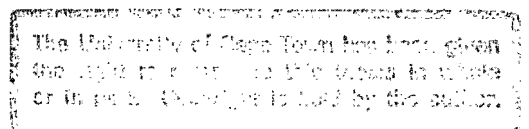


**Enhancements to the Markowitz Mean-  
Variance Optimisation Process of Asset  
Allocation**

A DISSERTATION PRESENTED TO  
THE FACULTY OF COMMERCE

IN FULFILMENT OF  
THE REQUIREMENTS FOR THE  
MASTERS OF COMMERCE DEGREE

BY  
WARREN McLEOD  
6 September 1998



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.

## ACKNOWLEDGEMENTS

My primary thanks goes to my Father |without whom this thesis would not have been possible. I also wish to express my appreciation to my family for the support they have always offered me, especially during the time of this thesis. I also thank my supervisor Professor Dave Bradfield for his advice, intuition and vision.

University of Cape Town

# TABLE OF CONTENTS

## CHAPTER 1 INTRODUCTION

|                                 |     |
|---------------------------------|-----|
| 1. INTRODUCTION.....            | 1.1 |
| 2. HISTORICAL BACKGROUND .....  | 1.1 |
| 3. ORGANISATION OF THESIS ..... | 1.6 |

## CHAPTER 2 LITERATURE REVIEW

|                                                                                |      |
|--------------------------------------------------------------------------------|------|
| 1. INTRODUCTION.....                                                           | 2.1  |
| 2. THE MARKOWITZ MODEL .....                                                   | 2.2  |
| 2.1. Problems with the Implementation of the Markowitz Model.....              | 2.5  |
| 3. THE TRACKING ERROR (TE) PORTFOLIO SELECTION MODEL .....                     | 2.8  |
| 4. THE MEAN VARIANCE TRACKING ERROR (MVTE) PORTFOLIO SELECTION MODEL .....     | 2.12 |
| 5. MEAN-VARIANCE-DOWNSIDE TRACKING ERROR (DTE) PORTFOLIO SELECTION MODEL ..... | 2.16 |
| 6. BAYESIAN ESTIMATES OF RETURN.....                                           | 2.19 |

## CHAPTER 3 THE OBJECTIVE FUNCTION

|                                                                               |      |
|-------------------------------------------------------------------------------|------|
| 1. SYNOPSIS.....                                                              | 3.1  |
| 2. INTRODUCTION.....                                                          | 3.3  |
| 2.1 Notation used for calculations.....                                       | 3.4  |
| 3. GENERATING THE EFFICIENT FRONTIER USING THE MARKOWITZ (MV) MODEL.....      | 3.7  |
| 3.1 Data.....                                                                 | 3.7  |
| 3.2 A Demonstration of the Markowitz (MV) Model.....                          | 3.9  |
| 3.3 The inclusion of cash in the opportunity set.....                         | 3.15 |
| 3.4 Insights for Practitioners .....                                          | 3.17 |
| 4. GENERATING THE EFFICIENT FRONTIER USING A TRACKING ERROR (TE) MODEL .....  | 3.18 |
| 4.1 The inclusion of cash in the opportunity set.....                         | 3.23 |
| 4.2 Insights for Practitioners .....                                          | 3.23 |
| 5. COMPARISON OF THE MV AND TE EFFICIENT FRONTIERS.....                       | 3.25 |
| 5.1 Insights for Practitioners .....                                          | 3.27 |
| 6. MEAN-VARIANCE TRACKING ERROR (MVTE) MODEL.....                             | 3.28 |
| 6.1 The inclusion of cash in the opportunity set.....                         | 3.34 |
| 6.2 Insights for Practitioners .....                                          | 3.35 |
| 7. SUMMARY COMPARISON BETWEEN MV, TE AND MVTE PORTFOLIOS .....                | 3.36 |
| 8. SUMMARY OF THE IMPLEMENTATION OF THE MVTE MODEL.....                       | 3.38 |
| 9. DOWNSIDE MEASURE OF THE TRACKING ERROR (DTE).....                          | 3.40 |
| 9.1 Insights for Practitioners .....                                          | 3.44 |
| 10. THE INFLUENCE OF THE BENCHMARK ON THE TE AND DTE MODELS .....             | 3.45 |
| 10.1 Benchmarks for TE Models .....                                           | 3.45 |
| 10.2 Benchmarks for DTE Models .....                                          | 3.48 |
| 10.3 The Selection of the Benchmark .....                                     | 3.51 |
| 11. COMPARISON OF THE TE, DTE AND MV EFFICIENT FRONTIERS .....                | 3.52 |
| 12. MEAN-VARIANCE TRACKING ERROR DOWNSIDE TRACKING ERROR MODEL (MVTEDTE)..... | 3.54 |
| 13. CONCLUSION .....                                                          | 3.55 |
| 14. APPENDIX .....                                                            | 3.56 |
| 14.1 Appendix 1.....                                                          | 3.56 |
| 14.2 Appendix 2.....                                                          | 3.59 |

**CHAPTER 4 INPUT ESTIMATION**

|                                                                                                                                                                                                                   |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 1. SYNOPSIS .....                                                                                                                                                                                                 | 4.1  |
| 2. INTRODUCTION .....                                                                                                                                                                                             | 4.3  |
| 3. THE BAYESIAN ESTIMATOR TECHNIQUE .....                                                                                                                                                                         | 4.6  |
| 3.1 Components of the Conditional Probability Density Function, $\pi(\mu   r_i)$ .....                                                                                                                            | 4.7  |
| 3.1.1 Insights for Practitioners .....                                                                                                                                                                            | 4.9  |
| 3.2 The Influence the Prior Distribution $\pi(\mu)$ and the Likelihood Function $\Lambda(\mu; r_1, r_2, \dots, r_n)$ have on the Conditional Probability Density Function $\pi(\mu   r_1, r_2, \dots, r_n)$ ..... | 4.9  |
| 3.3 The Effect on the Conditional Probability Density $\pi(\mu   r_1, r_2, \dots, r_n)$ , when altering the Prior .....                                                                                           | 4.14 |
| 3.4 Obtaining Estimates for a Security's Expected Return and Variance .....                                                                                                                                       | 4.19 |
| 3.5 The Effect Bayesian Estimates have on Efficient Frontiers .....                                                                                                                                               | 4.24 |
| 4. EMPIRICAL RESEARCH .....                                                                                                                                                                                       | 4.26 |
| 4.1 Outline of Empirical Research .....                                                                                                                                                                           | 4.26 |
| 4.2 Bayesian Estimates using a 90% Weight on the Portfolio Manager's Subjective Expectation... ..                                                                                                                 | 4.29 |
| 4.3 Bayesian Estimates using the Standard Errors of the Estimates in Calculating the Weights placed on the Portfolio Manager's Subjective Expectations .....                                                      | 4.35 |
| 4.4 The Effects of the "Biased" Bayesian Variance Estimate on the Efficient Frontier and Optimal Portfolios .....                                                                                                 | 4.43 |
| 4.5 Practical Insights for Practitioners .....                                                                                                                                                                    | 4.49 |
| 5. CONCLUSION .....                                                                                                                                                                                               | 4.50 |
| 6. APPENDIX .....                                                                                                                                                                                                 | 4.51 |
| 6.1 Appendix 1 .....                                                                                                                                                                                              | 4.51 |
| 6.2 Appendix 2 .....                                                                                                                                                                                              | 4.53 |
| 6.3 Appendix 3 .....                                                                                                                                                                                              | 4.55 |
| 6.4 Appendix 4 .....                                                                                                                                                                                              | 4.61 |
| 6.5 Appendix 5 .....                                                                                                                                                                                              | 4.65 |

**CHAPTER 5 BETA ESTIMATION**

|                                                                     |      |
|---------------------------------------------------------------------|------|
| 1. INTRODUCTION .....                                               | 5.1  |
| 2. THE CAPITAL ASSET PRICING MODEL (CAPM) .....                     | 5.2  |
| 3. OUTLIERS AND INFLUENTIAL POINTS .....                            | 5.4  |
| 3.1. Influential Points .....                                       | 5.4  |
| 3.2. Student's t Statistic .....                                    | 5.8  |
| 3.3. Cook's Distance .....                                          | 5.9  |
| 4. OUTLIER SHIFT MODEL .....                                        | 5.10 |
| 5. EMPIRICAL RESEARCH .....                                         | 5.12 |
| 5.1. Cook's Distances .....                                         | 5.12 |
| 5.2. Outlier Shift Model .....                                      | 5.13 |
| 5.3. Forward Stepwise Regression .....                              | 5.13 |
| 5.4. Backward Stepwise Regression .....                             | 5.14 |
| 5.5. Consequence of Removing the Outlier .....                      | 5.15 |
| 5.6. Dates which Consistently Influenced the $\beta$ Estimate ..... | 5.18 |
| 5.7. Insights for Practitioners .....                               | 5.21 |
| 6. CONCLUSION .....                                                 | 5.22 |

**CHAPTER 6 CONCLUSION**..... 6.1**BIBLIOGRAPHY**..... 1

**CHAPTER 1**  
**INTRODUCTION**

University of Cape Town

## 1. INTRODUCTION

The primary objective of a portfolio manager is to maximise the return of the portfolio he/she is regulating at the predetermined level of risk or similarly minimising the expected risk of the portfolio at the predetermined level of return. Effectively a portfolio manager needs to forecast the expected returns and risks of the assets he/she has in his/her opportunity set. This procedure is not as simple as it appears at face value. Many statisticians are very sceptical of the traditional forecasts used. Recently more comprehensive measures of an asset's expected return and risk have been proposed, which we have examined and compared in the South African context. These measures include adjusting the forecast of a security's expected return to include the manager's own subjective expectations and accepting and implementing the notion that a portfolio manager is measured relative to some benchmark. Consequently the portfolio manager's deviation from this benchmark could well be a better and more reasonable measure of a portfolio's risk than the traditional sample variance measure.

## 2. HISTORICAL BACKGROUND

"The Benefits of International Diversification in Bonds" was the heading of Levy and Lerman's<sup>1</sup> article in the September-October 1988 *Financial Analysts Journal*. Some readers of this article may well have been misled into believing that Levy and Lerman were proposing a *new* technique that is aimed at satisfying the objectives of an investor. If they did then they were deceived. Already back in 1959 Harry Markowitz<sup>2</sup> proposed and showed that an investor could reduce the overall *risk* of his/her investment by forming a *well-diversified* portfolio. Markowitz addressed this issue of portfolio risk and diversification in his 1959 dissertation.

The emphasis of Markowitz's dissertation was the measurement of a portfolio's risk. He adopted variance as a measure of risk since it was simple, logical and easy to

---

<sup>1</sup> Levy, H and Lerman, Z, "The Benefits of International Diversification in Bonds", *Financial Analysts Journal*, September-October, 1988.

<sup>2</sup> Markowitz, HM, "Portfolio Selection: Efficient Diversification of Investment", 1959.

calculate. Markowitz also considered other alternative risk measures and he decided that the most theoretically robust measure of risk was semivariance. Semivariance is the expected value of squared *negative* deviations about a specified "target" rate of return. Another term for semivariance is downside tracking error. Markowitz adopted simple variance over semivariance as a measure of risk in his proposed model because of the purported computational problems associated with calculating the semivariance statistic (he also did not have the computer power what we have today back in 1959)!

It was Harlow<sup>3</sup> (1991) who also acknowledged that semivariance may well be a better measurement of portfolio risk. If return distributions are approximately symmetrical about a specified "target" (benchmark) then variance is a sufficient measurement of portfolio risk and the differences between the two measurements is small. An assumption often used in the models of Capital Market Theory is that returns are distributed normally, and a normal distribution is symmetrical. However, Harlow challenges this assumption. Harlow bases his challenge on the research conducted in the finance, economic and psychological fields. The researchers in these fields have noted that over the last three decades individuals have viewed return dispersion in an asymmetric manner; that is losses weigh more heavily than gains (Libby and Fishburn<sup>4</sup>, 1977). Consequently Harlow (1991) proposed a Downside Tracking Error Portfolio Selection Model, which is based on semivariance as a measurement of portfolio risk.

Franks<sup>5</sup> (1992) and Roll<sup>6</sup> (1992) also examined portfolio selection models at the same time as Harlow. They too adopt the view that investors or portfolio managers are concerned about the performance of a portfolio about a specified "target" (benchmark). Consequently Rolls and Franks argue that a portfolio's risk should be measured by the *total* squared deviations (tracking error squared) about a specified benchmark, not just merely the *negative* deviations, which Harlow argues should be

---

<sup>3</sup> Harlow, WV, "Asset Allocation in a Downside-Risk Framework", *Financial Analysts Journal*, September-October, 1991, pp. 28-40.

<sup>4</sup> Libby, R and Fishburn, P, "Behavioural Models of Risk Taking in Business Decisions: A Survey and Evaluation", *Journal of Accounting Research*, Autumn 1977.

<sup>5</sup> Franks, ECF, "Targeting Excess-of-Benchmark Returns", *The Journal of Portfolio Management*, Summer 1992, pp. 6-12.

<sup>6</sup> Roll, R, "A Mean/Variance Analysis of Tracking Error", *The Journal of Portfolio Management*, Summer 1992, pp. 13-22.

the appropriate measurement of a portfolio's risk. Chow<sup>7</sup> (1995) took Franks and Roll's proposals one step further. He too considers tracking error as a useful measure of risk but he argues that it needs to be used in conjunction with simple variance. Chow's justification for his proposed Mean-Variance-Tracking Error (MVTE) Portfolio Selection Model is based on his view that investors acknowledge the importance of benchmarks in the decision process but at the same time they are not indifferent to the variance of absolute returns. Chow encompasses both of these measures of risk in his proposed portfolio selection model.

Black and Litterman<sup>8</sup> (1992) have also conducted much research in the field of portfolio selection in the 1990's. The work performed by Harlow, Franks and Roll (and many others) focussed on the measurement of the portfolio's risk whereas Black and Litterman focused on the measurement of a portfolio's expected return. Black and Litterman justified their research by pointing out that quantitative models have failed to assist with the critical allocation decision (due to the unreasonable nature in the results). They believe the weakness of most portfolio selection models have been in their estimations of the securities' expected returns.

Expected returns are difficult to forecast. Most of the models require an estimate for the expected return for each and every asset in the opportunity set. However, investors typically have knowledgeable views about returns in only a few markets. A weakness of many models is that they do not differentiate between knowledgeable views and weakly held views. A further weakness Black and Litterman have highlighted is that the optimal portfolio asset weights of standard asset allocation models are extremely sensitive to the return assumptions used. Black and Litterman believe that their model offers an intuitive solution to these two problems. The Black and Litterman model combines two tenets of modern portfolio theory - the mean-variance framework of Markowitz, and the Capital Asset Pricing Model (CAPM) of Sharpe<sup>9</sup> and Lintner<sup>10</sup> (1965). Black and Litterman's estimate is a weighting of an investor's own subjective

---

<sup>7</sup> Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995, pp. 54-60.

<sup>8</sup> Black, F and Litterman, R, "Global Portfolio Optimization", *Financial Analysts Journal*, September-October 1992, pp. 28-42.

<sup>9</sup> Sharpe, WF, "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk", *Journal of Finance*, September 1964.

expectation in a security's return and the CAPM estimate. The statistical term used to describe this type of estimate is a Bayesian Estimate.

The general Bayesian estimate itself receives mixed support. The Bayes estimate is based on the Bayes' Theorem. The Reverend Thomas Bayes was born in 1702. The mathematical research the reverend undertook was small in quantity but certainly of the very highest quality. The work of his which has certainly gained much attention is an essay he wrote with the title "An Essay Towards Solving a Problem in the Doctrine of Chances". This essay was only recognised after Bayes passed away and it is the foundation of the Bayes' Theorem.

The Bayesian philosophy went through a quiet period in the 1800's. It was the work of Sir Harold Jeffreys, born in 1891, that resuscitated Bayesianism to give it form (logical status) and substance (solutions to the pressing statistical paradigms of the day). It was also in Jeffreys' time that Bayesianism came under the sharpest criticism. A component of Bayesianism theory is the application of a subjection distribution, termed the prior distribution. The authorities of the 1930's and 1940's were Fisher and Neyman who attacked Jeffreys because the prior is *non-frequentist* and subjective by nature. The frequentist approach is based on repeating an experiment (preferably many times) and obtaining information from the experiment whereas the Bayesian approach is the quantification of ones uncertainty about the outcome of the experiment and is thus a personal or subjective notion. Jeffreys' own contribution to Bayesianism was the proposal of using an *improper* prior to represent the scenario where one does not have any subjective beliefs. The prior he proposed is termed to be improper as it does not satisfy the preconditions for a probability density function i.e. it does not integrate out to 1. Fisher and Neyman's criticism focused on this improper prior because effectively Jeffreys was attempting to *quantify* having *no* subjective belief (he was doing it because the prior, a subjective distribution, was a necessary component of Bayesianism and without it Bayesianism would fail). Although the traditionalists of Jeffreys' day differed in *opinion* to him some did acknowledge his

---

<sup>10</sup> Lintner, J, "The Validation of Risk Assets and the Selection of Risk, Investments in Stock Portfolios and Capital Budgets", *Review of Economics and Statistics*, February 1965.

work. Lindley<sup>11</sup> (1956) explains that he and other students "... had to admit that there was a cogent argument [with the Bayesian argument] that we rather arrogant young men could not demolish".

