

*Imagination decides everything....*

Blaise Pascal

# Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Background

Harry Markowitz [7] put forward the mean-variance framework which explains that any rational investor would want to maximize his/her portfolio returns whilst minimizing the risk exposure of that portfolio. This mean-variance framework has been popularized as a tool for picking the stocks for a portfolio in the area of portfolio management. In order to estimate and come up with these mean-variance efficient portfolios the investor has to estimate the expected returns for each stock and the covariance matrix of the stock returns. The estimated covariance matrix of the stock returns can be seen as representative of risk control whilst the expected returns are an indication of the expected stock performance. The computation of these efficient portfolios is of major concern to most Finance professionals.

One challenge that arises in the implementation of the Markowitz framework is the estimation of the ingredients mentioned above namely the expected returns and the covariance matrix. Using forecasted stock returns is problematic in that it usually results in portfolio weights that are sensitive to small changes in the stock returns forecasts themselves. It is however possible to consider a portfolio which does not require the expected returns as an input for the computation of the efficient portfolio. At the leftmost tip of the efficient frontier there is a portfolio whose weights are independent of the forecasted expected returns and it is called the Minimum Variance Portfolio. This portfolio has the unique property that it minimizes risk without a given or expected return level.

For the Minimum Variance Portfolio, all we need to estimate is the covariance matrix of the asset returns. However, there are problems that are associated with estimating the covariance matrix for large datasets. The traditional method is to use historical returns data of the stocks to find their sample covariance matrix. This methodology has a number of drawbacks that are associated with it, these being:

- When the number of stocks making up the portfolio is greater than the number of data points (more assets as compared to returns observations) the covariance matrix that is estimated using the traditional method is non-invertible. One of the requirements of the Markowitz framework is that the covariance matrix must be invertible.

- Sampling error occurs when the degrees of freedom per estimated parameter are not sufficient. In this case the number of observations in the sample is not enough as compared to the parameters being estimated for accurate estimation. [5]

- A problem that is associated with trying to impose a particular structure on the tradition covariance model is called the Specification error. In this case the estimator might be too specific than reality. [5]

It is desirable to have an estimated covariance matrix that has reduced estimation error but not too much specification error. However, it is important to note that there is a trade-off between sampling error and specification error (i.e reducing one usually results in an increase of the other). Thus we want to reduce the sampling error whilst maintaining some acceptable structural form in the covariance matrix.

## 1.2 Aims and Objectives

A number of methods have been put forward that try to impose some structure on the estimated covariance matrix. The most popular of these methods being the Ledoit shrinkage to a constant correlation, the Principal Components Analysis based covariance, the Single Index Model based covariance and an equally weighted portfolio of the Sample covariance, the Diagonal covariance and the Single Index Model based covariance. These methods have been tested extensively on USA's portfolios. Christoffer Bengtsson and Jan Holst [3] also investigated these methods and their modifications on the much smaller Swedish market. However, not much research has been done on the South African Market portfolios.

The aim of this paper is to investigate the above mentioned methods in a South African setting and establish the best method of covariance matrix estimation using the Shareholder Weighted Index (SWIX) as the stock universe of interest. All the portfolios constructed are based on the constrained minimum variance portfolio.

## 1.3 Layout

Section 2 (Literature Review) gives a review of the Markowitz framework, optimization and the various methods used for covariance estimation. Section 3 (Data and Methodology) reviews the implementation and data used in the analysis. Section 4 (Results) gives the empirical findings and their analysis. Section 5 (Conclusions and Recommendations) summarises the findings of this paper and suggests areas of further study.

# 2 Literature Review

This section gives a brief review of the existing literature on the framework and inputs to the portfolio optimization problem. In this section boldface capital letters refer to matrices and capital letters in italics represent vectors.

## 2.1 Portfolio Optimization

Consider a portfolio of N assets with weights given by $W$ ((Nx1) vector) and covariance matrix of the assets given by $\mathbf{\Sigma}$ (NxN matrix). The $i^{th}$ entry in $W$ represents the weight of asset $i$ in the portfolio. The $i^{th}$ diagonal element of $\mathbf{\Sigma}$ is the variance of the returns of asset $i$, whilst the off diagonal $(ij)^{th}$ elements are the covariances of asset $i$ and asset $j$. If the returns of the assets are $\mu = (E(r_1), ..., E(r_N)))^T$ [1](where $E(r_i)$, the return of the portfolio is then given by $W^T\mu$ denotes the expected return of asset $i$).The risk exposure of this portfolio is given by $W^T\mathbf{\Sigma}W$. The optimization problem for the minimum variance portfolio is then defined as:[4]

$$\min_{W} \frac{1}{2}W^T\mathbf{\Sigma}W \quad \text{subject to} \quad W^T\mathbf{1} = 1 \tag{1}$$

where $\mathbf{1}$ represents a Nx1 vector of ones.

The above is the unconstrained optimization problem in which the weights of the resultant portfolio are not restricted to being between 0 and 1 and it represents a situation in which short selling is allowed as weights can be negative.

[1]The superscript T refers to the transpose

The constrained problem has the additional condition that the weight for asset $i$ must be non-negative and less than one. The setup then becomes:

$$\min_{W} \frac{1}{2} W^T \mathbf{\Sigma} W \quad \text{subject to} \quad \begin{aligned} & W^T \mathbf{1} = 1 \\ & 0 \leq W_i \leq 1 \end{aligned} \tag{2}$$

In this scenario short selling is not allowed as weights cannot be negative. This is more realistic as usually portfolio managers are prohibited from short selling (for example in mutual funds) and many investors find it difficult to short sell.



## 2.2 Covariance Matrix Estimation Methods

In this section a brief discussion of the various covariance estimation methods is given.

### 2.2.1 Sample Covariance Matrix

Suppose that $\mathbf{R}$ represents the (NxT) matrix of asset returns with T time points and N assets present in the portfolio. Define $\bar{R}$ as the N×1 mean vector of $\mathbf{R}$ so that $\bar{R} = \frac{1}{T}\mathbf{R}\mathbf{1}$ (where $\mathbf{1}$ is an N×1 vector of ones). Let $\mathbf{X}$ be the N×T matrix of de-meaned asset returns, that is each column of $\mathbf{X}$ is defined as $R_i - \bar{R}$ ($R_i$ represents the $i^{th}$ column of $\mathbf{R}$ for $i = 1, 2, ..., T$). The Sample covariance is then defined as:

$$\mathbf{\Omega} = \frac{1}{T}(\mathbf{X}\mathbf{X}^T) \tag{3}$$

This estimator will also be called the Sample in this paper.

The Sample covariance matrix is the Maximum Likelihood estimator under the assumption of normality, which means it is the best unbiased estimator. However, for this assumption we put all of our confidence in the data which may not necessarily be sufficient. As a result we might over-fit the data and get an estimate that might perform well in-sample but performs poorly out-of sample [3]. Another issue is that we have to estimate a large number of parameters. A total of N(N+1)/2 parameters have to be estimated, for example if we consider the SWIX which has an average of 165 assets we need to estimate 13,695 parameters. To rectify this we need a lot of data which is usually not available. It is also important to note that if the number of assets $N > T$ then the sample covariance matrix is non-invertible.

### 2.2.2 Diagonal Covariance Matrix

The Diagonal covariance matrix is essentially the sample covariance matrix with all the off-diagonal elements being zeros. An assumption that the all the stocks are uncorrelated is made and results in a high degree of specification error.



### 2.2.3 Ledoit Shrinkage to a Constant Correlation Covariance Matrix

This estimator is introduced when the number of assets under consideration is large and hence estimation error might be significant. The idea is to shrink the sample covariance matrix towards a more structurally stable estimator like the constant correlation matrix. The principles used in the estimation of the new covariance matrix are based on Bayesian Shrinkage of a given matrix towards an assumed prior matrix and the formula is given as:

$$\mathbf{\Omega_{BS}} = \lambda \mathbf{\Omega}_{\text{prior}} + (1 - \lambda)\mathbf{\Omega} \tag{4}$$

For the Ledoit shrinkage to a constant correlation matrix (also referred to as Ledoit in this paper) the prior $\mathbf{\Omega}_{\text{prior}}$ is assumed to the constant correlation matrix [9] and $\mathbf{\Omega}$ is the Sample covariance matrix. $\mathbf{\Omega_{BS}}$ is the resultant covariance estimator. There is however a need for us to be able to estimate $\lambda$ so that it gives a result that is an optimal weighted sum of the two inputs. This optimal estimate is a function of the spread between variances and covariances and is given by:

$$\lambda = \frac{\text{SUM}(\text{SQ}(\mathbf{R})\text{SQ}(\mathbf{R})^T) - \frac{\text{SUM}(\text{SQ}(\mathbf{\Omega}))}{T}}{\text{SUM}(\text{SQ}(\mathbf{\Omega} - \mathbf{\Omega_{prior}})}[10] \tag{5}$$

where SQ() denotes the element-by-element squaring function for a matrix argument. SUM () is the sum of the argument matrix elements. For the implementation by Ledoit and Wolf [9] $\mathbf{R}$ is considered as the matrix of demeaned returns and the Matlab code was adapted from the one provided on Ledoit's website [9].



### 2.2.4 Constant Correlation Matrix

Let S denote the sample covariance matrix as defined in section 2.2.1 with entries $s_{ij}$, then the sample correlations between the returns on stocks i and j are given by:

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}$$

The average sample correlation is then given by [9]:

$$\bar{r} = \frac{2}{(N-1)N} \sum_{i=1}^{N-1} \sum_{j=(i+1)}^{N} r_{ij}$$

We define the sample constant correlation matrix $\mathbf{Q}$ by means of the sample variances and the average sample correlation:

$$q_{ii} = s_{ii} \qquad and \qquad q_{ij} = \bar{r}\sqrt{s_{ii}s_{jj}}$$

$\mathbf{Q}$ is then used as $\mathbf{\Omega}_{prior}$ in equation 5.

### 2.2.5 Principal Components Analysis Covariance

Principal component analysis (PCA) aims to explain the variation structure using a few linear combinations of the original stochastic variables. It is quite useful as it achieves data reduction and interpretability at the same time revealing some relationships that are not so obvious but might prove to be critical. This method of analysis argues that for any N×N covariance matrix $\mathbf{S}$, there exists $K < N$ principal components that can be used to explain the variation of the entire covariance matrix without losing much information.

The principal components for the sample are linear combinations

$$B_i = H_i^T = \sum_{m=1}^{N} H_{mi} R_m \quad i = 1, 2, ..., N \tag{6}$$

such that the variances and covariances are:

$$\text{Var}(B_i) = H_i^T \mathbf{S} H_i \tag{7}$$

$$\text{Cov}(B_i, B_j) = H_i^T \mathbf{S} H_j \qquad i, j = 1, 2, ..., N \tag{8}$$

of the N stochastic variables. These principal components should give the direction of the maximum variance in such a way that the principal components themselves are uncorrelated to each other and $H_i^T H_i$=1.

It is possible to find the optimal weight functions for the covariance without solving a number of optimizations. A proposition is given without proof: [3]

**Proposition 1** *For any covariance matrix* $\mathbf{S}$ *with eigenvalue-eigenvector pairs* $((\lambda_1, V_1), ..., (\lambda_N, V_N))$ *such that* $\lambda_1 \geq \lambda_2 \geq, ..., \lambda_N \geq 0$ *with* $V_k = (V_{1k}, ..., V_{Nk})$ *then the* $i^{th}$ *principal component is defined as:*

$$B_i = V_i^T R = \sum_{m=1}^{N} V_{mi} R_m \quad i = 1, 2, ..., N \tag{9}$$

*where* $R$ *is the* $N \times 1$ *vector of stochastic observations with covariance matrix* $\mathbf{S}$. *Such that:*

$$Var(B_i) = V_i^T \mathbf{S} V_i = \lambda_i \tag{10}$$

$$Cov(B_i, b_j) = V_i^T \mathbf{S} V_j = 0 \tag{11}$$

A critical question is how then do we determine the eigenvalues to be significant or not. Define $\mathbf{D}$ as the matrix whose diagonal elements are sample variances and the off-diagonal elements are zeros. $\mathbf{X}$ is the N×T matrix of de-meaned asset returns for the universe of assets. Then we can define a new input into the principal components algorithm as:

$$\mathbf{P} = \mathbf{D^{1/2}X} \tag{12}$$

$$\mathbf{S_p} = \frac{1}{T}\mathbf{PP^T} \tag{13}$$

The principal components of $\mathbf{S_P}$ are defined in the same way as those of the Sample covariance matrix using proposition 1. Bengtsson and Holst[3] use the rules defined in econophysics [11] to select the significant principal components. They define V=T/N and choose the principal components whose eigenvalues deviate significantly from $\lambda_{max}$ defined as:

$$\lambda_{min}^{max} = 1 + \frac{1}{V} \pm 2\sqrt{\frac{1}{V}} \tag{14}$$

In this study, the principal components are determined from the T×T stock cross return product matrix $\mathbf{X^TX}$. Because of its symmetry it is a type of covariance matrix. This is transformed using the method of Bengtsson and Holst described above to obtain the K significant eigenvalues. After the significant eigenvalues K have been chosen, select the corresponding principal components to come up with a K×T matrix $\mathbf{F}$ of the principal components. The next step is to regress this matrix on the original de-meaned data $\mathbf{X}$ such that [10]:

$$\mathbf{C} = (\mathbf{FF^T})^{-\mathbf{1}}(\mathbf{FX^T}) \tag{15}$$

$$\mathbf{Z} = \mathbf{X} - \mathbf{C^TF} \tag{16}$$

where $\mathbf{C}$ is a K×N of factor exposures for the stocks and $\mathbf{Z}$ is the residual returns matrix.
The covariance matrix that is based on the first K significant principal components is then defined as:

$$\mathbf{\Omega} = \frac{1}{T}\mathbf{C}^T(\mathbf{F}\mathbf{F}^T)\mathbf{C} + \frac{1}{T}\mathbf{V} \tag{17}$$