Jeffreys applied his work in the natural sciences but he often did propose that there were implications for Bayesianism in the fields of social science and economics. Those who have applied Bayesianism in the field of economics include Grossman<sup>12</sup> (1980) who in his paper illustrated the use of Bayes' Theory in econometric models of markets subject to uncertainty. An economic related field that has certainly applied Bayesianism analysis for a number of years is actuarial science (Miller<sup>13</sup> 1980).

The focus of this thesis is on the practical application of portfolio selection. It is a field that receives much attention, no more so than after the world market crashes (i.e. October 1997) which highlighted the importance of risk management. Consequently there is a need to examine the current tools in current use to create our portfolios and to look at ways in which they could be improved. The Bayesians have certainly contributed in this area, and more noticeably in the 1990's. We shall examine their contributions quite extensively in this thesis.

---

<sup>11</sup> Lindley, DV, "On a Measure of the Information Provided by an Experiment", *Annals of Mathematical Statistics* 22, 1956, pp. 986-1005.

<sup>12</sup> Grossman, SJ, "Rational Expectations and Econometrics Modelling of Markets subject to Uncertainty: A Bayesian Approach", *A Bayesian Analysis in Econometrics and Statistics*, 1980, pp. 143-158.

<sup>13</sup> Miller, RB, "Actuarial Applications of Bayesian Statistics", *Bayesian Analysis in Econometrics and Statistics*, 1980, pp. 197-212.

### 3. ORGANISATION OF THESIS

This thesis is concerned with examining proposed practical portfolio selection models and investigating them empirically in the South African context. All models, especially portfolio selection models are based on various assumptions and in this thesis we firstly examine the basis of the models before implementing and probing them empirically. In this way we shall have a better understanding of the empirical results.

All the portfolio selection models which we shall examine consist of two core elements that need to be estimated. These two elements are a measure of a portfolio's risk and an estimate of its expected return. All the chapters in this thesis will be dealing with either one or both of these two measures.

The thesis comprises of six chapters.

Chapter 1 discusses the historical background of portfolio selection models and the historical advancements made in the measures of a portfolio's risk and return.

Chapter 2 is a literature review. The literature review essentially covers all the aspects of portfolio selection which we shall examine empirically. It discusses the work that has already been accomplished in the areas of portfolio selection which we shall explore and implement ourselves in the South African context.

Chapter 3 is an examination of the measure of a portfolio's risk. It firstly discusses the measures currently used and then *compares* the proposed measures empirically. The main theme in the chapter is the measure of a portfolio's risk relative to a benchmark rather than simply using its variance as the measure of its risk, which has been the traditional measure of a portfolio's risk.

Chapter 4 takes an extensive look at the estimation of a portfolio's expected return. It has been proposed that a Bayesian estimate can produce a far more realistic and practical estimate for a portfolio's expected return than a simple sample mean of past

returns. Section 3 of Chapter 4 discusses quite extensively how these Bayesian estimates are obtained. Section 4 then examines these Bayesian estimates empirically.

Chapter 5 contains a somewhat different strand of research, but also has some connections with forecasting a portfolio's expected return. Essentially in Chapter 4 the Capital Asset Pricing Model (CAPM) is used in the Bayesian context to produce an estimate for a security's expected return, however, in the CAPM model itself one needs to estimate a security's beta (a measure of a security's market risk). Chapter 5 looks at the possible shortcomings of the estimate of a security's beta and proposes and examines ways to avoid these shortcomings. Consequently we propose a refined adjustment to the measure of a security's CAPM estimate.

Chapter 6 is a final conclusion and it looks at what we have achieved in this thesis.

**CHAPTER 2**

**LITERATURE REVIEW**

University of Cape Town

## 1. INTRODUCTION

When portfolio managers are asked the question, "What is your objective?" many surely respond in a similar manner, "To maximise the return of my portfolio while simultaneously minimising its risk". Hence there are two primary parameters one needs to estimate - a portfolio's expected return and a portfolio's risk, or uncertainty. There have been numerous proposals for the measurement of a portfolio's risk. Some portfolio managers have suggested that a portfolio's risk is simply a measure of its uncertainty and this uncertainty is captured by its aggregate variance (i.e. deviations from the *mean*), whereas other portfolio managers acknowledge that they are rewarded for their portfolio's performance as measured by the aggregate deviations from a predefined benchmark. Hence they believe that a superior measure of a portfolio's risk would be its relative risk or variance of relative returns which is termed its tracking error.

Similarly there have been numerous proposals for using different measures for a portfolio's expected return. There are those who simply use sample means as the estimates for the securities' expected returns whereas others have proposed using Bayesian expectations as these estimates capture the portfolio managers' own subjective expectations (which the simple sample means certainly cannot do). Markowitz (1952<sup>1</sup>, 1959<sup>2</sup>) addressed this issue of portfolio selection. The consequent portfolio selection model he proposed has been labelled the "Markowitz" model and forms the basis of many developments in the area of portfolio selection/asset allocation.

---

<sup>1</sup> Markowitz, H.M, "Portfolio Selection", *Journal of Finance*, March 1952, pp. 77-91.

<sup>2</sup> Markowitz, H.M, "Portfolio Selection: Efficient Diversification of Investments", 1959.

## 2. THE MARKOWITZ MODEL

Markowitz certainly acknowledged that risk and return are two core parameters that need to be estimated. Markowitz simply used sample standard deviations as measures of the securities' risks. The return of a *portfolio* is taken to be the weighted sum on each of the individual security's returns, and the portfolio's standard deviation (or variance) is a measure of its risk.

The investor's choice of an optimal portfolio using the Markowitz model is based on expected return and the uncertainty associated with that return and the covariance between individual shares. The choice is based on the assumption that a portfolio manager will undoubtedly prefer an "efficient" portfolio to an "inefficient" portfolio. A portfolio is termed to be "efficient" if it is impossible to find a portfolio with a greater expected return without incurring greater risk. Similarly a portfolio is also "efficient" if it is impossible to reduce the portfolio's risk without decreasing its expected return.

The objective of the portfolio manager is to maximise a portfolio's return while simultaneously minimising its risk (or equivalently its variance). In order to achieve this objective the portfolio manager needs to stipulate his/her risk tolerance. The risk tolerance is the number of units of variance the portfolio manager is willing to take on in order to gain a unit of return. Effectively the objective, as proposed by Sharpe<sup>3</sup>, (1970) is to

$$\text{MAX Portfolio's Expected Return} - \frac{\text{Expected Variance}}{\text{Risk Tolerance}}$$

The following notation will be used to simplify the ensuing expressions:

$E_i$  = Expected Return on the  $i^{\text{th}}$  security

$\sigma_i$  = Standard deviation of return on the  $i^{\text{th}}$  security

$E_p$  = Expect return on the portfolio

<sup>3</sup> Sharpe, W.F, "Portfolio Theory and Capital Markets", 1970.

$\sigma_p$  = Standard deviation of return on the portfolio

$\sigma_{ij}$  = Covariance between security i and security j

$\rho_{ij}$  = Correlation coefficient for the returns on security i and j

$X_i$  = Proportion of funds invested in security i

$N$  = Total number of securities considered

$\frac{1}{\lambda}$  = Risk Tolerance

also

$$\begin{aligned} E_p &= \sum_{i=1}^N X_i E_i \\ &= X' E \end{aligned}$$

and

$$\begin{aligned} \sigma_p^2 &= \sum_{i=1}^N \sum_{j=1}^N X_i X_j \sigma_{ij} \\ &= X' \Sigma X \end{aligned}$$

where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_N^2 \end{pmatrix}$$

Hence the objective proposed by Sharpe, in accordance with the Markowitz Model, is to

$$\text{MAX } E_p - \lambda \sigma_p^2$$

Each optimal portfolio will be efficient. By considering all levels of risk tolerance (letting  $\lambda$  run from 0 to  $\infty$ ) and maximising the objective function at each of these levels, one will obtain all the efficient portfolios. A plot of all these efficient

portfolios in mean-standard deviation space results in a curve that is termed the efficient frontier. More simply, the efficient frontier is the set of efficient portfolios for different levels of risk.

Figure 1 is an example of an efficient frontier. The data for this example was extracted from the hypothetical scenario Beninga<sup>4</sup> (1997) discusses in his book "Financial Modelling".

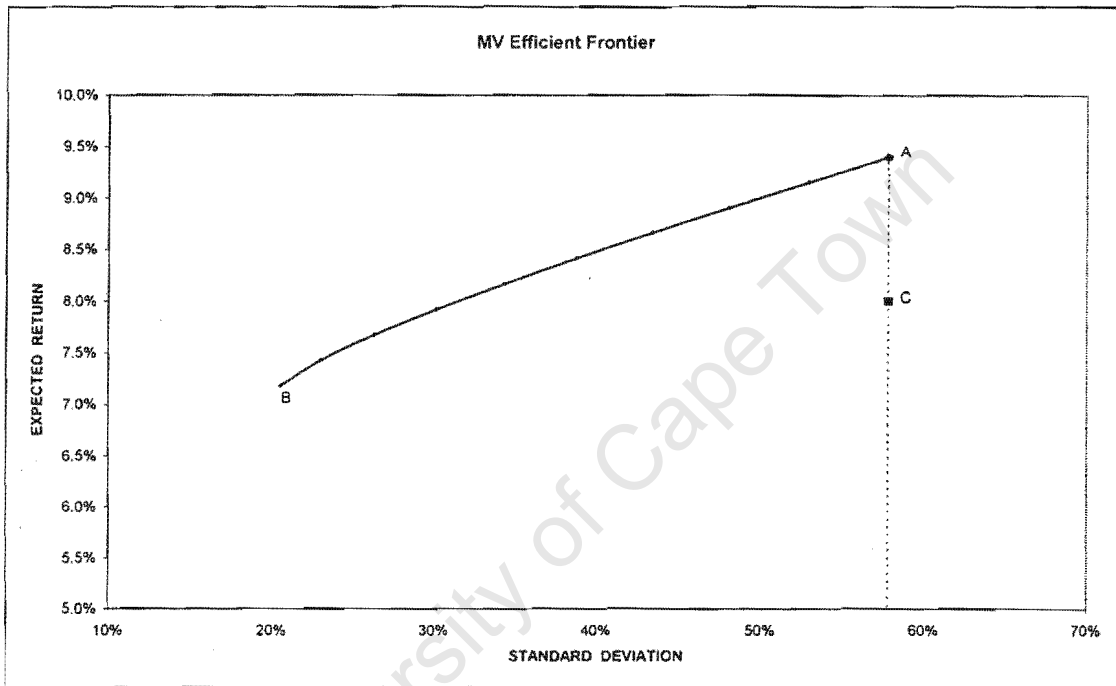


Figure 1 The Efficient Frontier

Points A, B and C in Figure 1 are three specific portfolios. Portfolio A consists of a single asset. It is the *asset* in the opportunity set with the highest expected return. It is consequently also the *portfolio* with the highest expected return. One could certainly not increase the expected return of this portfolio by investing in any assets that have lower expected returns than what it itself already has. Portfolio C is inefficient, since portfolio A has an equivalent standard deviation (measure of risk) to portfolio C, but simultaneously it has a higher expected return. In fact, any portfolio to the south-east of the efficient frontier is inefficient. Portfolio B is the portfolio with the minimum standard deviation (risk). It does not necessarily comprise simply of the asset with the smallest standard deviation, since a portfolio of several assets can collectively have a

<sup>4</sup> Beninga, S, "Financial Modelling", 1997.

smaller standard deviation than the asset in the opportunity set with the smallest individual standard deviation. If a risk-free asset were in the opportunity set, then the minimum risk portfolio would simply be the risk-free asset and it would have a standard deviation of zero.

The Markowitz Model certainly appears logical and simple and we could certainly expect it to have much support. Yet practically it is not used as extensively as we would expect (Fischer<sup>5</sup> and Statman 1997). There are claims that the traditional implementation of the Markowitz Mean-Variance Portfolio Selection Model has some shortcomings.

### **2.1. Problems with the Implementation of the Markowitz Model**

The quality of portfolios constructed using historical means, variances and correlations do not appeal to portfolio managers' intuition since these optimal portfolios comprise of very few of the securities in the opportunity set, and the weights of the securities included in the portfolios are extreme (Fischer and Statman 1997). Michaud<sup>6</sup> (1989) believes that these extreme weightings are displayed since *historical* means, variances and correlations are used as estimates for the true parameters. Michaud asserts that the consequence of using historical estimates is that too much portfolio weight is shifted to those securities which show high returns, low variance and low correlations *historically* with other securities. He claims that these features are ideal, but because they happened in the past it certainly does not ensure they will exist in the future. Michaud proposes that better estimates with less estimation error should be used. Green and Hollifield<sup>7</sup> (1992) also picked up on these extreme weightings but their argument leaned the other way. They believe that eliminating estimation errors would not eliminate extreme weightings and would not make efficient mean-variance portfolios intuitively appealing. Green and Hollifield recommend that investors abandon their intuition about the features of desirable portfolios and accept optimised mean-variance portfolios.

<sup>5</sup> Fischer, K and Statman, M, "The Mean-Variance Optimization Puzzle", *Financial Analysts Journal*, July/August 1997, pp. 41-50.

<sup>6</sup> Michaud, R, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?", *Financial Analysts Journal*, January-February 1989, pp. 31-42.

Much of the criticism of the Markowitz Model is levelled at the estimation of returns. Frost and Savirano<sup>8</sup> (1988) used the Capital Asset Pricing Model (CAPM) equation in Bayesian terms for determining the expected returns of securities but again they found that the (Markowitz) Mean-Variance Optimisation Model resulted in only a few of the securities being prevalent in the efficient Mean-Variance portfolios. There are obviously contradictory views on whether better estimators will in fact result in more intuitively appealing portfolios. This forms one of the primary aims of this thesis.

Black<sup>9</sup> (1995) also admits that estimating expected returns is difficult. He claims that daily data hardly helps at all - only longer time periods are useful, but he believes, in order to produce a reliable estimate, we need an extremely long period, over 200 months, to estimate the average. He suggests that other models could well be used for estimating a security's expected return, but he stipulates that when there is no theory to back the factors used in the proposed model then the past average return can well give the best estimate, but that is normally a highly inaccurate estimate.

Much criticism has been levelled at estimating the expected returns of securities in the Markowitz Model. Equivalently, much attention has been aimed at the measure of risk. Markowitz proposed using sample variance as a measure of risk. Today most portfolio managers are judged or rewarded by the total return performance of the portfolio relative to a pre-specified benchmark, usually a broadly diversified index of assets (Roll<sup>10</sup> 1992). Roll claims that the variability of this *relative return*, termed the tracking error (TE) squared, could well be a superior measure of risk in the context of the portfolio selection model. The tracking error might well be a superior estimate for the portfolio's risk from the portfolio manager's perspective since he/she is measured against a benchmark, however, from an *investor's* perspective variance might well still be the best measure of risk (as it captures the chance of absolute loss). This could well be the case if an investor is merely concerned with the investment's total performance not its relative performance (as measured against some benchmark).

---

<sup>7</sup> Green, R and Hollifield, B, "An Empirical Bayes Approach to Efficient Portfolio Selection", *Journal of Financial and Quantitative Analysis*, September 1992.

<sup>8</sup> Frost, P and Savarino, J, "An Empirical Bayes Approach to Efficient Portfolio Selection", *Journal of Financial and Quantitative Analysis*, September 1988.

<sup>9</sup> Black, F, "Estimating Expected Return", *Financial Analysts Journal*, 1995, pp. 168-171.

Franks<sup>11</sup> (1992) and Roll (1992) have separately used the tracking error as a measure of risk in their portfolio selection models. We have termed the portfolio selection model, which bases its measure of risk on the variability of relative return, the Tracking Error (TE) portfolio selection model.

<sup>10</sup> Roll, R, "A Mean/Variance Analysis of Tracking Error", *Journal of Portfolio Management*, Summer 1992, pp. 13-22.

<sup>11</sup> Franks, E.C, "Targeting Excess-of-Benchmark Returns", *The Journal of Portfolio Management*, Summer 1992, pp. 6-12.