Where $\mathbf{V}$ is a matrix that has its diagonal elements equal to the diagonal elements of $\mathbf{Z}\mathbf{Z}^{\mathbf{T}}$ and all other elements as zero. Multiplication with 1/T here is made so that the resultant covariance matrix is comparable to the sample covariance matrix defined in equation 3.



### 2.2.6 Single Index Model Covariance

Sharpe's Single Index Model (also referred to as Single Index in this paper) assumes the market to be the only factor that is significant in determining asset returns and has been popularized as the Capital Asset Pricing Model (CAPM). This model can be expressed for a single asset as:

$$r_{it} = \alpha_i + \beta_i R_{Mt} + \epsilon_{it} \tag{18}$$

Considering a portfolio with N assets, we can rewrite this in matrix form as:

$$\mathbf{R_t} = \alpha + \beta R_{Mt} + \epsilon_{\mathbf{t}} \tag{19}$$

where $\epsilon_{\mathbf{t}}$ is the N×1 vector containing the mean zero uncorrelated residuals $\epsilon_{it}$. Then the covariance matrix implied by this market model is:

$$\mathbf{\Omega} = \sigma_M^2 \beta \beta^T + \mathbf{\Omega}_\epsilon \tag{20}$$

where $\sigma_M^2$ is the market portfolio variance.

Using the available data to estimated the covariance matrix gives us:

$$\hat{\mathbf{\Omega}} = \hat{\sigma}_M^2 \hat{\beta} \hat{\beta}^T + \hat{\mathbf{\Omega}}_\epsilon \tag{21}$$

This is obtained by regressing the portfolio returns on the market returns. The significance of this model is that it adds a lot of structure as it assumes only one factor (the market) to be significant and it reduces greatly the parameters to be estimated. In total we need only estimate 2N+1 parameters as compared to N(N+1)/2 for the sample covariance matrix. For example the SWIX which has an average of 165 stocks, we would need to only estimate 331 parameters as opposed to 13,695 for the sample covariance. This represents a great reduction in parameters to be estimated and as such the estimation error. However, it is also highly unreasonable to assume that only one factor adequately explains asset returns, hence, this model will introduce some specification error.



### 2.2.7 Portfolio of Estimators Covariance

Different methods give errors in different directions, we can weight different methods to come up with one estimator in which it is hoped the errors cancel out. Bengtsson and Holst [3] define an equally weighted portfolio of the Sample, Diagonal and Single Index covariances.

Let **S** represent the Sample covariance, **D** the diagonal covariance and **SI** the Single Index covariance, then the Portfolio of estimators is given by:

$$\mathbf{\Omega} = \frac{1}{3}\mathbf{S} + \frac{1}{3}\mathbf{D} + \frac{1}{3}\mathbf{SI} \tag{22}$$

## 2.3 Portfolio Analysis

There are a number of ways that can be used to analyze the performance of a portfolio. A review of some of those relevant to this paper are given in the sections that follow.

### 2.3.1 Concentration

Portfolio diversification can be assessed based on its concentration measure. The Herfindahl-Hirschman Index (HHI) is a measure of concentration that is obtained by first squaring the investment weights in a portfolio and then summing them. Consider a portfolio of N assets whose weights are represented by an N×1 vector $W$ so that the weight of the $i^{th}$ asset is given by $W_i$ $i = 1, 2, ..., N$. The HHI is given as:

$$\text{HHI} = \sum_{i=1}^{N} W_i^2 \tag{23}$$

From the above formulation of concentration we note that HHI will be a minimum for an equally weighted portfolio and gets larger as the weights are skewed (concentrated). If the portfolio is made up of only one stock then the HHI will be unity which is its maximum value. Portfolios with smaller measures of concentration are considered to be more diversified and hence less risky than portfolios with large concentrations.

### 2.3.2 Beta

The beta of a stock measures how sensitive the returns of that particular stock are to the returns of the market as a whole. In other words it measures the exposure of a particular stock to market risk. A positive beta implies that the stock follows the market movements and negative beta implies the stock follows the market inversely. A beta of zero means the stock is not sensitive to any market movements.

For any stock $i$ with returns history $X_i = (x_{1i}, x_{2i}, ..., x_{Ti})$ and a market returns history $Y = (y_1, y_2, ..., y_T)$, the beta of stock $i$ is given as:

$$\beta_i = \frac{\sum_{k=1}^{T}(y_k - \bar{y})(x_{ki} - \bar{x_i})}{\sum_{k=1}^{T}(x_{ki} - \bar{x_i})^2} \tag{24}$$

Now let $\beta$ be the N×1 vector whose $i^{th}$ element is $\beta_i$, the beta of asset $i$. Then the beta of the portfolio with weight vector $W$ is given by:

$$\beta_{\mathbf{p}} = \mathbf{W^T}\beta \tag{25}$$

Like the beta of a stock the beta of a portfolio measures the exposure of the portfolio to market risk. Risky portfolios tend to have a high beta value.

### 2.3.3 Total Risk

Total Risk of a portfolio also provides a measure of portfolio performance as it measures its risk exposure. For a portfolio with weight vector $W$ and covariance matrix $\mathbf{\Sigma}$ the Total Risk is given as:

$$\sigma^2_{\text{total}} = W^T \mathbf{\Sigma} W \tag{26}$$

This gives the in-sample risk of the portfolio. The above result is multiplied by 52 to get the annualized Total Risk. Usually risk is reported in terms of standard deviation so that we record:

$$\sigma_{\text{total}} = \sqrt{\mathbf{W}^T \mathbf{\Sigma} \mathbf{W} \times 52} \tag{27}$$

### 2.3.4 Tracking Error

The tracking error measures how much a portfolio follows an index to which it is benchmarked.

For a portfolio of weights $W_p$ invested on a market with benchmark weights $W_b$ the portfolios out-performance over the benchmark is calculated as $x = W_p - W_b$. The risk of these active weights is then the Tracking error and is given by:

$$\sigma_{\text{Tracking error}} = \sqrt{x^T \mathbf{\Sigma} x \times 52} \tag{28}$$

where multiplication by 52 is to annualize the result when using weekly data.

The tracking error of a portfolio that perfectly follows an index to which it is benchmarked is zero. The benchmark used in this paper is the SWIX.

# 3 Data and Methodology

## 3.1 Data

The datasets for this project were extracted from the Shareholder Weighted Index (SWIX) from 31 December 2006 through to 31 July 2009 and from weekly stock returns from 5 October 2003 to 31 August 2009. Weekly data was used as it allowed each stock that is present in the SWIX at any given date to have a 170 time point return history (about three and a half years of weekly data). Choossing 170 data points means that the Sample covariance matrix will be invertible as the SWIX usually has 165 stocks.

## 3.2 Methodology

The procedure of analysis is given below:

1. At the end of each month (31 December 2006 to 31 July 2009), extract the stocks that are present in the SWIX.

2. Retrieve the weekly return data for each stock present in the SWIX such that each stock has a 170 data points return history.

3. Determine the covariance matrix for the selected return data using a particular covariance estimation method.

4. Feed the covariance matrix to the optimizer for the constrained Minimum Variance Portfolio.

5. Retain the portfolio weights of the resultant minimum variance efficient portfolio.

6. Hold this portfolio for an out-of-sample period of one month storing the weekly returns for that month. This is because weekly data is used. Each month will then have four returns for the portfolio strategy.

7. Repeat the procedure for the end of each month to come up with a series of 128 weekly returns for a particular portfolio strategy.

8. The standard deviation of this series is the critical metric, record it for each of the different covariance methods used.

9. Also for each month record the concentration, portfolio beta and total risk of each covariance estimation method. The total risk is calculated assuming the sample covariance as the true covariance for comparison purposes.

10. Repeat this process for all covariance estimation methods under consideration.

11. Retain the out-of-sample returns of the market portfolio (SWIX) for the 128 weeks and then regress the returns of each method on these market returns to get the overall beta of the method through the 128 weeks. The market returns are obtained by assuming the SWIX weights are also held through the out-of-sample period.

12. For comparison purposes also record the concentration and total risk of the SWIX portfolio.

13. Implement the naive investment strategy in which an equal weighting is given to all the stocks present in the portfolio at any time.

14. For the Ledoit shrinkage method record the shrinkage parameter used at each date to see how much the sample covariance was shrunk.

15. For the Principal Components Analysis method analyze how many principal components would be significant explaining most of the variation in the data. The PCA covariance is then constructed based on these significant components.

# 4 Results

## 4.1 Ex-ante Analysis

In this section an analysis of the resultant optimal portfolios within the sample is given.

### 4.1.1 Ledoit Shrinkages

Figure 1 gives the shrinkage parameters that were used for the Ledoit covariance matrix at each date. From December 2006 up to October 2008 this shrinkage parameter is greater than 0.4 but drops to above 0.25 from November 2008 until July 2009. These parameters show that the sample covariance matrix is shrunk quite significantly towards the constant correlation matrix hence a reasonable degree of structure is imposed. However, in-sample results discussed later show that it still performs just as the Sample covariance.



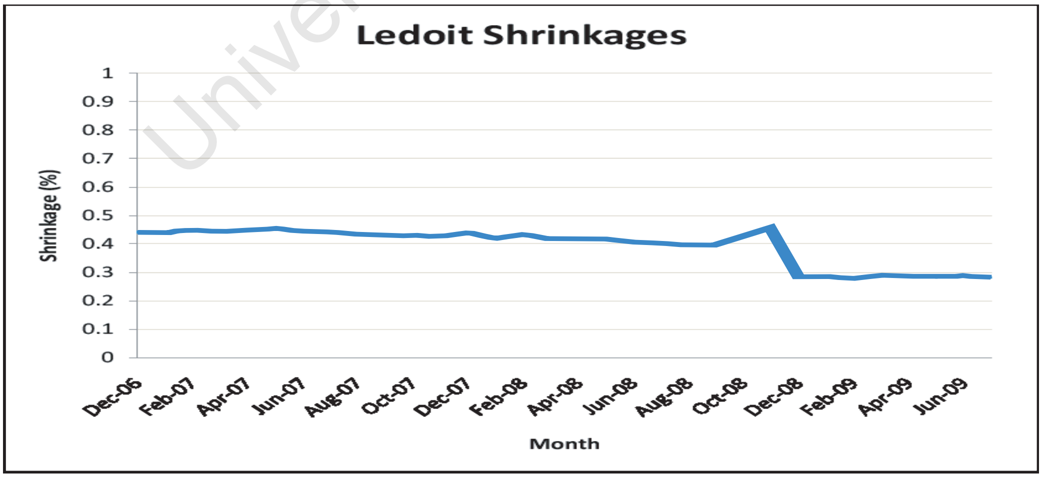


Figure 1: Ledoit Shrinkages

### 4.1.2 Principal Components Analysis

Figure 2 shows the contribution to the total variation of the first few components. The number of significant components in the covariance estimation has a lowest of 3 and and a highest of 7 and these explain between 71% and 81% of the total variation. Components can be seen to be predominatly 5 or 6 for the SWIX as shown in Figure 2, this is consistent with literature which suggests that usually the first 5 components are sufficient for explaining most of the variation [10].