### 3. THE TRACKING ERROR (TE) PORTFOLIO SELECTION MODEL

Portfolio managers' performances are often measured relative to a benchmark. Consequently they are concerned with deviation from this benchmark. As with the Markowitz Mean-Variance Portfolio Selection Model, the objective of the TE Portfolio Selection Model is to maximise the portfolio's return and to minimise its risk, but risk is now defined as tracking error squared, not variance.

The objective of the TE model is to

$$\text{MAX Portfolio's Expected Return} - \frac{(\text{Tracking Error})^2}{\text{Tracking Error Tolerance}}$$

The tracking error tolerance sets the price of tracking error in terms of return. For simplification  $\phi$  ( $\psi$ ) will be used to represent  $\frac{1}{\text{Tracking Error Tolerance}}$ .

The following notation together with that specified on page 2.2 will be used to simplify the ensuing expressions:

$R_{it}$  = the return for asset  $i$  at time  $t$

$R_{bt}$  = the return of the benchmark at time  $t$

$RR_{it}$  = the relative return for asset  $i$  at time  $t$

$$RR_{it} = R_{it} - R_{bt}$$

$\Sigma^*$  = the variance-covariance matrix of the relative returns ( $RR_i$ 's) of the assets in the opportunity set

$\sigma_p^{*2}$  = the variance of relative return on the portfolio

where

$$\sigma_p^{*2} = X' \Sigma^* X$$

Effectively the objective of the Franks and Roll tracking error portfolio selection models is to

$$\text{MAX } E_p - \psi \sigma_p^{*2}$$

Each portfolio will be efficient in mean-tracking error space. By considering all levels of tracking error tolerance (letting  $\psi$  run through the values 0 to  $\infty$ ) one will obtain all the TE efficient portfolios and the plot of these efficient portfolios in mean-TE space is termed the TE efficient frontier.

Figure 2 is an example of a TE efficient frontier. The TE efficient frontier was constructed from the data Chow<sup>12</sup> (1995) examined when he himself discussed the TE efficient frontier.

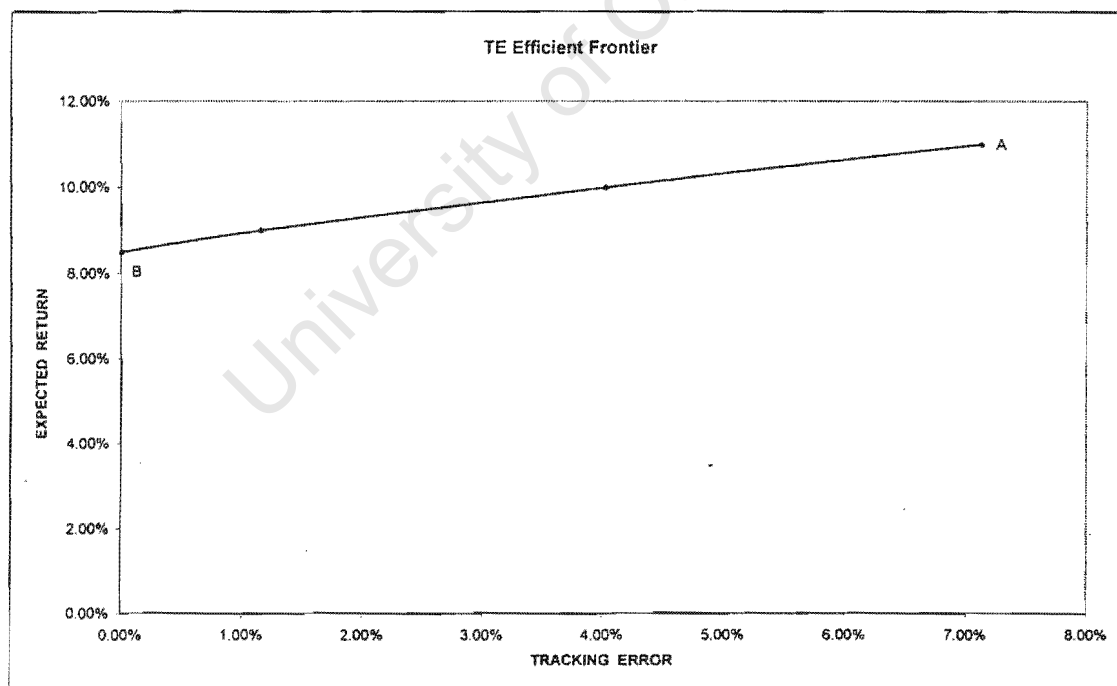


Figure 2 The Tracking Error (TE) Efficient Frontier

<sup>12</sup> Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995, pp.54-60.

Points A and B in Figure 2 are two specific portfolios. Portfolio A consists of a single asset. This is the asset that has the highest expected return (regardless of its tracking error). The expected return of portfolio A cannot be improved by adding any other assets to it. Portfolio B is the portfolio with zero tracking error. This occurs when the benchmark itself is in the opportunity set, or if the benchmark is simply a weighting of the expected returns of the assets in the opportunity set. This point was made by Franks (1992) and by Black<sup>13</sup> and Litterman (1992). Equivalence to the benchmark in the Markowitz Model is the risk-free asset. The risk-free asset in the Markowitz Model forms the vertical axis intercept of the mean-variance efficient frontier (where the standard deviation (risk) is zero). Hence the benchmark (in Mean-Tracking Error space) performs a similar role to the risk-free asset (in Mean-Variance space). The benchmark will lie south east of the *mean-variance* efficient frontier in mean-variance space and similarly the risk-free asset will lie south east of the tracking error efficient frontier in *mean-tracking* error space. Effectively we are claiming that the benchmark is typically inefficient in a Markowitz mean-variance portfolio selection model (in an ex-post sense) and similarly the risk-free asset is inefficient in a Tracking Error Portfolio Selection Model.

The TE portfolio selection model has received much support. Roll explicitly advocates using TE as a measure of risk on two accounts:

1. Ideal portfolio managers would outperform the benchmark every month by a fixed amount implying zero tracking error volatility hence one could ascertain with complete statistical reliability that the portfolio manager is adding value over an index fund alternative (the benchmark).
2. The portfolio managers' performances are more than likely reviewed relative to a benchmark and the variability about this benchmark will be a good measure of risk.

Although the TE portfolio selection model has received much support, one needs to understand the foundations of the model before simply using it as a "black box". Great attention needs to be drawn to the benchmark. In fact a portfolio's variance

---

<sup>13</sup> Black, F and Litterman, R, "Global Portfolio Optimisation", *Financial Analysts Journal*, September-October 1992, pp. 28-43.

would be identical to its tracking error squared if the benchmark were a risk-free asset (Black<sup>14</sup> and Litterman 1991). The benchmark may well be different for different portfolio managers. Many portfolio managers use indices as their benchmarks. All too often these indices have zero weight in cash, a risk-free asset, hence a portfolio with a positive weight in cash will suffer from tracking error (Connor<sup>15</sup> 1995). Many portfolio managers certainly wish to have portfolios containing cash since cash lowers expected transaction costs; but efficient TE portfolios seldom if ever contain cash. Consequently these efficient TE portfolios are often considered to be unreasonable to many portfolio managers. Even if various portfolio managers have identical opportunity sets, expected returns and objectives, the efficient portfolios will differ if their benchmarks differ.

Chow (1995) acknowledges that tracking error might well be a good measure of risk (relative risk) from a portfolio manager's perspective since he/she is often measured against a benchmark, but simultaneously Chow believes that a portfolio selection model must also acknowledge a portfolio's absolute risk, as captured by its variance, since *investors* are often more concerned with this measure of risk. One would obviously need to stipulate how important each of these two measures of risk are, and this is captured in the Mean-Variance Tracking Error (MVTE) Portfolio Selection Model, which is proposed by Chow.

---

<sup>14</sup> Black, F and Litterman, R, "Global Asset Allocation with Equities, Bonds and Currencies", *Fixed Income Research*, October 1991.

<sup>15</sup> Connor, G and Leland, H, "Cash Management and Index Tracking", *Financial Analysts Journal*, November-December 1995, pp. 75-80.

#### 4. THE MEAN VARIANCE TRACKING ERROR (MVTE) PORTFOLIO SELECTION MODEL

Markowitz introduced the concept of portfolio selection based on return and variance. This was followed by the recognition that many portfolio managers' performances are evaluated relative to a benchmark and hence led to the establishment of a tracking error model which is based on return and relative risk. Chow (1995) claims that both approaches fail to yield satisfactory results. Chow believes that portfolio managers acknowledge the importance of benchmarks in decision making but they, or more specifically investors, are not indifferent to the variance of absolute returns. Hence Chow proposed a portfolio selection model that encompasses both measures of risk - variance and tracking error.

The objective of Chow's Mean-Variance-Tracking Error (MVTE) portfolio selection model is to

$$\text{MAX Portfolio's Expected Return} - \frac{\text{Expected Variance}}{\text{Risk Tolerance}} - \frac{(\text{Tracking Error})^2}{\text{Tracking Error Tolerance}}$$

or in the notation we have established

$$\text{MAX } E_p - \lambda \sigma_p^2 - \psi \sigma_p^*$$

In layman's terms the portfolio manager seeks portfolios with high expected returns, low variances (volatility on an absolute basis) as well as low tracking error squares (volatility relative to a benchmark). The portfolio manager needs to define the relative importance of these three goals. This is achieved by assigning appropriate values to the risk tolerance and tracking error tolerance.

The Markowitz and the TE portfolio selection models are two-dimensional models whereas the MVTE portfolio selection model is a three dimensional model. The MVTE model would be reduced to the equivalent Markowitz model or TE model if the portfolio manager's respective TE tolerance or risk tolerance were infinite. We

established that a plot of the efficient portfolios, achieved by utilising the Markowitz model, in mean-variance space would produce a mean-variance efficient frontier and similarly a plot of the efficient portfolios, attained by applying the TE model, in mean-tracking error space would produce the TE efficient frontier. However, if we were to plot the efficient portfolios of the MVTE model in mean-variance-tracking error space we would not observe a curve, but rather a 3 dimensional surface. Figure 3 is a plot of a surface generated from a hypothetical set of efficient MVTE portfolios.

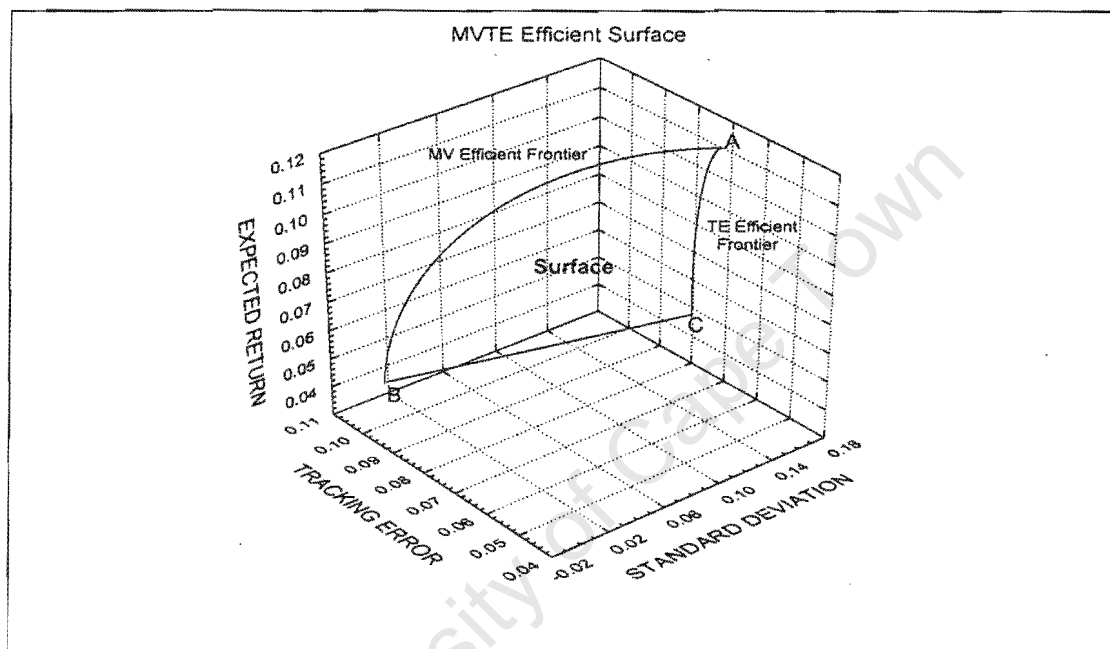


**Figure 3** A plot of the MVTE efficient surface

The mean-variance efficient frontier and the TE efficient frontier each delineate a border of this surface. The curve extending from point A to point B is the MV efficient frontier, and the curve extending from point A to point C is the TE efficient frontier. Both frontiers have been plotted in mean-standard deviation-tracking error space. These two frontiers meet at point A in Figure 3 where both the risk tolerance and tracking error tolerance are infinite. Point A is the portfolio with maximum expected return, and it simply consists of the asset in the opportunity set that has the highest expected return.

Chow analysed his proposed MVTE model (together with the Markowitz and TE models) empirically by considering an opportunity set containing US stocks, US small stocks, non-US stocks, US bonds, real estate, cash, international bonds and

international stocks. He first set the expected return to a value (we shall examine the case where he set the expected return to 9%) and he then determined the corresponding efficient portfolios using the Markowitz portfolio selection model and the TE portfolio selection model. He performed the same operation using the MVTE portfolio selection model. In Chow's analysis the MVTE optimal portfolio does not lie on either the mean-variance or on the TE efficient frontiers. Chow ensured this by setting the risk tolerance to 57.8 and the tracking error tolerance to 25. Chows findings are listed in Table 1.

**Table 1 Optimal portfolios obtained by Chow when using the Markowitz, Tracking Error and Mean-Variance-Tracking Error (MVTE) portfolio selection models respectively**

| CHARACTERISTICS      | PORTFOLIO SELECTION TECHNIQUE |        |        |
|----------------------|-------------------------------|--------|--------|
|                      | MARKOWITZ                     | TE     | MVTE   |
| Expected Return      | 9%                            | 9%     | 9%     |
| Standard Deviation   | 10.30%                        | 11.26% | 10.78% |
| Tracking Error       | 4.69%                         | 1.15%  | 1.79%  |
| <b>ALLOCATION</b>    |                               |        |        |
| Cash                 | 0%                            | 0%     | 0%     |
| International Bonds  | 9.4%                          | 0.9%   | 3.5%   |
| International Stocks | 23.9%                         | 7.9%   | 12.7%  |
| Real Estate          | 10.6%                         | 1.2%   | 4.1%   |
| Small Stocks         | 7.9%                          | 2.2%   | 3.9%   |
| US Bonds             | 29.1%                         | 36%    | 33.9%  |
| US Stocks            | 19.2%                         | 51.8%  | 41.9%  |

There is a notable difference in the portfolio weights between the Markowitz and TE optimal portfolios; but the allocation for the MVTE optimal portfolio certainly appears to be a compromise between the two. The standard deviation and tracking error of the MVTE model is also certainly a compromise between those of the Markowitz and TE models.

A sharp criticism of the Markowitz and TE models was that they produced fairly extreme weightings on few of the assets in the opportunity set; and these select few assets differed when using the TE portfolio selection model rather than the Markowitz portfolio selection model. Arguably the MVTE model produces far more appealing optimal solutions. Importantly this is achieved without having to stipulate constraints for the maximum and minimum proportion of the portfolio that can be invested in any

single asset in the opportunity set. The MVTE model is certainly not applying any constraints but in contrast it takes account of the relative importance of return, risk and tracking error through the values of risk tolerance and tracking error tolerance.

Both the TE and MVTE portfolio selection models do take heed of the notion that the portfolio manager's performance is measured relative to a benchmark. The portfolio's tracking error captures the risk or variability about this benchmark. But the tracking error term in the objective function effectively penalises the portfolio manager even if he/she manages to outperform the benchmark! It is the Mean-Variance-Downside Tracking Error portfolio selection model that addresses this concern and we discuss this model in the next section.

University of Cape Town

## 5. MEAN-VARIANCE-DOWNSIDE TRACKING ERROR (DTE) PORTFOLIO SELECTION MODEL

The TE and MVTE portfolio selection models determine portfolios that jointly minimise returns below and above some benchmark. This implies that portfolio selection models based on a tracking error term have assets that have fewer months of abysmal performance relative to a benchmark as their optimal solution. Simultaneously they favour assets with fewer months of superb results relative to this benchmark (Roll 1992)! There is certainly a strong case for arguing that a portfolio manager's objective should be focused on minimising only the under performance of the benchmark (Harlow<sup>16</sup> 1991) which is the objective of the downside tracking error (DTE) portfolio selection model.