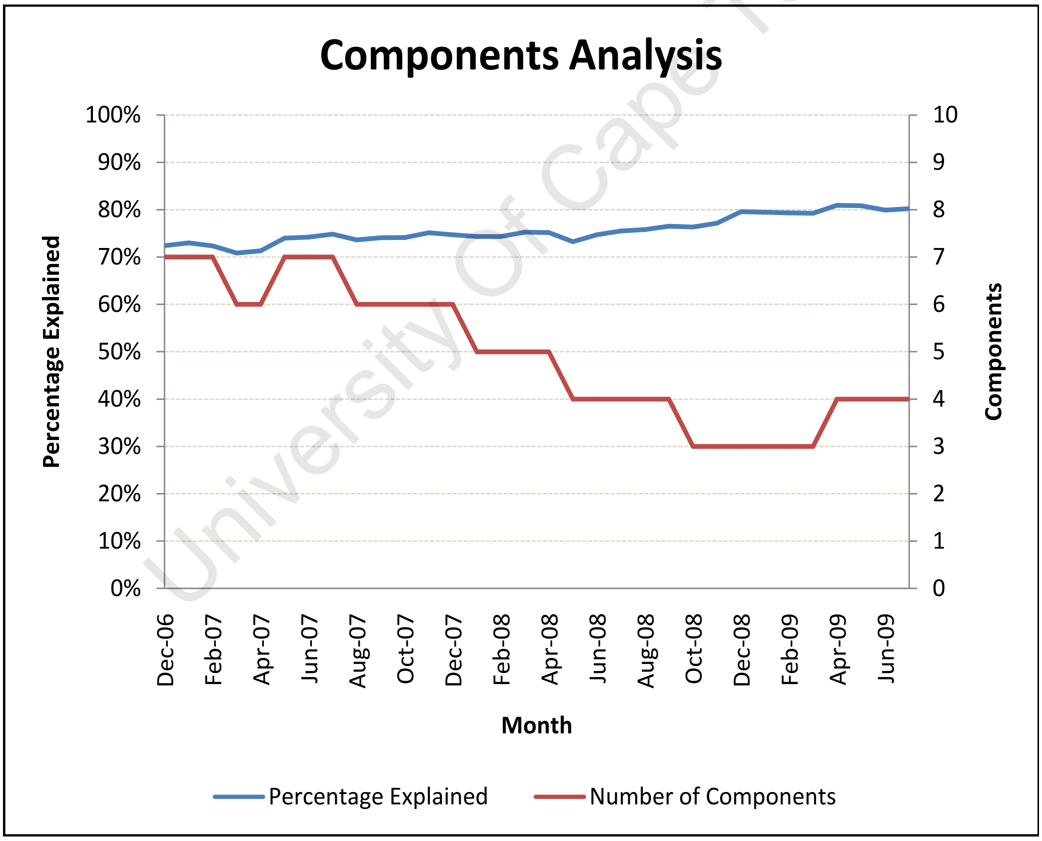


Figure 2: Principal Components Analysis

#### 4.1.3 Total Risk

Figure 3 shows that the benchmark SWIX has the highest ex-ante total risk when we assume the sample covariance matrix as the true covariance matrix. The naive method has the second highest total risk and the Diagonal method is third with movements similar to the naive approach. All the other methods (Ledoit, PCA, Sample and the Portfolio of estimators) have total risk that is almost identical through time. It is therefore difficult to separate the better estimator using the total risk measure. The Single Index Model was not recorded because its total risk gets so large because it places large weights on a few stocks and is not plotted here.



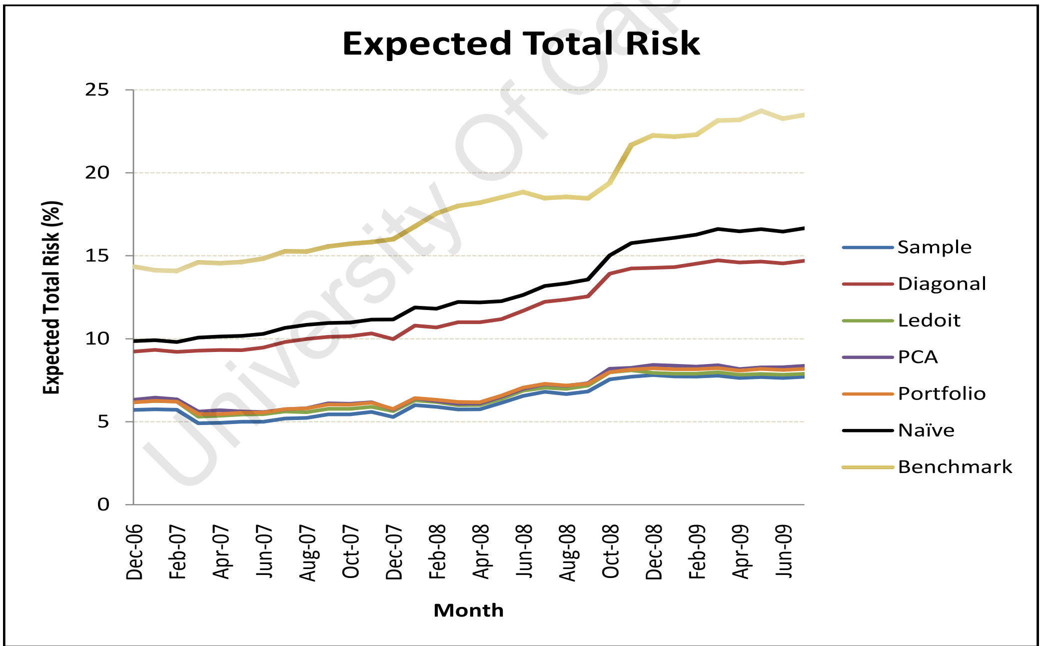


Figure 3: Total Risk

The Naive investment strategy gives equal weighting to all the stocks present in the portfolio at any time as explained in section 3.

#### 4.1.4 Tracking Error

The tracking error is closely related to risk. It gives the risk that a portfolio is exposed to by deviating from the benchmark portfolio. Figure 4 shows that the naive approach has the lowest tracking error followed by the Diagonal method. This is consistent with what we have observed for the total risk that these two move close to each other. The other four methods (Sample, Ledoit, PCA and Portfolio of estimators) all have tracking errors that are virtually superimposed over each other. Again it proves to be difficult to separate these methods using tracking error. The Single Index Model is left out because its values are too large and thus are outliers.