The tracking error (TE) was defined as the standard deviation of the relative return. Denoting the return on asset  $i$  at time  $t$  as  $R_{it}$  and the return on the benchmark as  $R_{bt}$  then the relative return at time  $t$  on security  $i$ ,  $RR_{it}$ , is written as

$$RR_{it} = R_{it} - R_{bt}$$

The downside relative return is an adjustment to this measure

$$\begin{aligned} \text{Downside } RR_{it} &= 0 && \text{for } R_{it} > R_{bt} \\ &= R_{it} - R_{bt} && \text{otherwise} \end{aligned}$$

The Downside Tracking Error (DTE) is the standard deviation of the downside relative returns.

The objective of the DTE portfolio selection model is to

$$\text{MAX Portfolio's Expected Return} - \frac{(\text{Expected Downside Tracking Error})^2}{\text{Downside Tracking Error Tolerance}}$$

Or using the notation defined earlier together with

$\Sigma^{**}$  = the variance-covariance matrix of the downside relative returns of the assets in the opportunity set

$$\psi^* = \frac{1}{\text{the Downside Risk Tolerance}}$$

$\sigma_p^{**2}$  = the variance of downside relative return on the portfolio where

$$\sigma_p^{**2} = X' \Sigma^{**} X$$

The objective of the DTE portfolio selection model is to

$$\text{MAX } E_p - \psi^* \sigma_p^{**2}$$

The empirical research performed by Harlow (1991) using the DTE portfolio selection model resulted in portfolios having less downside exposure than those determined using the Markowitz portfolio selection model ex post. Hence portfolio managers averse to below target return dispersions achieve a more attractive risk/return trade-off with the downside-risk framework. The asset allocation example considered by Harlow produced a significantly higher average bond allocation than the traditional Markowitz asset allocation model. Harlow argued that bonds increase downside protection while offering the same or greater levels of expected return. This may be true in Harlow's empirical research, but it is questionable whether bonds always offer greater downside protection. In order to make this claim one needs to examine the distribution of bond returns (and those of the other asset classes). Unfortunately Harlow did not perform this investigation! Harlow claims that if the distribution is symmetrical about the benchmark then the TE and DTE measures will be identical, and consequently TE will be an adequate measure of risk. As the return benchmark increases so more of the distribution falls below this specified benchmark. The downside component of returns consequently becomes larger, increasing the numerical value of this risk measure. Accompanying this shift is an increase in the allocation to stocks. Harlow merely states this as an observation; and he believes that

<sup>16</sup> Harlow, W, "Asset Allocation in a Downside-Risk Framework", *Financial Analysts Journal*, September-October 1991, pp. 28-40.

it is not an intuitive result but rather it "represents the complete interaction between the portfolio distribution of returns and the benchmark".

The objective of this thesis is to examine the various portfolio selection models that have been proposed. Thus far we have observed that the Markowitz model was certainly one of the first proposed and it has received much support. Numerous people viz. Roll, Franks and Harlow have proposed adjustments to the basic Markowitz model. Their emphasis has been on adjusting the model's measure of risk. They have proposed that TE (standard deviation of relative returns) and DTE (standard deviation of adjusted relative returns) could be better measures of risk in the context of the portfolio selection objective.

The simple objective of portfolio managers is to maximise a portfolio's return and simultaneously minimise its risk (however it may be measured). Thus far we have focused on the measurement of risk in portfolio selection models. We now focus on the measurement of return in these same models.

## 6. BAYESIAN ESTIMATES OF RETURN

One of the objectives of portfolio selection models is to maximise the portfolio's expected return. This is a parameter that needs to be estimated. Markowitz (1952), in his proposed portfolio selection model, simply used sample means of the securities' past returns as estimates for the securities' expected returns. Jobson<sup>17</sup> and Korkie (1980) found that at least 300 observations of monthly return data was necessary for these traditional estimators to produce an optimal weight vector that was reasonably comparable to its "true" value. The simulation conducted by Jobson and Korkie used a stationary distribution but in reality it is quite unlikely that the actual return data is being generated by a stationary distribution over an interval as long as 300 months! This lead Jobson and Korkie and others to consider different estimators of a security's expected return.

Statisticians claim that an estimator is "admissible" if no other estimator exists that dominates it for a given loss function. The loss function is defined as follows :

1. If the estimated value for the parameter equals the true parameter value then the loss is zero.
2. If the estimated value for the parameter does not equal the true parameter value then the loss is non-zero.

Under the condition of "admissibility" based on the Loss function, Stein<sup>18</sup> (1955) has shown that when the opportunity set contains two or more securities, sample means are not admissible estimators of expected returns. Sample means are sub-optimal because they ignore the inherent multivariate nature of the problem, which is the case when there are two or more securities in the opportunity set (Michaud 1989).

If sample means are not the best estimates for the securities' expected returns, then other estimators are needed. These "other" estimators could certainly be functions of

<sup>17</sup> Jobson, JP and Korkie, B, "Estimation for Markowitz Efficient Portfolios", *Journal of the American Statistical Association*, September 1980.

<sup>18</sup> Stein, C, "Inadmissibility of the Usual Estimate for the Mean of a Multivariate Distribution", *Proceedings of the 3<sup>rd</sup> Berkley Symposium on Probability and Statistics 1*, 1955, pp. 197-206.

various factors e.g. the debt/equity ratio might be considered to be a factor. Black (1995) is actually very sceptical of such estimators and he claims that when we have no theory for how a factor should be priced, past average returns might well give the best estimate, but he does admit that it is normally a highly inaccurate estimate.

An estimation approach, which has gained much support in the area of portfolio selection in the 1990's, is the Bayesian approach. The proponents of the Bayesian approach acknowledged that a sample mean often does not capture all the information present when estimating the true mean. They realised that the user of the estimate may have some personal expectations that are not captured by the sample. Consequently, the Bayesian estimate is a compromise between the sample mean and the user's own personal/subjective expectations. Each of these two components receives weighting. The more certain the user is in his/her expectation the greater the weight placed on this personal expectation. The sample mean can be considered to be the neutral view, i.e. if the user has no particular views then the sample mean is the estimate of the true mean. Similarly, if the user is totally certain in his/her own expectations then this personal expectation is the estimate of the true mean.

Black<sup>19</sup> and Litterman (1992) have applied the Bayesian estimate in the Markowitz portfolio selection model. Their Bayesian estimate is slightly different to the one proposed by the statisticians but it is based on the same principles. Black and Litterman claim that the neutral position would be better represented by a CAPM (Capital Asset Pricing Model) estimate than by a sample mean of past returns. Hence if the portfolio manager has no subjective expectations in the securities' expected returns then the CAPM estimates will be the neutral estimates for the securities' expected returns. If the portfolio manager does have one or more views about the performance of the securities in the opportunity set he/she can adjust the neutral view established by the CAPM equilibrium in accordance with those views. The portfolio manager can control how strongly a particular view influences portfolio weights in accordance with the degree of confidence with which he/she holds the view.

---

<sup>19</sup> Black, F and Litterman, R, "Global Portfolio Optimization", *Financial Analysts Journal*, September-October 1992, pp. 28-43.

Black and Litterman claim that the Markowitz portfolio selection model will produce extreme optimal solutions if one simply used sample means as estimates for the securities' expected returns. In contrast, using the CAPM estimates, one obtains the market portfolio as the optimal solution. The optimal solution is then drawn away from this equilibrium position as the portfolio manager stipulates his/her own expectations. In summary, Black and Litterman use the CAPM estimates to represent a portfolio manager's neutral views.

If one used the Bayesian estimate as proposed by the statisticians one would need to stipulate one's personal expectations for each security in the opportunity set (if one had no view one could simply state the sample mean as one's "personal" expectation). Black and Litterman have proposed a model where the portfolio manager need only stipulate his/her expectations for the securities for which he/she has these personal expectations. Although the portfolio manager has not stipulated a personal expectation for each and every security, all the securities' expected returns will nevertheless be affected by those personal expectations that are stated. This is because the model contains a common factor.

It is certainly difficult to judge a Bayesian estimate because a Bayesian estimate requires a set of personal expectations. Hence when one examines how the model performs in simulations one is not only testing the model but also the strategy from which one obtains the portfolio manager's subjective expectations.

Michaud (1989) supports the notion that Bayes estimation can significantly improve the Markowitz portfolio selection model. The Bayesian estimator Michaud discussed differs to the one proposed by Black and Litterman in that a pooled mean is used as the neutral expectation, rather than the CAPM expectation. Observed sample means, rather than portfolio managers' subjective expectations, are then "shrunk" to this pooled mean. The degree of shrinkage is a function of the samples' standard deviations and the pooled mean's own standard deviation. The smaller the sample standard deviation, the greater is the weight that is placed on that sample mean. Unfortunately, Michaud did not support his discussion with any empirical research.

Jorion<sup>20</sup> (1986) also examined a Bayesian estimate based on a pooled mean. His justification for using a pooled sample mean as the neutral estimate is that it will have a smaller estimation error than the individual sample estimates. Jorion did examine the Bayesian estimates empirically. He compared the Bayesian estimate to the "traditional" sample mean estimate. The comparison is based on a loss function. As stated earlier, the loss function is defined as follows:

1. If the estimated value for the mean equals the true mean then the loss is zero.
2. If the estimated value for the mean does not equal the true mean then the loss is non-zero.

Jorion found that by applying the *negative exponential* loss function the Bayesian estimate's loss was less than the loss obtained when using the traditional sample mean estimate. The two losses do approach one another in magnitude when the sample size increases and nears 200 (Jorion 1986). As already mentioned, Jobson and Korkie (1980) are very sceptical about merely increasing the sample size since it is quite unlikely that the actual return data is being generated by a stationery distribution over an interval of 200 months! One must nevertheless be wary of Jorion's results since it is the task of the practitioner to stipulate the loss function (Jorion applied the negative exponential loss function). We believe one might well obtain contradictory results if one were to use a different loss function. Jorion does justify his results by stating that he also considered a quadratic loss function, which essentially gave the same results. Nevertheless, there are an infinite number of loss functions that one could actually use.

A Bayesian estimate is a subjective measure. One needs the portfolio manager to stipulate his/her personal expectations. In essence one is not merely measuring the estimate itself but also the portfolio manager's knowledge and his/her ability to state this knowledge. The whole concept of Bayesian estimates has come under much criticism from many "traditionalists" but there are also those statisticians who believe that the Bayesian concept can more realistically represent the expectations portfolio managers have for the returns on the securities.

---

<sup>20</sup> Jorion, P, "Bayes-Stein Estimation for Portfolio Analysis, *Journal of Financial and Quantitative Analysis*, September 1986, pp.279-292.

**CHAPTER 3**

**THE OBJECTIVE FUNCTION**

University of Cape Town

## 1. SYNOPSIS

Several strands of research have emerged which generally yield portfolios that tend to be more satisfactory to practitioners than the traditional implementation of the Market Model. One such strand of research focuses on the realism of the objective function of asset allocation. A key component of this objective function is the appropriate measure of the portfolio's risk. Several measures have been proposed which shall be examined empirically in the context of the Johannesburg Stock Exchange in this chapter. Some refinements to these measures will be considered in this chapter.

In recent times much emphasis has been placed on the evaluation of funds relative to a benchmark. One important objective of contemporary fund managers that is not captured by the traditional mean-variance objective function formulation of Markowitz<sup>1</sup> (1959), therefore, is the consideration of the risk of the portfolio departing from the benchmark. In Section 2 we give a background to the problem by first describing the well known Markowitz mean-variance asset allocation model. We use this as an opportunity to give an outline of the necessary notation conventionally used in the implementation of asset allocation models. In Section 3 we give an empirical demonstration of the implementation of the Markowitz mean-variance approach using the major indices of the JSE opportunity set. In this section Sharpe's<sup>2</sup> (1970) step by step procedure of constructing the efficient frontier is followed. In Section 3.3 we introduce the risk-free asset into the opportunity set and examine the consequences of this action both from a graphical and portfolio composition perspective using the South African data set.

For many investors however the mean-variance model fails to provide satisfactory results in the form of realistic portfolios. Franks<sup>3</sup> (1992) acknowledged the practical shortcomings of the mean-variance model and he proposed instead that the true objective of portfolio managers should be to minimise the deviations from a benchmark. We label Franks' (1992) model as the tracking error (TE) model and demonstrate it empirically in Section 4 using our South African data set. It is

---

<sup>1</sup> Markowitz, HM, Portfolio Selection: Efficient Diversification of Investment", 1959.

<sup>2</sup> Sharpe, WF, "Portfolio Theory and Capital Markets", 1970.

<sup>3</sup> Franks, ECF, "Targeting Excess-of-Benchmark Returns", *The Journal of Portfolio Management*, Summer 1992.

interesting to see the impact of the risk-free asset when viewed in a framework that includes tracking error. In such a framework the risk-free asset relative to a benchmark will no longer be a constant time series and consequently will not dominate portfolios to the same extent as in a framework which includes variance (where its own variance is constant). Franks' tracking error model does not however take any account of the variance of return as measured by the traditional mean variance (MV) model. Chow<sup>4</sup> (1995) by contrast considers both models by proposing a joint model that takes account of both the variance and tracking error of a portfolio in the objective function resulting in a 3 dimensional feasible region. This model is referred to as the mean variance tracking error model (MVTE). Section 6 demonstrates this MVTE model on the South African data set and in Section 7 we show that the composition of a MVTE optimal portfolio is essentially a compromise of the compositions of the equivalent underlying mean-variance and tracking error optimal portfolios. Essentially the MV and TE models are each a special case of the MVTE model. Finally in Section 9 we demonstrate a model where the tracking error is curtailed by the sign of the relative return i.e. whether it is positive or negative. We refer to this model as the *downside tracking error* (DTE) model. We argue that fund managers may not want to minimise performance in excess of the benchmark and we empirically demonstrate the inclusion of this view in the efficient frontier generating objective function.

Note that readers who are familiar with the implementation of the traditional Markowitz mean-variance optimisation procedure can essentially skip Sections 2 and 3.

---

<sup>4</sup> Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995.

## 2. INTRODUCTION<sup>5</sup>

In the investment world an investor selects a single portfolio from an infinite set of possibilities. For example, if an investor is faced with the option of creating a portfolio from only two assets, say asset A and asset B, then the investor can invest anything from 0% to 100% in either of the two assets. Effectively the opportunity set of portfolios is infinite because proportions are continuous variables. Although there is an infinite set of possibilities there may just be a single portfolio that suits each investor's particular objective. Markowitz<sup>6</sup> (1959) described an optimisation approach that has become well known to scholars in the field of finance. The Markowitz model assumes that an investor seeks portfolios with **high expected returns** and **low expected risks** or *uncertainties*. Effectively the investor has two goals:

1. to maximise return.
2. to minimise the uncertainties of the portfolio's return, which can be measured by the variance of the portfolio.

As a consequence of these two goals Markowitz described how an *efficient frontier* could be generated by combining these two goals into one objective function which simultaneously maximises return for minimum risk. Such an objective function can be formulated as:

$$MAX \text{ Expected Return} - \frac{\text{Expected Variance}}{\text{Risk Tolerance}}$$

The risk tolerance sets the price of variance in terms of return. Sharpe<sup>7</sup> showed how varying the risk tolerance incrementally yielded an algorithmic approach to generating the efficient frontier.

<sup>5</sup> Readers who are familiar with the implementation of the traditional Markowitz mean-variance optimisation procedure can essentially skip Sections 2 and 3.

<sup>6</sup> Markowitz, HM, "Portfolio Selection: Efficient Diversification of Investment", 1959.

<sup>7</sup> Sharpe, WF, "Portfolio Theory and Capital Markets", 1970.

## 2.1 Notation used for calculations

The following notation will be used for all the calculations that we perform:

$E_i$  = Expected return on the  $i^{\text{th}}$  security

$\sigma_i$  = Standard deviation of return on the  $i^{\text{th}}$  security

$E_p$  = Expected return on the portfolio

$\sigma_p$  = Standard deviation on the portfolio

$\sigma_{ij}$  = Covariance between security  $i$  and security  $j$

$\rho_{ij}$  = Correlation coefficient for the returns on securities  $i$  and  $j$

$X_i$  = Proportion of funds invested in security  $i$

$n$  = Total number of shares considered

also

$$E_p = \sum_{i=1}^n X_i E_i = X'E$$

and

$$\begin{aligned}\sigma_p^2 &= \sum_{i=1}^n \sum_{j=1}^n X_i X_j \sigma_{ij} \\ &= \sum_{i=1}^n \sum_{j=1}^n X_i X_j \sigma_i \sigma_j \rho_{ij} \\ &= X' \Sigma X\end{aligned}$$

where

$$X = \begin{pmatrix} X_1 \\ \vdots \\ \vdots \\ X_n \end{pmatrix}$$

$$E = \begin{pmatrix} E_1 \\ \vdots \\ E_n \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \dots & \sigma_n^2 \end{pmatrix}$$

The Markowitz Mean-Variance (MV) objective function for generating an efficient frontier can now be written as follows:

$$\text{MAX } E_p - \lambda \cdot \sigma_p^2$$

or in matrix notation

$$\text{MAX } X'E - \lambda X'\Sigma X$$

Where  $\lambda = \frac{1}{\text{risk tolerance}}$

It was Sharpe (1970) who first considered this objective function. The objective function is usually accompanied by the following two constraints:

1.  $\sum_{i=1}^n X_i = 1$  that is, the entire portfolio must be invested.
2.  $X_i \geq 0$  for all  $i = 1, 2, \dots, n$ . That is no securities may be held in negative quantities thus no short sales are permitted.