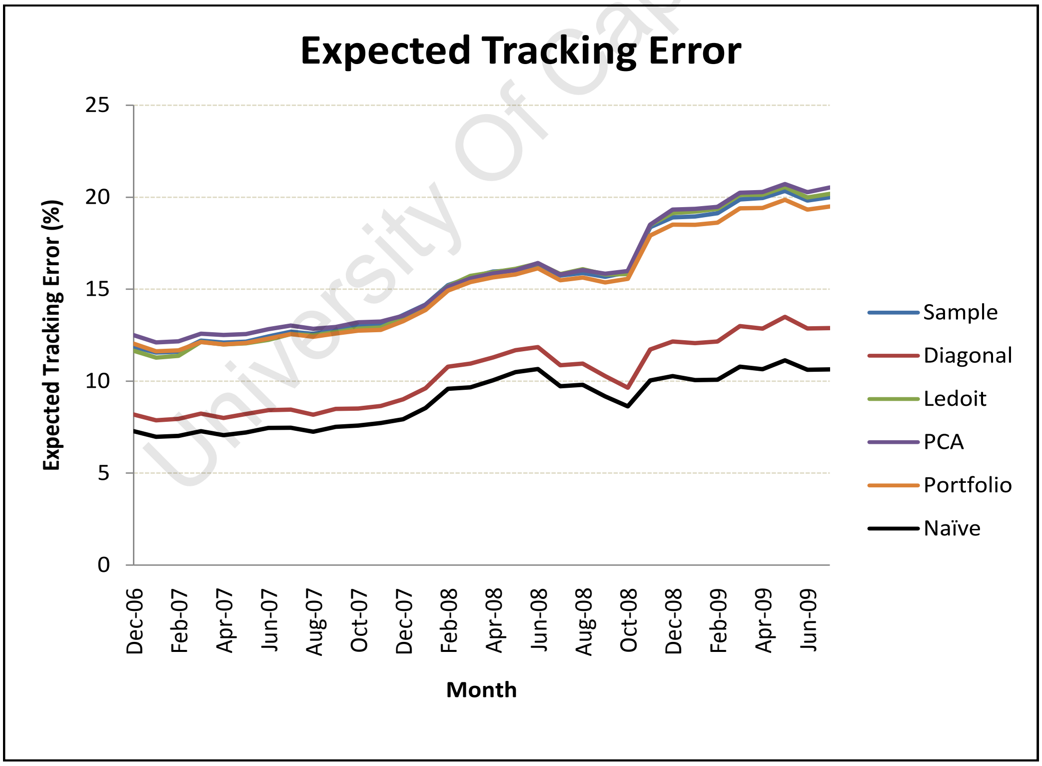


Figure 4: Tracking Error

### 4.1.5 Portfolio Sizes

Table 1 shows the maximum and minimum portfolio sizes of the optimal portfolios generated by each covariance estimation method as well as for the entire market (SWIX) extracted from the equally weighted portfolio. It can be seen from this table that the Single Index Model based covariance produces the smallest portfolio sizes with a minimum of just one stock and a maximum of three stock. This explains why this method produces risky portfolios as compared to all the other methods as it gives virtually no diversification at all (diversification is at the centre of risk reduction). The Diagonal matrix uses the same number of stocks as the naive portfolio which uses all the stocks present in the SWIX. As such it will be observed that because they use a lot of stocks these methods will produce the smallest portfolio weights which result in small concentration measures. The Sample has a minimum of 25 and a maximum of 43, Ledoit a minimum of 27 and a maximum of 50, PCA a minimum of 35 and a maximum of 53. These three methods all have performance patterns that are almost similar in-sample. The Portfolio estimators which is the best performing method in our contest has a minimum portfolio size of 46 stocks and a maximum of 74 stocks. This can be seen as the difference in performance of the methods, as the other methods namely the Sample, Ledoit and PCA compromise on diversification by using a smaller number of stocks. A well diversified portfolio requires a reasonable number of stocks to be present. The Portfolio of estimators uses a number of stocks that gives diversification without introducing too much specification or estimation error.

| Method | Minimum | Maximum |
|---|---|---|
| Sample | 25 | 43 |
| Diagonal | 159 | 167 |
| Ledoit | 27 | 50 |
| PCA | 35 | 53 |
| Single Index | 1 | 3 |
| Portfolio | 46 | 74 |
| Equal Weights (Naive) | 159 | 167 |

Table 1: Portfolio Sizes

#### 4.1.6 Market Exposure

Figure 5 shows the betas of the optimal portfolios with respect to the market. The naive approach has the highest risk as measured by its betas. Again the Diagonal covariance has the second highest risk as measured relative to the market. The Single Index Model method has the lowest betas as a result of how it is estimated based on the betas of the stocks. The optimiser picks the stocks with the lowest betas and as such the resultant portfolio will have low market exposure as measured by the beta. The Ledoit, PCA and the Sample methods have betas that are almost superimposed on each other. The Portfolio of estimators has higher betas than these three but is very close to them and mimics their movements. In terms of exposure to the market measurement the best performing is the Single Index Model and the worst performing is the Diagonal method.

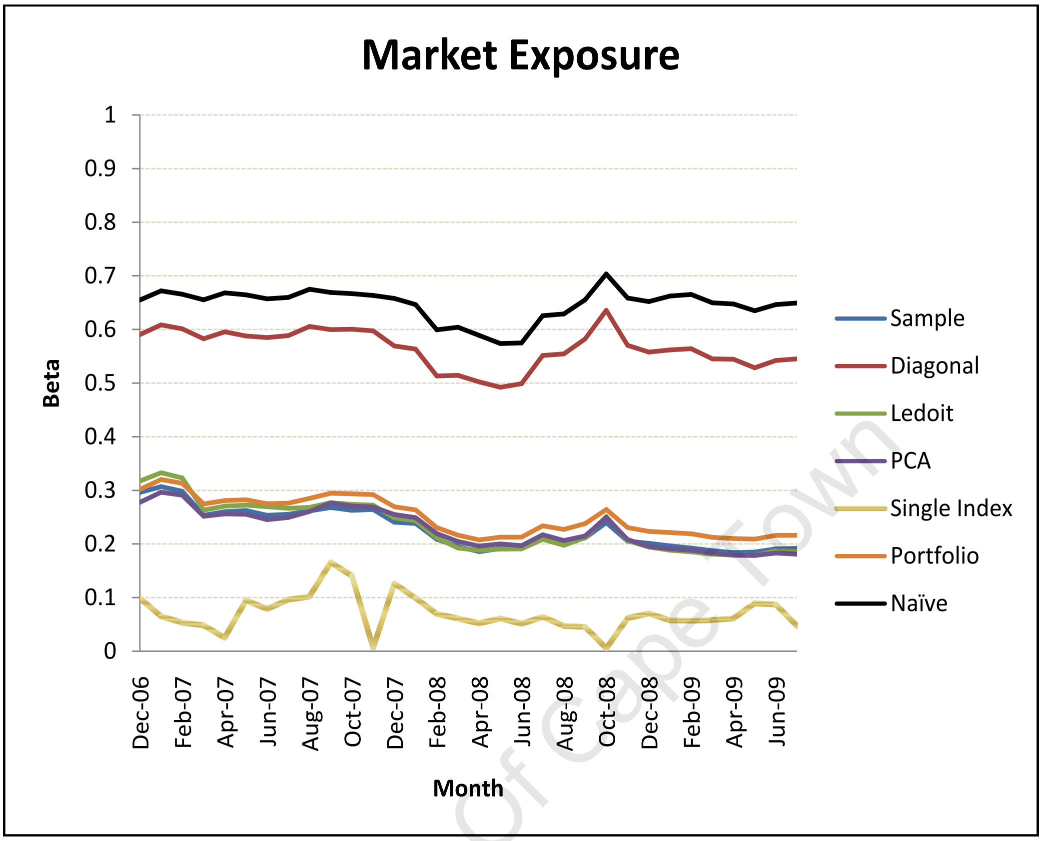


Figure 5: Market Exposures(Betas)