A portfolio is represented by a vector  $X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$  where  $X_i$  is the proportion of funds

invested in the  $i^{\text{th}}$  security. The value of the objective function thus changes for varying values of  $X_i$ .

An inefficient approach practised in the late 1960's was to allocate random proportions to  $X_i$  and then observe which allocation maximises  $X'E - \lambda X'\Sigma X$ . A more effective and efficient solution can be attained by using the following steps (Sharpe<sup>8</sup>, 1970):

1. assign a particular numerical value between 0 and  $\infty$  to  $\lambda$ ;
2. solve for the  $X_i$ 's
3. change the value of  $\lambda$  and repeat

<sup>8</sup> The solution requires a quadratic programming technique since the function which one is maximising contains terms in  $X_i^2$ . A technique that can be used and which has been used in the empirical problem is Newton's method.

### 3. GENERATING THE EFFICIENT FRONTIER USING THE MARKOWITZ (MV) MODEL<sup>9</sup>

In this section we demonstrate the implementation of the Markowitz model on a data set consisting of the major indices on the JSE.

#### 3.1 Data

One approach to assess the practical implementation of the model is to inspect the composition of the optimal portfolios and to assess whether the resulting portfolios are similar to those typically constructed by portfolio managers in practice. The data set consisted of the seven second-tier indices of the Johannesburg Stock Exchange (JSE). The seven indices can be considered as a representation of the cross-structure of the JSE All Share Index. An estimate for the expected returns for each of these assets was required. The average monthly return for each asset covering the period January 1992 to December 1996 (5 years) was used in the estimation of inputs. The seven indices that were considered are listed in Table 1 on page 7 together with their average monthly returns.

Table 1 Average monthly returns of the constituent indices

| INDEX                   | Proportion of Overall Index 1992 | Proportion of Overall Index 1996 | AVERAGE MONTHLY RETURN |
|-------------------------|----------------------------------|----------------------------------|------------------------|
| Coal Index              | 0.82%                            | 1.42%                            | 0.01823                |
| Diamond Index           | 8.27%                            | 4.88%                            | 0.009605               |
| All Gold Index          | 18.98%                           | 6.51%                            | 0.007329               |
| Metals & Minerals Index | 6.73%                            | 3.21%                            | 0.006646               |
| Mining Financial Index  | 23.92%                           | 15.87%                           | 0.013337               |
| Financial Index         | 7.41%                            | 17.38%                           | 0.018482               |
| Industrial Index        | 33.87%                           | 50.54%                           | 0.010939               |

The proportions do not sum to 100% because of rounding errors

From Table 1 it is evident that over our 5-year period of investigation (1992-1996) the component indices used did not have constant weights in the overall index. As we will argue later the portfolio construction techniques are therefore not able to *precisely* reconstruct the Overall Index (which is used as a benchmark in Sections 4, 6, 7 and

<sup>9</sup> Readers who are familiar with the implementation of the traditional Markowitz mean-variance optimisation procedure can essentially skip Section 3.

10) with a *constant* set of investment weights. This has various implications which will be discussed later! In some of the analysis we include the benchmark in the opportunity set of assets to assess the graphical consequences of being able to reconstruct the benchmark!

To commence the demonstration we consider the average absolute monthly returns of the seven South African indices as an estimate for their expected returns. The expected returns vector  $E$  would be:

$$E = \begin{pmatrix} 0.01823 \\ 0.009605 \\ 0.007329 \\ 0.006646 \\ 0.013337 \\ 0.018482 \\ 0.010939 \end{pmatrix}$$

and the covariance matrix  $\Sigma$  is:

$$\Sigma = \begin{pmatrix} 0.0079 & 0.0017 & 0.0032 & 0.0032 & 0.0026 & 0.0006 & 0.0011 \\ & 0.0061 & 0.0044 & 0.0023 & 0.0035 & 0.0016 & 0.0016 \\ & & 0.0123 & 0.0057 & 0.0059 & 0.001 & 0.0011 \\ & & & 0.006 & 0.004 & 0.001 & 0.0015 \\ & & & & 0.0045 & 0.0013 & 0.0016 \\ & & & & & 0.0021 & 0.0015 \\ & & & & & & 0.0017 \end{pmatrix}$$

Following the traditional implementation approach we shall apply the expected return matrix,  $E$ , and the covariance matrix,  $\Sigma$ , to the Markowitz Model.

### 3.2 A Demonstration of the Markowitz (MV) Model

As stated before, the objective of the Markowitz Model (MV) is to maximise the objective function, i.e.

$$\text{MAX } X'E - \lambda X'\Sigma X$$

Firstly, by setting  $\lambda = 0$ , the following solution was achieved for our South African data set

$$X = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}$$

The solution suggests that in order to maximise the objective function where  $\lambda = 0$  all the investor's funds (100%) should be invested in the FINANCIAL index. The achieved return for such a portfolio was:

$$X'E = \begin{pmatrix} 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} 0.01823 \\ 0.009605 \\ 0.007329 \\ 0.006646 \\ 0.013337 \\ 0.018482 \\ 0.010939 \end{pmatrix} = 0.018482$$

In the situation where  $\lambda = 0$  the investor is effectively maximising the return of the portfolio with no concern for risk (variance). This is achieved by investing all the funds in the asset that offers the highest expected return irrespective of its variance. This asset, as already indicated, is the FINANCIAL index and forms the point of highest return on the efficient frontier for the period examined. The variance of the portfolio created by investing all ones funds in the FINANCIAL index would simply

be the variance of the FINANCIAL index. This is evident when examining the calculations of the portfolio's variance:

$$X' \Sigma X = \begin{pmatrix} 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} 0.0079 & 0.0017 & 0.0032 & 0.0032 & 0.0026 & 0.0006 & 0.0011 \\ & 0.0061 & 0.0044 & 0.0023 & 0.0035 & 0.0016 & 0.0016 \\ & & 0.0123 & 0.0057 & 0.0059 & 0.001 & 0.0011 \\ & & & 0.006 & 0.004 & 0.001 & 0.0015 \\ & & & & 0.0045 & 0.0013 & 0.0016 \\ & & & & & 0.0021 & 0.0015 \\ & & & & & & 0.0017 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} = 0.0021$$

The risk of the portfolio which is measured by the standard deviation of the portfolio would be  $\sqrt{0.0021} = 0.04583$  which is simply the standard deviation of the FINANCE index. The expected risk (standard deviation) and expected return of this optimal portfolio can be plotted in 2 dimensional space. It is represented in Figure 1 on page 10 and is denoted A:

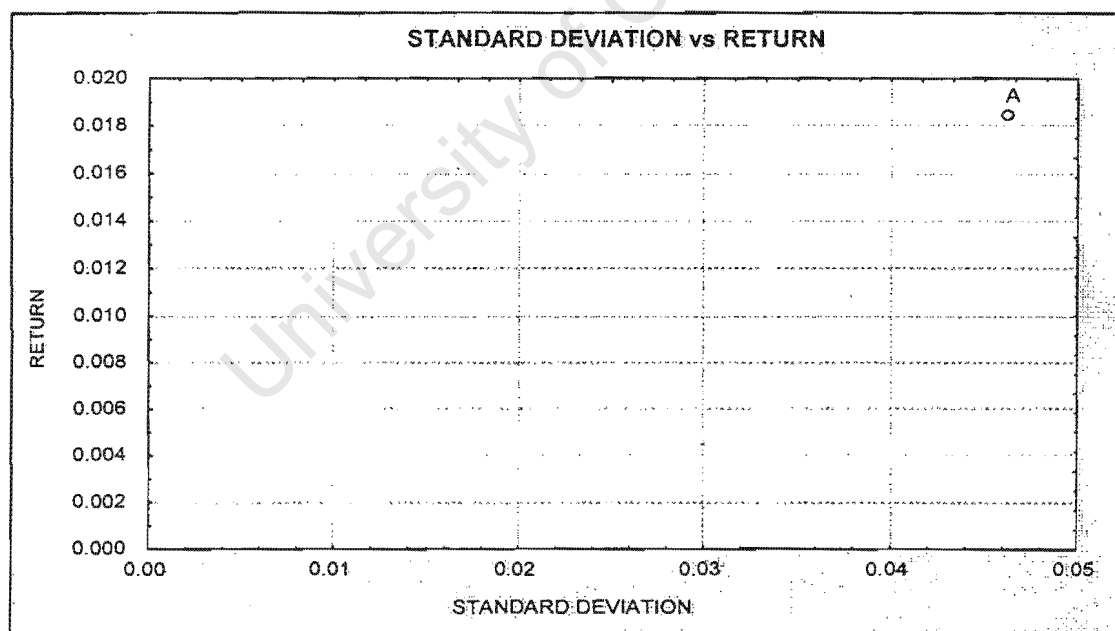


**Figure 1 Plot of the portfolio which obtains the maximum possible expected return based on the 1992-1996 data**

The other extreme value for  $\lambda$  that one could consider is  $\lambda = \infty$ . Effectively one is obtaining the minimum possible standard deviation of the portfolio with no consideration for the expected return of the portfolio. This will clearly not be obtained by simply investing all ones funds in the asset that has the minimum standard

deviation since this value can be reduced further by means of diversification within the opportunity set. The objective function (with  $\lambda = \infty$ ) yielded the following solution:

$$X = \begin{pmatrix} 0.0887 \\ 0 \\ 0.0323 \\ 0 \\ 0 \\ 0.2781 \\ 0.6009 \end{pmatrix}$$

It is clear that only four of the seven assets had non-zero weighting. The expected return for this optimal portfolio would be:

$$\begin{aligned} X'E &= (0.0887 \ 0 \ 0.0323 \ 0 \ 0 \ 0.2781 \ 0.6008)E \\ &= 0.0136 \\ &= 1.36\% \text{ per month} \end{aligned}$$

The variance of the portfolio would be calculated as follows:

$$\begin{aligned} X'\Sigma X &= (0.0887 \ 0 \ 0.0323 \ 0 \ 0 \ 0.2781 \ 0.6008)\Sigma \begin{pmatrix} 0.0887 \\ 0 \\ 0.0323 \\ 0 \\ 0 \\ 0.2781 \\ 0.6008 \end{pmatrix} \\ &= 0.0016 \end{aligned}$$

The risk (standard deviation) of the portfolio is  $\sqrt{0.0016} = 0.04 = 4\%$  per month.

Similarly for the case where  $\lambda = 0$ , the risk and return of this portfolio can be plotted in two-dimensional space. Figure 2 on page 12 shows a plot of the optimal solutions obtained by setting  $\lambda = 0$  and by setting  $\lambda = \infty$  respectively.

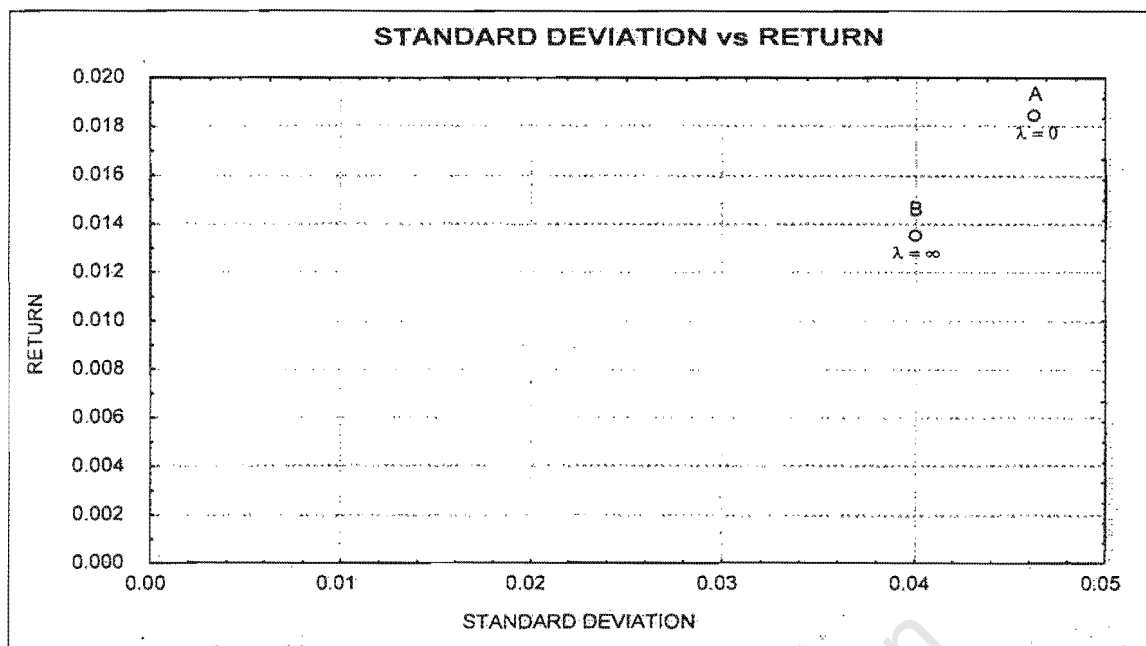


Figure 2 Optimal portfolios where  $\lambda = 0$ , and where  $\lambda = \infty$  based on the 1992-1996 data

Now the full optimal set can be achieved by setting  $\lambda$  equal to all the values between 0 and  $\infty$ . The expected return and the expected standard deviation can be calculated for each optimal portfolio once the investment weights ( $X_i$ ) have been established using the procedure described earlier. These calculations were performed and the expected returns and the expected standard deviations of these optimal portfolios have been plotted in two-dimensional space:

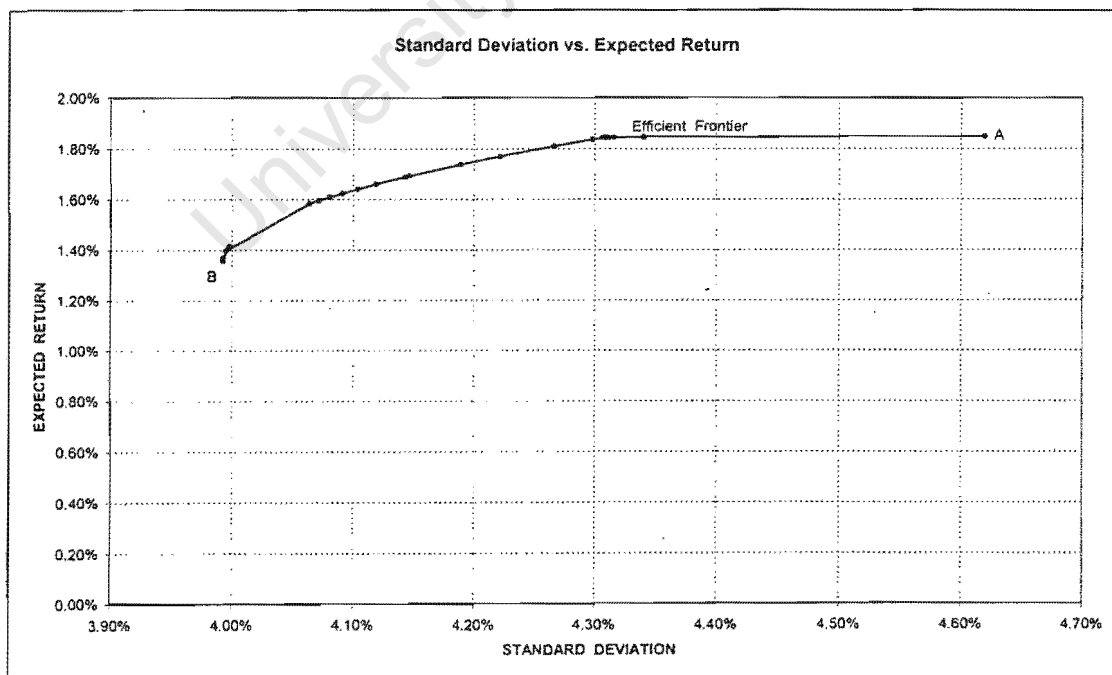


Figure 3 Plot of the efficient frontier based on the 1992-1996 JSE data. Information concerning these efficient portfolios is listed in Table 2 on page 14

The plots of the returns and risks (standard deviations) of the optimal portfolios form a solid curve in two-dimensional space. This curve is termed the "efficient frontier". A portfolio is said to be "efficient" if it is impossible to find a portfolio which has a greater expected return without incurring greater risk, also, it is impossible to obtain smaller risk without decreasing expected return. Each of the optimal portfolios obtained using the Markowitz mean-variance (MV) optimisation procedure is an efficient portfolio. Investors will obviously always prefer efficient portfolios, however, the higher the expected return of an efficient portfolio the higher is the associated level of risk. It is for this reason that the efficient frontier will either be positively linear or convex but never concave. This is evident when one examines Figure 3 on page 12.

Table 2 on page 14 summarises the efficient portfolios generated by solving the objective function each time for various incremental changes in  $\lambda$  as specified in the first column of the table.

**Table 2** Efficient portfolios for the specified  $\lambda$  values based on the JSE Indices covering the period 1992-1996

| $\lambda$ | RETURN  | STD     | PROPORTIONS |      |      |      |      |      |      |
|-----------|---------|---------|-------------|------|------|------|------|------|------|
|           |         |         | coal        | diam | gold | metm | mfin | fina | indu |
| 1500      | 0.01357 | 0.03992 | 9%          | 0%   | 3%   | 0%   | 0%   | 28%  | 60%  |
| 500       | 0.01365 | 0.03992 | 9%          | 0%   | 3%   | 0%   | 0%   | 29%  | 59%  |
| 400       | 0.01365 | 0.03992 | 9%          | 0%   | 3%   | 0%   | 0%   | 29%  | 59%  |
| 300       | 0.01369 | 0.03993 | 9%          | 0%   | 3%   | 0%   | 0%   | 29%  | 59%  |
| 200       | 0.01415 | 0.03997 | 10%         | 0%   | 3%   | 0%   | 0%   | 34%  | 53%  |
| 100       | 0.01399 | 0.03995 | 10%         | 0%   | 3%   | 0%   | 0%   | 32%  | 55%  |
| 20        | 0.01584 | 0.04064 | 13%         | 0%   | 1%   | 0%   | 0%   | 53%  | 33%  |
| 19        | 0.01596 | 0.04071 | 13%         | 0%   | 1%   | 0%   | 0%   | 54%  | 31%  |
| 18        | 0.01610 | 0.04080 | 13%         | 0%   | 1%   | 0%   | 0%   | 56%  | 30%  |
| 17        | 0.01625 | 0.04091 | 14%         | 0%   | 1%   | 0%   | 0%   | 58%  | 28%  |
| 16        | 0.01642 | 0.04103 | 14%         | 0%   | 1%   | 0%   | 0%   | 59%  | 26%  |
| 15        | 0.01661 | 0.04119 | 14%         | 0%   | 1%   | 0%   | 0%   | 62%  | 23%  |
| 14        | 0.01689 | 0.04142 | 15%         | 0%   | 0%   | 0%   | 0%   | 65%  | 20%  |
| 13        | 0.01694 | 0.04146 | 15%         | 0%   | 0%   | 0%   | 1%   | 65%  | 19%  |
| 12        | 0.01739 | 0.04189 | 16%         | 0%   | 0%   | 0%   | 1%   | 70%  | 13%  |
| 11        | 0.01771 | 0.04221 | 16%         | 0%   | 0%   | 0%   | 2%   | 74%  | 9%   |
| 10        | 0.01811 | 0.04266 | 17%         | 0%   | 0%   | 0%   | 2%   | 78%  | 3%   |
| 9         | 0.01837 | 0.04298 | 17%         | 0%   | 0%   | 0%   | 1%   | 82%  | 0%   |
| 8         | 0.01843 | 0.04306 | 17%         | 0%   | 0%   | 0%   | 0%   | 82%  | 0%   |
| 7         | 0.01844 | 0.04307 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 6         | 0.01844 | 0.04307 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 5         | 0.01844 | 0.04307 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 4         | 0.01844 | 0.04307 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 3         | 0.01844 | 0.04307 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 2         | 0.01844 | 0.04308 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 1         | 0.01844 | 0.04309 | 16%         | 0%   | 0%   | 0%   | 0%   | 84%  | 0%   |
| 0.75      | 0.01844 | 0.04311 | 16%         | 0%   | 0%   | 0%   | 0%   | 84%  | 0%   |
| 0.5       | 0.01844 | 0.04315 | 15%         | 0%   | 0%   | 0%   | 0%   | 85%  | 0%   |
| 0.25      | 0.01845 | 0.04340 | 12%         | 0%   | 0%   | 0%   | 0%   | 88%  | 0%   |
| 0         | 0.01848 | 0.04619 | 0%          | 0%   | 0%   | 0%   | 0%   | 100% | 0%   |

These results indicate that the efficient portfolios over the period 1992-1996 never consisted of more than four of the seven indices considered. Even more specifically no efficient portfolio consisted of more than 3% invested in the GOLD index. The fact that only four out of seven indices were selected is not consistent with the types of compositions managers of funds find typically appealing in South Africa (i.e. diversified). It is for this reason that many investors reject the traditional implementation of the Markowitz mean-variance (MV) procedure and consciously select alternative portfolios (Chow,<sup>10</sup> 1995). Our results in the South African context are also damaging for the proponents of the traditional mean-variance implementation

<sup>10</sup> Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995, pp. 54-60.

approach. As a consequence attempts have been made to obtain greater realism in the solutions.

One area that is worth considering is the modification of the inputs via a Bayesian approach in accordance with Black<sup>11</sup> and Litterman (1992). Another, but not independent approach is to modify the objective function. We look at the latter here. It was also partly for this reason that Chow (1995) examined the tracking error (TE) model which we will discuss in Section 4.

### **3.3 The inclusion of cash in the opportunity set**

Portfolio managers typically have cash, a risk free asset, included in their opportunity set. Figure 4 is a plot showing the inclusion of the risk-free asset i.e. cash in the opportunity set. The opportunity set of the lower efficient frontier consists of varying combinations of only the seven JSE indices (risky assets) whereas the other efficient frontier (the Capital Market Line (CML)) comprises of one particular combination of the seven JSE indices together with varying combinations of the risk-free asset ( $r_f$ ). The Treasury Bill had an average monthly return of 0.087% over the period 1992-1996 and it has been used as an estimate for the risk-free rate.

---

<sup>11</sup> Black, F and Litterman, R, "Global Portfolio Optimization", *Financial Analysts Journal*, September-October 1992, pp. 28-42.

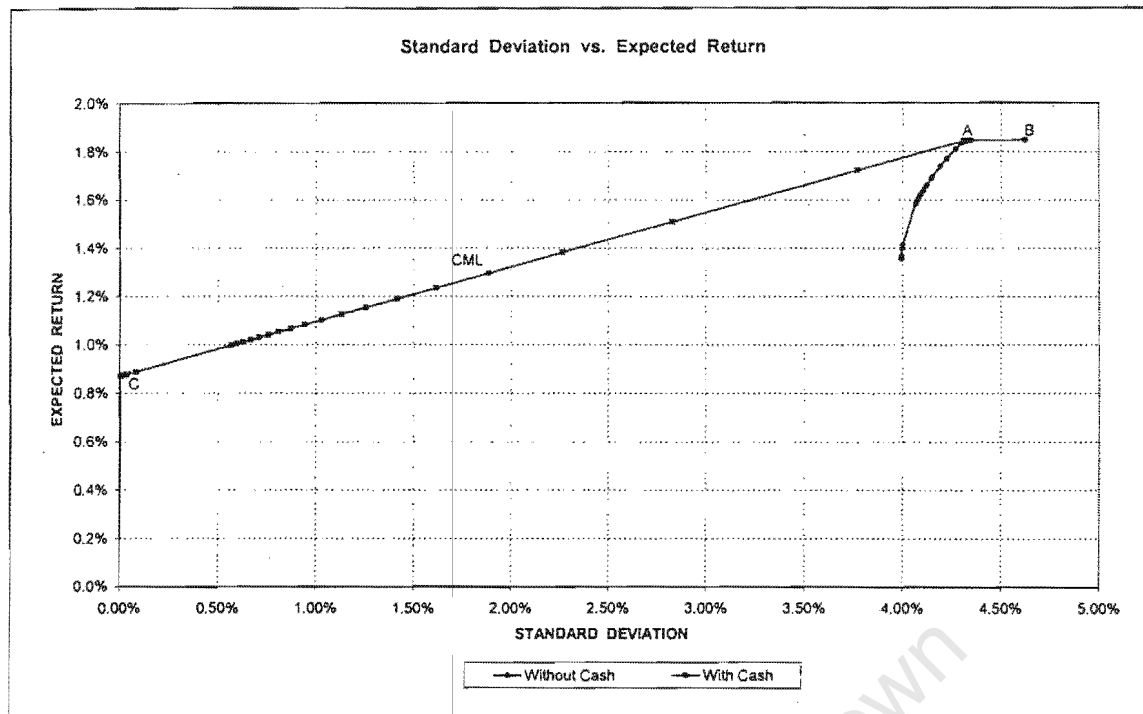


**Figure 4** The MV efficient frontier of an opportunity set containing cash together with the MV efficient frontier of an opportunity set that does not contain cash in the opportunity set based on the JSE indices spanning the period 1992-1996. The compositions of a few of these optimal portfolios are listed in Table 3

The Capital Market Line (containing the risk-free asset) generally spans a larger return interval, as well as a larger risk interval than the efficient frontier that does not include the risk-free asset. Note also the relationship between expected risk and expected return over the interval A to C in Figure 4 is linear. Over the interval A to B no cash is contained in the optimal portfolios. It is evident that the return obtained at point A can be increased marginally as one approaches point B but the associate risk increases significantly over the interval A to B. If we lifted the restriction of no borrowing then the Capital Market Line would extend further (to infinity).

The compositions of a selection of optimal MV portfolios (for a given return) obtained from the opportunity set that includes the risk-free rate ( $r_f$ ), and the composition of a selection of optimal MV portfolios from the opportunity set that is identical except it does not include a risk-free asset ( $r_f$ ) are listed in Table 3.

**Table 3** Comparative MV portfolio compositions with and without the risk-free asset at specified expected return levels based on the JSE data spanning the period 1992-1996

| CHARACTERISTICS           | TECHNIQUE     |            |               |            |
|---------------------------|---------------|------------|---------------|------------|
|                           | Without $R_f$ | With $R_f$ | Without $R_f$ | With $R_f$ |
| <i>Expected Return</i>    | 1.4%          | 1.4%       | 1.7%          | 1.7%       |
| <i>Standard Deviation</i> | 4%            | 2.4%       | 4.2%          | 3.7%       |
| <b>Allocation</b>         |               |            |               |            |
| Coal Index                | 10%           | 9%         | 15%           | 15%        |
| Diamond Index             | 0%            | 0%         | 0%            | 0%         |
| All Gold Index            | 3%            | 0%         | 0%            | 0%         |
| Metals & Minerals Index   | 0%            | 0%         | 0%            | 0%         |
| Mining Financial Index    | 0%            | 0%         | 0%            | 0%         |
| Financial Index           | 32%           | 45%        | 67%           | 71%        |
| Industrial Index          | 55%           | 0%         | 17%           | 0%         |
| Risk-free Asset ( $R_f$ ) |               | 46%        |               | 14%        |

### 3.4 Insights for Practitioners

- The optimal portfolios we achieve when we implement the Markowitz mean-variance optimisation procedure using historic returns as estimates for the securities' expected returns are not very appealing. The reason for these optimal portfolios being unattractive to many investors is because they are *not* very diversified.
- The optimal portfolio that has the least quantity of risk (i.e. variance) is the one that consists solely of cash.
- The Capital Market Line (CML), which has cash in its opportunity set, spans a wider risk and return interval than one that does not have cash in its opportunity set.

#### 4. GENERATING THE EFFICIENT FRONTIER USING A TRACKING ERROR (TE) MODEL

In this section we focus on the measurement of risk. We argue that managers of portfolios are concerned about the performance of the portfolio relative to a benchmark. Such a measurement is termed the *relative risk (relative variance)* of the investment. In the traditional Markowitz procedure the variance of the portfolio is more specifically a measurement of the *absolute risk*. Chow (1995) proposes that the quantity being maximised in the Markowitz objective function is not the true measure of what managers actually wish to maximise. He suggests that perhaps investors are more concerned about comparing the performance of the portfolio against a benchmark rather than purely minimising the absolute risk of the portfolio. In the next section (Section 5) we bring back the notion of absolute risk (variance) and combine it together with the tracking error. However in this section we consider the proposal of replacing variance with tracking error in the formulation of the objective function. We define tracking error as the standard deviation of *relative returns*.

Measuring the performance of the portfolio against the performance of the benchmark instead captures the notion that investors would be uncomfortable with large deviations from the benchmark. Franks proposed this approach in 1992. Frank's model typically assumes the investor wishes to maximise expected return and minimise *relative variance*. In this case the objective is<sup>13</sup>:

$$\text{MAX Expected Return} - \frac{(\text{Expected Tracking Error})^2}{\text{Tracking Error Tolerance}}$$

The tracking error tolerance sets the price of tracking error in terms of the return. For simplification  $\psi$  ( $\psi$ ) will be used to represent  $\frac{1}{\text{Tracking Error Tolerance}}$ . The objective can be written as maximising the expression:

$$\text{MAX Expected Return} - \psi (\text{Expected Tracking Error})^2$$

<sup>13</sup> Franks actually proposed maximising the excess expected return (beyond the benchmark) not the expected return itself. Nevertheless, the composition of the optimal portfolios using either objective function will be identical.

To demonstrate the implementation of this approach empirically the JSE-Overall Index was taken as the benchmark. Denoting the return of asset,  $i$ , at time  $t$  as  $R_{it}$  and the return of the JSE Overall Index as  $R_{bt}$  then the *relative return* at time  $t$  for asset  $i$  ( $RR_{it}$ ) is written as:

$$RR_{it} = R_{it} - R_{bt}$$

Let  $\Sigma^*$  represent the variance-covariance matrix of the relative returns ( $RR_{it}$ 's) of the assets in the opportunity set. Effectively the objective of the Franks' Tracking Error (TE) model is to maximise the expression:

$$\text{MAX } X'E - \psi X'\Sigma^*X$$

As with the Markowitz mean-variance model (MV) this objective function can be maximised by

1. assigning a particular numerical value between 0 and  $\infty$  to  $\psi$
2. solve for the  $X_i$ 's
3. change the value of  $\psi$  and repeat

In the empirical example being considered here:

$$E = \begin{pmatrix} 0.01823 \\ 0.009605 \\ 0.007329 \\ 0.006646 \\ 0.013337 \\ 0.018482 \\ 0.010939 \end{pmatrix}$$

and the matrix  $\Sigma^*$  in our empirical example is:

$$\Sigma^* = \begin{pmatrix} 0.0066 & -0.0004 & 0.0003 & 0.0011 & 0.0003 & -0.0006 & -0.0001 \\ & 0.003 & 0.0005 & -0.0007 & 0.0003 & -0.0004 & -0.0005 \\ & & 0.0075 & 0.0019 & 0.0019 & -0.0019 & -0.0018 \\ & & & 0.0031 & 0.0009 & -0.001 & -0.0005 \\ & & & & 0.0011 & -0.0009 & -0.0006 \\ & & & & & 0.0012 & 0.0005 \\ & & & & & & 0.0005 \end{pmatrix}$$

The optimal solutions were computed and the returns and **tracking errors**<sup>14</sup> (not standard deviations) of each optimal solution have been plotted in two-dimensional space.