### 4.1.7 Concentration

Figure 6 shows that the Single Index Model has the highest concentration which can be as large as unity. The optimizer looks for the stocks with the lowest beta (market exposure) and places large weights in these few stocks. Using this method can lead to a portfolio that has only one stock (concentration of unity). All the other estimation methods have a concentration that is smaller than 0.2. The Ledoit has the second highest concentration and the Diagonal has the lowest. The naive and Diagonal concentrations are superimposed over each other. It is important to note that the smaller the weights the lower the concentration, hence because of their concentration

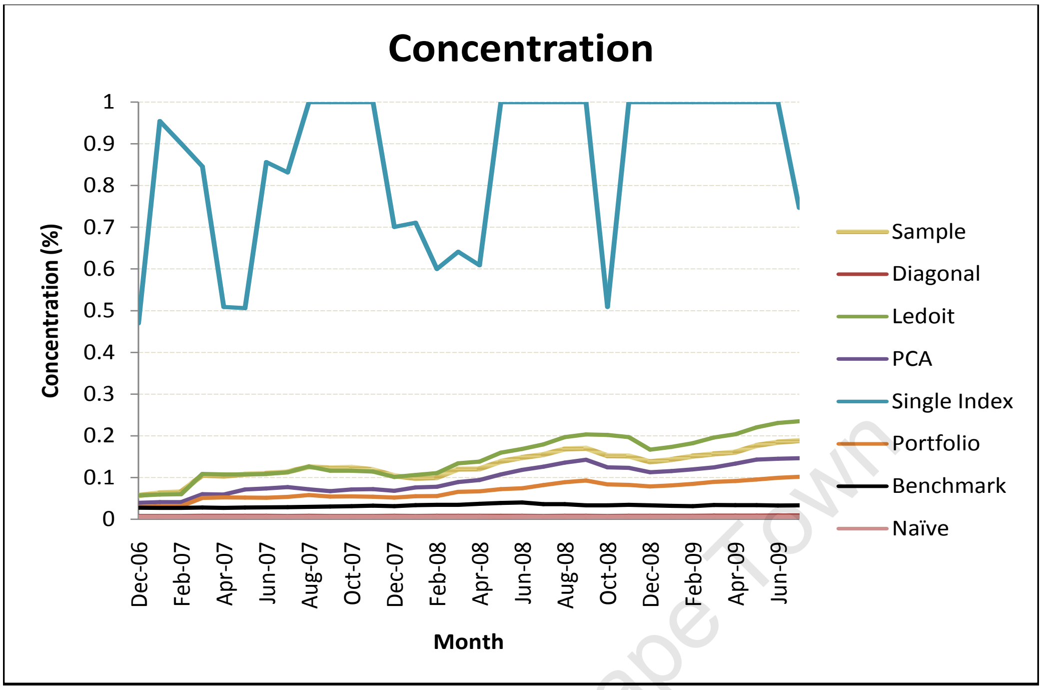


Figure 6: Concentration

the Diagonal method and the naive approach produce the smallest weights. This is consistent with Bengtsson and Holst [3] who record that the Diagonal method has the smallest portfolio weights. However, we should also realize that the Diagonal matrix is based on a method that has high levels of specification error and hence the portfolios it produces are based on the flawed assumption that the stocks are uncorrelated. The Portfolio of estimators has the third smallest concentration after the benchmark SWIX itself. This essentially means that in terms of concentration no estimation method can be used for better portfolio determination over the SWIX. But for the methods in our contest the Portfolio of estimators performs the best after the Diagonal method in terms of concentration. The Sample covariance also performs better than the Ledoit shrinkage in this regard. Confirming some of the arguments in the literature that the sample covariance might actually be a good consideration in-sample.

#### 4.1.8 Discussion

The naive approach and the Diagonal method seem to achieve diversification as they have the lowest concentrations, however, these have high risk associated with them as measured by the total risk and the betas. For all the other methods it is difficult to separate them based on in-sample performance in terms of risk reduction. Even the Sample covariance method itself has properties that are similar to the contestants and outperforms some of them marginally in-sample. In terms of total risk, the Sample covariance is better than all the other methods. For market exposure it out performs the Portfolio of estimators and is very similar to the PCA and the Ledoit estimators. When portfolio concentration is considered it is better than the Ledoit estimator. This means that in-sample even the Sample covariance estimator can be considered and will outperform other methods. In order to establish the better estimation method an analysis of the covariance estimators Ex-post is given in the next section.



## 4.2 Ex-post Analysis

Here the out of sample performance of each of the covariance estimation methods is given. The standard deviation of the returns and the beta of each estimation method are recored in table 2.

### 4.2.1 Risk and Market Exposure

The results in table 2 show that the Single Index Model which assumes the SWIX as the market has the highest risk (standard deviation) of 34.34 (annualised). Large weights are placed into a few stocks, there is not much diversification as a result the risk is high. The SWIX has the second highest out of sample standard deviation (risk) of 22.94, it performs better than the Single Index Model because it introduces diversification and invests in more than a few stocks (3 for the Single Index). The naive approach has the third highest risk measure of 16.26. This method does not consider covariance estimates at all but reduces portfolio concentration using equal weights for all the stocks. The Diagonal is the fourth highest with a risk of 14.93, it assumes that all the stocks in the portfolio are uncorrelated and hence has a high level of specification error. The Sample covariance matrix is outperformed by the Ledoit, PCA and the Portfolio of estimators. It has the fifth highest risk measure of 10.55. This is expected as the Sample covariance matrix has a high degree of estimation error caused by the large number of parameters that have to be estimated. As a result the optimizer places the biggest weights in these estimated errors (this is the problem of error maximization). The sixth highest standard deviation is for the Ledoit method whose risk measure is 10.46. This is the third lowest standard deviation in the estimation contest. The shrinkage parameters have a high of 0.457 and a low of 0.278, this shows that the sample covariance matrix is shrunk significantly towards the constant correlation matrix and hence a lot of structure is given to the resultant covariance estimate. The PCA has the second best out-of-sample standard deviation of 10.33. This variance reduction was achieved using between 3 and 7 principal components only as shown in section 4.1.2. The best performing covariance estimation method in the Out-of-Sample period is the

Portfolio of estimators with a risk estimate of 10.10. This is consistent with what was reported by Bengtsson and Holst [3] in their contest. The returns of each estimation method were regressed on the returns of the market and the betas were recorded. As is shown in table 2 the diagonal method has the highest risk as measured relative to the SWIX with a beta of 0.55. The Sample, Ledoit, PCA and Portfolio of estimators have betas that are close to each other in the range 0.27 to 0.29 hence their market risk exposures are almost similar with a difference of 0.02 between the highest and lowest exposures. The beta of the Ledoit and the PCA are the same although their risks are different, so they have the same exposure to the market. The Single Index Model has the lowest beta of 0.27 as minimizing the risk will amount to choosing the stocks that have the lowest betas because of the way the Single Index Model covariance matrix is defined.