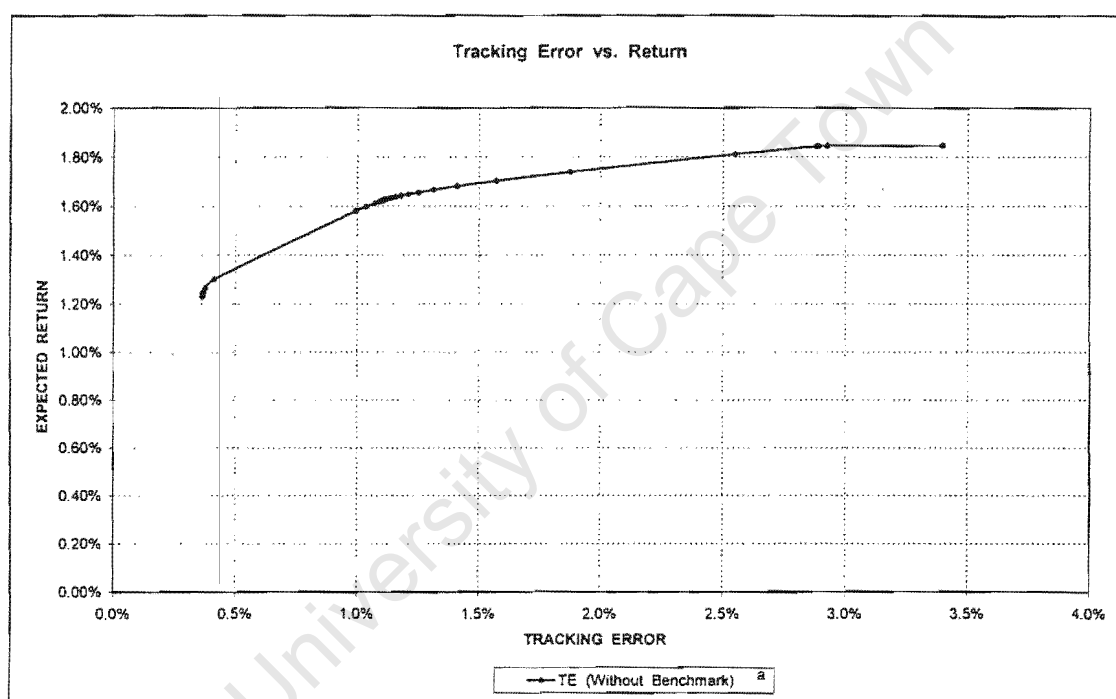


**Figure 5** Plot of the tracking-error efficient frontier based on the JSE data spanning the period 1992-1996. The compositions of these optimal portfolios are listed in Table 4

<sup>a</sup> The expression "Without Benchmark" in parenthesis above indicates that the benchmark itself is excluded from the opportunity set of assets.

From Figure 5 it is evident that the plots of the returns and tracking errors of the optimal portfolios will form a solid curve in two-dimensional space. This curve is termed the "efficient frontier" with respect to the tracking error. A portfolio is said to be *tracking error efficient* if it is impossible to find a portfolio that has a greater

<sup>14</sup> Recall with mean-variance approach of Section 3 the standard deviation (not variance) is typically used in graphical representations of Markowitz theory. Similarly we used the tracking error (not the tracking error squared) here.

expected return without incurring a greater tracking error. It is also impossible to obtain a smaller tracking error without decreasing the expected return. Each of the optimal portfolios is an efficient portfolio. As is evident in Figure 5 the efficient frontier is either convex or positively linear.

Table 4 summarises the numerical results of optimal portfolios obtained when setting  $\psi$  to the corresponding value specified in the first column of the table.

**Table 4 TE efficient portfolios achieved for incremental values of  $\psi$ . The JSE data spans the period 1992-1996**

| $\psi$ | RETURN  | TE      | PROPORTIONS |      |      |      |      |      |      |
|--------|---------|---------|-------------|------|------|------|------|------|------|
|        |         |         | coal        | diam | gold | metm | mfin | fina | indu |
| 3500   | 0.01229 | 0.00363 | 1%          | 7%   | 8%   | 4%   | 22%  | 17%  | 42%  |
| 2500   | 0.01230 | 0.00363 | 1%          | 7%   | 8%   | 4%   | 22%  | 17%  | 42%  |
| 1500   | 0.01232 | 0.00364 | 1%          | 7%   | 8%   | 4%   | 22%  | 17%  | 42%  |
| 500    | 0.01242 | 0.00365 | 1%          | 7%   | 7%   | 3%   | 22%  | 18%  | 41%  |
| 400    | 0.01246 | 0.00367 | 1%          | 6%   | 7%   | 3%   | 23%  | 19%  | 41%  |
| 300    | 0.01252 | 0.00369 | 1%          | 6%   | 7%   | 3%   | 23%  | 19%  | 40%  |
| 200    | 0.01264 | 0.00376 | 1%          | 6%   | 7%   | 3%   | 24%  | 20%  | 39%  |
| 100    | 0.01302 | 0.00412 | 2%          | 5%   | 6%   | 2%   | 26%  | 23%  | 35%  |
| 20     | 0.01581 | 0.00991 | 4%          | 1%   | 0%   | 0%   | 43%  | 47%  | 5%   |
| 19     | 0.01598 | 0.01034 | 4%          | 1%   | 0%   | 0%   | 43%  | 49%  | 2%   |
| 18     | 0.01613 | 0.01072 | 5%          | 0%   | 0%   | 0%   | 44%  | 51%  | 1%   |
| 17     | 0.01620 | 0.01090 | 5%          | 0%   | 0%   | 0%   | 44%  | 51%  | 0%   |
| 16     | 0.01623 | 0.01098 | 5%          | 0%   | 0%   | 0%   | 44%  | 52%  | 0%   |
| 15     | 0.01625 | 0.01103 | 5%          | 0%   | 0%   | 0%   | 43%  | 52%  | 0%   |
| 14     | 0.01627 | 0.01109 | 5%          | 0%   | 0%   | 0%   | 43%  | 52%  | 0%   |
| 13     | 0.01629 | 0.01117 | 5%          | 0%   | 0%   | 0%   | 42%  | 52%  | 0%   |
| 12     | 0.01632 | 0.01127 | 5%          | 0%   | 0%   | 0%   | 42%  | 53%  | 0%   |
| 11     | 0.01635 | 0.01140 | 6%          | 0%   | 0%   | 0%   | 41%  | 53%  | 0%   |
| 10     | 0.01639 | 0.01156 | 6%          | 0%   | 0%   | 0%   | 40%  | 54%  | 0%   |
| 9      | 0.01644 | 0.01177 | 6%          | 0%   | 0%   | 0%   | 39%  | 54%  | 0%   |
| 8      | 0.01650 | 0.01207 | 6%          | 0%   | 0%   | 0%   | 38%  | 55%  | 0%   |
| 7      | 0.01657 | 0.01249 | 7%          | 0%   | 0%   | 0%   | 37%  | 56%  | 0%   |
| 6      | 0.01668 | 0.01311 | 8%          | 0%   | 0%   | 0%   | 35%  | 58%  | 0%   |
| 5      | 0.01682 | 0.01407 | 8%          | 0%   | 0%   | 0%   | 32%  | 60%  | 0%   |
| 4      | 0.01703 | 0.01570 | 10%         | 0%   | 0%   | 0%   | 28%  | 63%  | 0%   |
| 3      | 0.01739 | 0.01873 | 12%         | 0%   | 0%   | 0%   | 21%  | 67%  | 0%   |
| 2      | 0.01811 | 0.02548 | 16%         | 0%   | 0%   | 0%   | 6%   | 77%  | 0%   |
| 1      | 0.01844 | 0.02876 | 18%         | 0%   | 0%   | 0%   | 0%   | 82%  | 0%   |
| 0.75   | 0.01844 | 0.02878 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| 0.5    | 0.01844 | 0.02885 | 16%         | 0%   | 0%   | 0%   | 0%   | 84%  | 0%   |
| 0.25   | 0.01845 | 0.02922 | 14%         | 0%   | 0%   | 0%   | 0%   | 86%  | 0%   |
| 0      | 0.01848 | 0.03394 | 0%          | 0%   | 0%   | 0%   | 0%   | 100% | 0%   |

By contrast to the MV results shown in Table 2, in Table 4 all seven (rather than merely four) of the assets are found in portfolios on the efficient frontier. Arguably many investors might well consider the portfolios obtained by means of the tracking

error (TE) to be more realistic than the portfolios obtained by simply using the MV optimisation technique since more assets were found to be included in the optimal TE portfolios than in the optimal MV portfolios.

It is worth mentioning that in our data set although the components of our selected benchmark, the JSE Overall Index, are found in the opportunity set, the exact weightings of these components have changed over the 5-year time horizon. Consequently the model is unable to reconstruct a portfolio which is identical to the benchmark. Hence a portfolio with zero tracking error could not be formed. Nevertheless the weighting of each component in the optimal portfolio, with the minimum tracking error, is a compromise of the weightings of each component of the JSE Overall Index (which have changed over the 5-year time horizon). In theory the efficient frontier will touch the return axis.

In Figure 6 below we show the graphical consequences of being able to reconstruct the benchmark from the opportunity set of assets. We do this by including the benchmark (here the JSE Overall Index) in the opportunity set of assets.

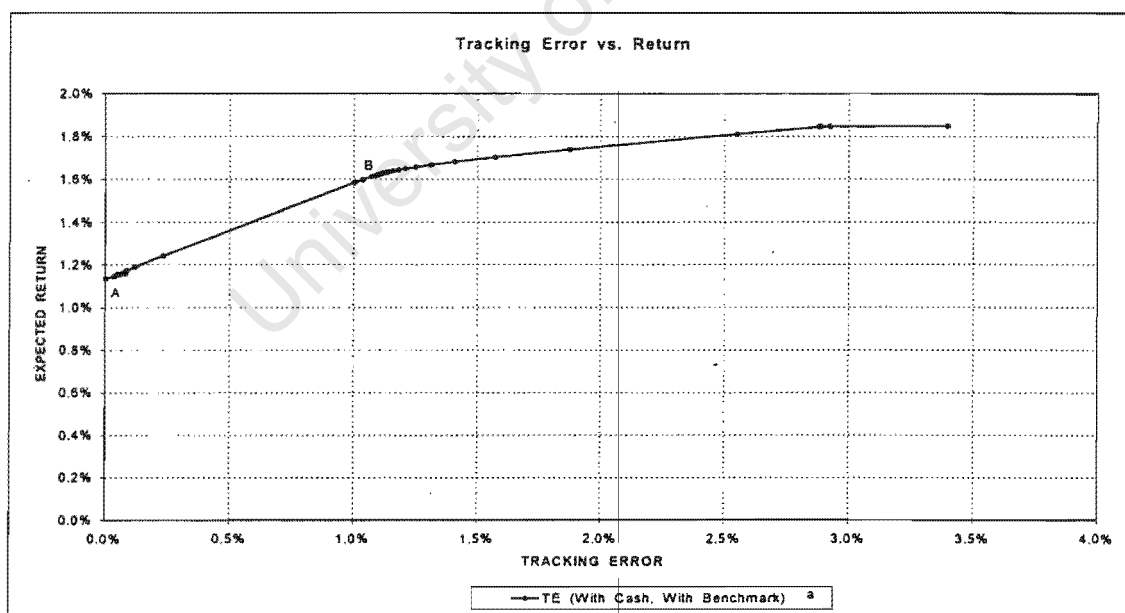


**Figure 6 Tracking Error Efficient Frontier of the opportunity set that includes cash and the benchmark. The JSE data spans the period 1992-1996.**

<sup>a</sup> The expression "With Cash, With Benchmark" in parenthesis above indicates that the benchmark itself and the risk-free asset i.e. cash are included in the opportunity set of assets.

The line segment  $AB^{15}$  consists of all portfolios that contain some weight of the benchmark. At point A we have 100% of the benchmark in the efficient portfolio whereas at point B the benchmark is excluded from the efficient portfolio. Hence in the mean-TE space the inclusion of the benchmark has a similar effect to including the risk-free asset in the MV efficient frontier (which results in the CML). The difference is that we can borrow or lend (sell short) the risk-free asset, but we cannot sell the benchmark short here

Clearly the choice of the benchmark will affect the resulting portfolios on the efficient frontier. It is thus imperative that much thought goes into the selection of the benchmark. A rationale for the inclusion of a benchmark in an objective function is that benchmarks are often imposed as a yardstick for performance comparison. In this empirical demonstration the JSE Overall Index was selected as the benchmark.

#### **4.1 The inclusion of cash in the opportunity set**

We included cash in the opportunity set of assets and found that the risk-free asset has not impacted on the efficient frontier. This is to some extent likely to be purely a consequence of the omission of the risk-free asset in our choice of benchmark (i.e. the JSE Overall Index). Note that although the risk-free rate has zero variance, in mean-tracking error space it will not be "tracking error free". In fact the tracking error of the risk-free asset will be the same magnitude of the variance of the benchmark. Thus we can no longer introduce the notion of a Capital Market Line.

#### **4.2 Insights for Practitioners**

- Each of the assets in the opportunity set considered was a member of at least one of the optimal tracking-error (TE) portfolios. This property did not hold when considering the optimal mean-variance portfolios. This testifies that the optimal portfolios achieved applying the Mean-Tracking-Error optimisation procedure are more diversified than the similar Mean-Variance optimal portfolios.

<sup>15</sup> In Appendix 1 we give a proof demonstrating that the segment AB is linear.

- The portfolio with the smallest tracking-error is the one whose composition is identical to the benchmark's composition.
- A portfolio comprising simply of cash has zero variance but certainly not zero tracking-error. This portfolio will certainly not be the portfolio with minimum tracking error, unless the benchmark itself is simply cash, which is very unlikely.

University of Cape Town

## 5. COMPARISON OF THE MV AND TE EFFICIENT FRONTIERS

The MV and TE efficient frontiers can be compared by plotting both frontiers in MV space, as shown in Figure 7.

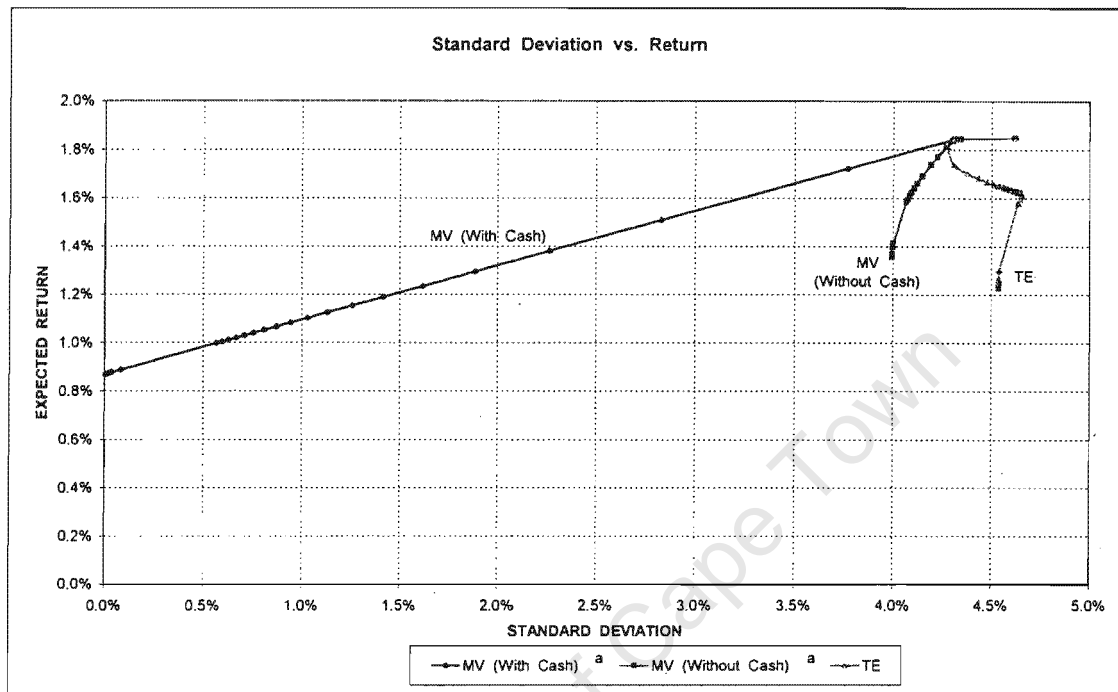


**Figure 7 MV and TE Frontiers in MV Space based on JSE data spanning the period 1992-1996.**

<sup>a</sup> The term in parenthesis "(With/Without Cash)" indicates that the risk-free asset has been included/excluded from the opportunity set.

By construction the TE efficient portfolios will be dominated by the MV efficient portfolios in MV space. Interestingly, the extent of domination rests upon how close the benchmark is to the MV efficient portfolio. In the above diagram this domination appears to be significant in MV space, indicating that TE efficient portfolios can certainly be MV inefficient (since their compositions vary substantially when compared to MV efficient portfolios). The TE frontier's composition certainly differs to the composition of the MV efficient frontier (with cash in the opportunity set) at lower levels of return since the TE efficient frontier never contains cash in its optimal portfolios whereas the MV optimal portfolios (with cash in their opportunity sets) certainly contain a substantial weighting in cash at these lower levels of return. The compositions of the TE and MV (whose opportunity set does not contain cash) efficient frontiers also differ significantly. At a return level of 1.4%, the minimum return for a MV efficient portfolio (to 1 decimal places), the MV optimal portfolio has

a weighting of 0% invested in the Mining Financial Index whereas the TE optimal portfolio at this return level has a weighting of 26% invested in the Mining Financial Index. Another significant difference in the investment weightings at this return level is that in the MV efficient portfolio (whose opportunity set does not contain cash) 60% is invested in the Industrial Index whereas in the TE efficient portfolio only 35% is invested in the Industrial Index.

Similarly the MV and TE efficient frontiers can be compared by plotting both frontiers in TE space, as done in Figure 8.