| **Method** | **Annualised Risk** | **Beta** |
|---|---|---|
| Single Index | 34.34 | 0.27 |
| SWIX | 22.94 | 1 |
| Naive | 16.26 | 0.63 |
| Diagonal | 14.93 | 0.55 |
| Sample | 10.55 | 0.29 |
| Ledoit | 10.46 | 0.28 |
| PCA | 10.33 | 0.28 |
| Portfolio | 10.10 | 0.30 |

Table 2: Ex-post Analysis

### 4.2.2 Discussion

In the out-of-sample analysis, the best performing estimator in terms of risk reduction is the Portfolio of estimators followed by the PCA and the Ledoit. The difference in risk between the Sample and Portfolio of estimators is 0.45. There is not much difference in terms of market exposure between the Single Index, Sample, Ledoit, PCA and the Portfolio of estimators. The difference in the betas between the largest and the smallest is just 0.02. The Portfolio of estimators reduces the risk from 10.55 for the sample to 10.10 a difference of 0.4. These results suggest that the Ledoit, PCA and Portfolio of estimators do reduce the risk of the optimal portfolio although the difference might not be significant.

# 5 Conclusions and Recommendations

## 5.1 Conclusions

The purpose of these paper was to investigate the various covariance matrix estimation methods for the constrained portfolio optimization problem in a South African setting. The primary reason being that it is one of the two required inputs in the Markowitz mean-variance optimization problem. The portfolios considered are minimum variance portfolios which do not require an input of the expected returns which usually prove to be difficult to generate and gives results that are very sensitive to any small movements in the estimated expected returns. The estimation methods considered are the Ledoit shrinkage to a constant correlation (Ledoit) proposed by Ledoit and Wolf [9], Principal Components Analysis (PCA), Single Index Model, Diagonal and a Portfolio of estimators. These were all put in to a horse race with the Sample covariance matrix.

The number of components used for the principal components method was between 3 and 7 and this explains most of the variation present (71% and above). This is in line with literature which argues that usually the first five components are significant to explain most of the variation. The Ledoit shrinkage towards the constant correlation matrix proved to have quite a significant impact as the shrinkage parameters used were quite high. However, its performance especially in the in-sample period did not deviate much from the Sample covariance. The Portfolio of estimators can be seen as one that diversifies the errors of the different estimation techniques. Different estimators give errors in different directions, hence, when they are combined the error is reduced significantly. The sample estimator has a lot of estimation error whilst the other methods (the Diagonal and the Single Index Model) have specification error and averaging these produces a better estimator.

This paper reveals that under the constrained model, in the Ex-ante period it is really difficult to pick exactly one method that can be said to perform better than the others especially for the Sample, Ledoit, PCA and the Portfolio of estimators which all have the same in-sample behaviors. However, the best performing covariance estimation method is in the Ex-post period is the Portfolio of estimators which has the lowest standard deviation followed by the PCA. This result is consistent with that of Disatnik and Benninga [5], who record the Portfolio of estimators as one of the best methods. The results suggest that the performance of the Sample covariance matrix can be improved by other methods that have the same in-sample exposures as it has but offer risk reduction capabilities for the out-of-sample period, which really is our major concern as we want to hold the portfolio out-of-sample. It is also interesting to note that simplicity is best as the portfolio of estimators actually performs better than the more complexly determined Ledoit shrinkage and PCA methods. This is the emphasis given by Disatnik and Benninga [5] that simplicity is best. However, it is important to note that the difference in the Ex-post standard deviations and betas between the estimators is small, hence one can consider using the sample covariance for constrained portfolio optimization in a South African setting.

## 5.2 Recommendations

This paper focuses on the constrained portfolio optimization of the Minimum Variance portfolio. In order to further investigate the performance of the covariance matrix estimators this study could be extended to the case when the portfolio is not Minimum Variance. Such a study would involve generating the expected future returns of the stocks present in the portfolio. In such a setup the optimizer would be fed with the estimated covariance as well as the expected returns. This means that the optimizer will produce portfolios that have reduced risk exposure for a given return level. Ledoit and Wolf [9] performed a similar analysis using shrinkage estimators and established that these out-perform the Sample covariance matrix.

The performance of the estimators could also be compared when the portfolio is unconstrained. Bengtsson and Holst [3] considered this scenario and it revealed that all the methods considered in this paper actually out-perform the Sample covariance in terms of risk.

Another scenario could be to allow the in-sample as well as the out-of-sample periods to vary and observe the performance of the covariance estimators.[5]

# References


[1] G. K. Basak, R. Jagannathan, and T. Ma. Assessing the Risk in Sample Minimum Risk Portfolios. *Working paper, University of Bristol, Northwestern University and University of Utah*, 2004.

[2] Z Bodie, A. Kane, and A.J. Marcus. *Investments*. Mc-Graw Hill, Newyork, fifth edition, 2003.

[3] Bengtsson C. and J. Holst. On Portfolio Selection: Improved Covariance Matrix Estimation for Swedish Asset Returns. *Working paper, Lund University and Lund Institute of Technology*, 2002.

[4] Luenberger D. *Investment Science*. Oxford University Press, Oxford, 1998.

[5] Disatnik D.J. and S. Benninga. Shrinking the Covariance Matrix. *Journal of Portfolio Management, Volume 33, Number 4*, Summer 2007.

[6] Brian H. and V.T. Daniel. *ESSENTIAL MATLAB for Engineers and Scientists*. ELSEVIER, Oxford, third edition, 2007.

[7] Markowitz H. Portfolio Selection. *Journal of Portfolio Management, Volume 7, Number 1*, March 1952.

[8] Wolf M. Resampling vs. Shrinkage for Benchmarked Managers. *Wilmott Magazine, Volume 1, pp. 76-82*, 2007.

[9] Ledoit O. and M. Wolf. Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management, Volume 31, Number 1*, Fall 2004.

[10] Clarke R., H. de Silva, and S. Thorley. Minimum-Variance Portfolios in the U.S. Equity Market. *Journal of Portfolio Management, Volume 33, Number 1*, Fall 2006.

[11] Guhr T. Econophysics. *Lecture Notes, Department of Mathematical Physics, Lund Institute of Technology*, 2001.

[12] C.G. Troskie, R. Mupambirei, R.K. Guo, and N. Hossain. Accurate Estimates of Risk using the Troskie-Hossain Improved Sharpe Single Index Model. Dynamic Time Series Models. *University of Cape Town*, 2009.