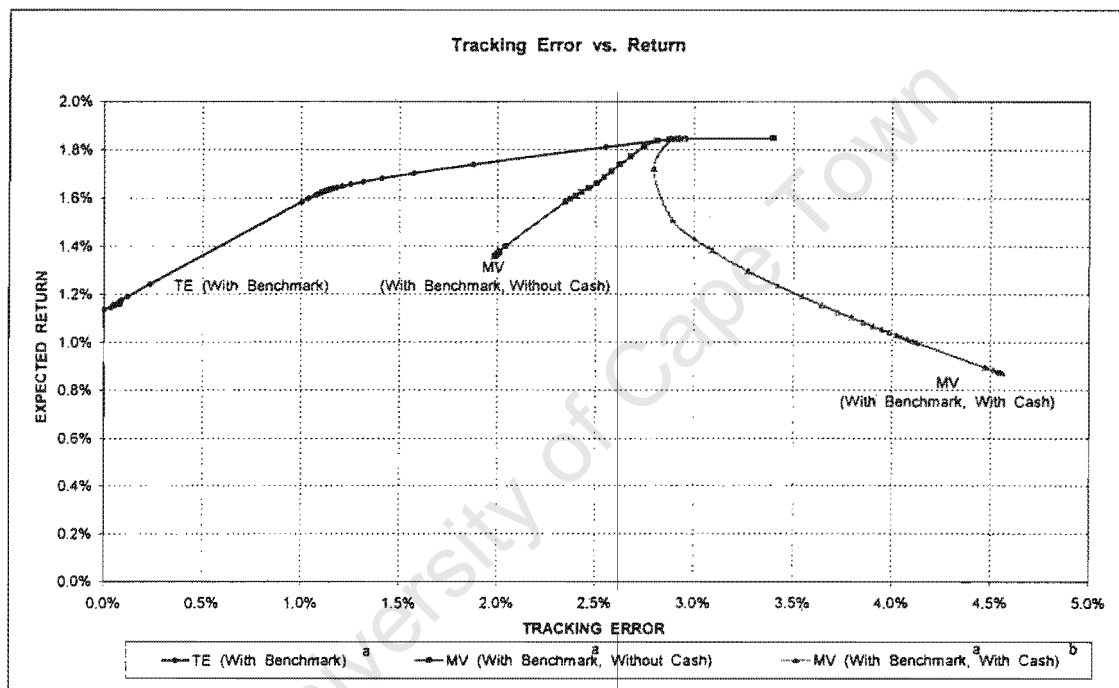


**Figure 8 MV and TE Frontiers in Mean-TE Space based on the JSE data spanning the period 1992-1996**

<sup>a</sup> The term in parenthesis "(with Benchmark)" indicates that the benchmark has been included in the opportunity set.

<sup>b</sup> The term in parenthesis "(with Cash)" indicates that Cash has been included in the opportunity set.

In the mean-tracking error space (by construction) the TE efficient frontier obviously dominates the MV frontier, as shown in Figure 8. An interesting feature of Figure 8 is the shape of the MV efficient frontier, which includes cash (i.e. the Capital Market Line (CML)). The CML "bends back" in TE space. The portion that bends back consists of the MV efficient portfolios containing substantial weighting in cash. This has two important implications for fund managers:

- (i) The backward sloping portion of the MV efficient frontier in TE space corresponds to the linear portion of the MV efficient frontier in MV space. This backward sloping portion of the MV efficient frontier bends back more and more as the component in *cash increases* - until finally the most *inefficient* portfolio in TE space is the one with 100% invested in cash. Thus as the cash component of the portfolio increases so the portfolio becomes *less efficient* in TE space.
- (ii) If portfolios are selected from efficient frontiers generated using the mean-variance framework, and they include cash - but managers are monitored relative to a benchmark performance, then they will be selecting from the backward sloping portion of the MV frontier in mean-TE space. This is highly *inefficient* in mean-TE space.

We note however that the benchmark here (the JSE-OVERALL INDEX) consists of equities alone. These insights may differ slightly if the benchmark instead consisted of some aggregate of funds e.g. the MV efficient frontier might well not bend back as sharply in mean-TE space if a proportion of the benchmark is cash.

### 5.1 Insights for Practitioners

- Mean-variance optimal portfolios dominate the mean-tracking-error optimal portfolios in mean-variance space. This dominance is more significant at low levels of return. The reason is that at low levels of return and variance the optimal mean-variance portfolios comprise mostly of cash, whereas at the same levels of low return (and low levels of tracking-error) the optimal tracking-error portfolios will comprise of very little, if any, cash.
- Similarly the mean-tracking-error optimal portfolios dominate the mean-variance optimal portfolios in mean-tracking-error space. Again this dominance becomes more evident as the mean-variance optimal portfolios start to comprise of more and more cash.

## 6. MEAN-VARIANCE TRACKING ERROR (MVTE) MODEL

Portfolio managers and their clients have slightly opposing objectives. Portfolio managers are rewarded for their performance measured relative to a benchmark. This uncertainty can be measured by the tracking error (TE). However, the client's primary concern might well be with losing money and this uncertainty could well be measured by the variance of returns. Hence there is an attempt to include both measures of risk. In this section we consider a joint objective function which combines the MV and TE approaches. One motivation for considering this combination is that efficient TE portfolios will generally not be MV efficient. If the benchmark is MV efficient (unlikely) then efficient TE portfolios can be MV efficient. The less MV efficient the benchmark is however, the less MV efficient TE portfolios will be and consequently the Sharpe ratio of TE efficient portfolios will be relatively poorer.

In the tracking error method of Franks<sup>16</sup> (1992) one is substituting relative risk for absolute risk. Chow<sup>17</sup> (1992) argues that the substitution of risk relative to the benchmark (relative risk) for absolute risk (variance) is too extreme for many investors. Although relative risk is important, investors are still usually concerned with the prospects of losing money. This notion is not fully captured by the tracking error model. When a portfolio declines with the benchmark the tracking error measure will not necessarily reflect an increase in tracking error or increase in risk. The model that Chow (1995) proposes is effectively a compromise between the Markowitz Mean-Variance optimisation and Franks' Tracking Error optimisation. Additionally both the MV and TE models can be considered as special cases of Chow's model. Chow (1995) claims that investors seek portfolios with high returns, low standard deviation and low tracking error. If this is true then the objective function should measure all three of these portfolio characteristics. Chow (1995) claims that his proposed method generates portfolios which more closely resemble the portfolios investors actually choose than do objective functions that do not take the tracking error into account. Portfolio managers reject traditional MV optimal portfolios and

<sup>16</sup> Franks, ECF, "Targeting Excess-of-Benchmark Returns", *The Journal of Portfolio Management*, Summer 1992, pp. 6-12.

<sup>17</sup> Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995, pp. 54-60.

consciously select portfolios that perform closer to some benchmark. Hence their objectives surely must include a measure of this preference!

The objective function upon which Chow's Mean-Variance Tracking Error (MVTE) is based on maximising the expression:

$$\text{MAX Expected Return} - \frac{\text{Expected Variance}}{\text{Risk Tolerance}} - \frac{(\text{Expected Tracking Error})^2}{\text{Tracking Error Tolerance}}$$

Reformulating the objective function in matrix notation one maximises:

$$\text{MAX } X'E - \lambda X'\Sigma X - \psi X'\Sigma^* X$$

$\lambda$  sets the price of variance in terms of return and  $\psi$  sets the price of tracking error in terms of return.

As before this objective function is typically subjected to the following two constraints:

1.  $\sum X_i = 1$
2.  $X_i \geq 0$  for all  $i$

Using this optimisation function the portfolio optimisation process searches for a solution with a high expected return and low portfolio volatility on an absolute basis, as well as relative to the benchmark. Chow argues there are effectively three goals. Investors are able to specify the relative importance of these goals when they assign risk tolerance  $\left(\frac{1}{\lambda}\right)$  and tracking error tolerance  $\left(\frac{1}{\psi}\right)$ . Tolerances can be set so that the MVTE collapses to either the MV or TE models as described below.

The set of efficient portfolios includes the MV efficient set, the mean tracking error (MTE) efficient set and all convex combinations of these two sets. Whereas the Markowitz MV method produces an efficient frontier in the risk-return space and the tracking error (TE) method produces an efficient frontier in the tracking error-return

space, the MVTE instead produces an efficient surface in the three dimensional return, risk and tracking error space. Two of the boundaries of this efficient surface are reasonably intuitive and will be described below. This efficient surface was created from the same opportunity set i.e. the seven second tier indices of the Johannesburg Stock Exchange, that was used to construct the MV and TE efficient frontiers:

The first boundary of the efficient surface is the subset of objectives of the MVTE that ignore the expected tracking error - i.e. the usual MV solution. To create this subset the tracking error tolerance is set to infinity or effectively  $\psi$  is set equal to zero. The tracking error term in the MVTE objective function disappears and the MVTE objective function reduces to the Markowitz MV objective function. Effectively this boundary of the MVTE objective function would be the efficient frontier of the MV model but plotted in three-dimensional space rather than in just two-dimensional space. The three dimensions are return, (absolute) risk and tracking error. The return and (absolute) risk (standard deviation) co-ordinates of this boundary would correspond to those obtained by simply using the MV objective function. However, the tracking error of each optimal solution needs to be calculated and this value is then used as the third co-ordinate in the three dimensional space of return, (absolute) risk and tracking error.

The second boundary of the efficient surface is the subset of objectives of the MVTE that ignore expected (absolute) risk. To create this subset the (absolute) risk tolerance is set to infinity or effectively  $\lambda$  is set equal to zero. The variance (absolute risk squared) term in the MVTE objective function disappears and the MVTE objective function reduces to the minimum tracking error (TE) objective function. The boundary of the MVTE objective being considered where  $\lambda = 0$  would be the efficient frontier of the TE model but plotted in the three-dimensional space of return, (absolute) risk and tracking error rather than in the two dimensional space of only return and tracking error. In the three dimensional space the co-ordinates of expected return and expected tracking error correspond to the co-ordinates of expected return and expected tracking error in two dimensional space. The third co-ordinate is (absolute) risk (standard deviation). The third co-ordinate is calculated by taking the

optimal portfolios obtained by means of the TE method and then calculating the standard deviation of each of these optimal portfolios.

Figure 9 is the efficient surface determined by implementing the MVTE model using our data of indices plotted in (absolute) risk, tracking error and return space.

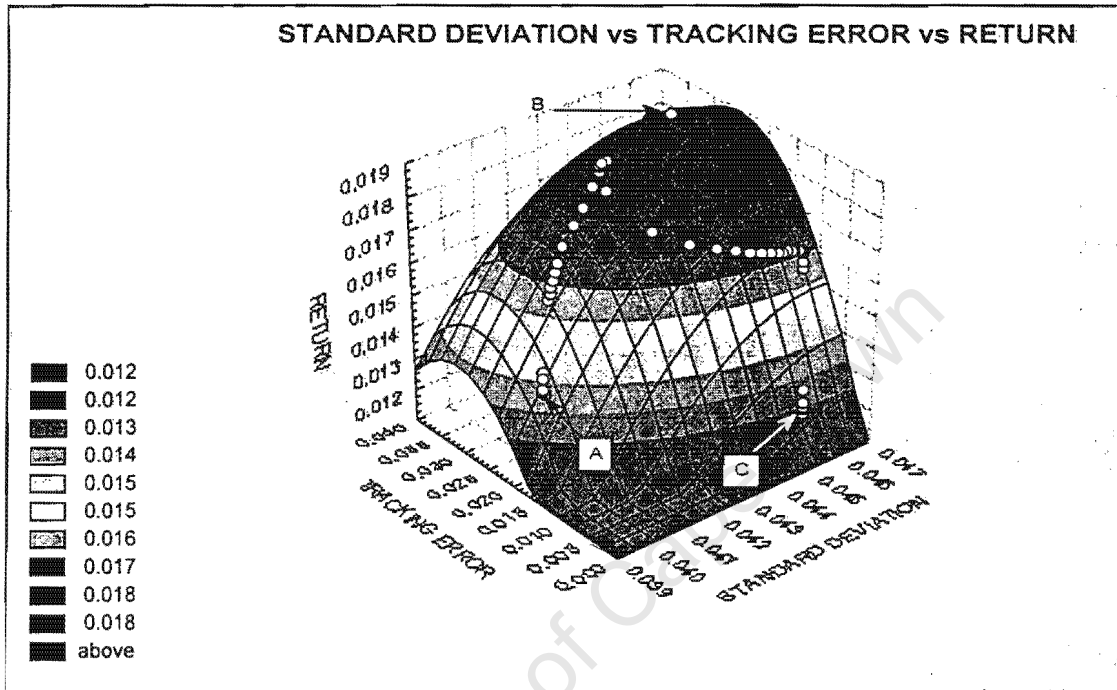


**Figure 9** Efficient surface of the MVTE model based on the JSE data spanning the period 1992-1996

The MV and MTE efficient frontiers delineate two of the boundaries of this surface. The MV efficient frontier originates with the risk-return co-ordinate of the lowest (absolute) risk portfolio (point A). From this point the expected (absolute) risk and return increase. However, depending on the choice of the benchmark the tracking error may be either increasing or decreasing along the MV frontier, which is evident in our empirical example. In Figure 9 the end-point, point B, is the co-ordinates of the highest expected return portfolio where both efficient frontiers join. The MTE efficient frontier starts with the tracking error co-ordinates of the lowest tracking error portfolio, point C. From this point the expected tracking error and expected return increase. It joins the MV frontier always at the highest expected return portfolio which is point B. A line that connects the co-ordinates of the portfolio with the lowest absolute risk (standard deviation) and lowest relative risk (tracking error deviance) completes the boundary for the MVTE efficient surface. The return, risk and tracking

error of any optimal portfolio generated by the MVTE optimiser lies on the convex adjoining the above two boundaries. An additional point worth noting is the degree to which the boundaries separate will be a function to which the benchmark departs from the MV efficient frontier!

In Figure 10 the MVTE optimal surface, as depicted in Figure 9, has been projected onto the two dimensional space of absolute risk and return. The MV and TE frontiers mark the boundaries of the MVTE surface.

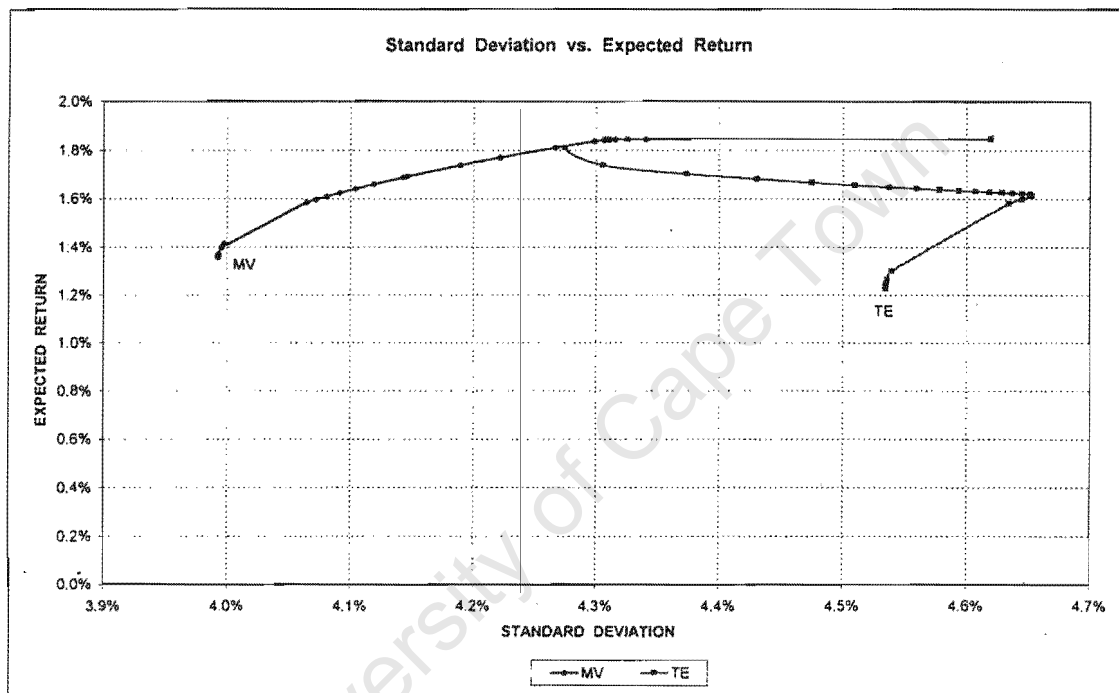


Figure 10 Boundaries of the MVTE surface plotted in (Absolute) Risk - Return space

Similarly, in Figure 11 the MVTE optimal surface has been projected onto the two-dimensional space of return and tracking error. Again the MV and MTE frontiers mark the boundaries of the MVTE surface. In this empirical example the tracking errors of the two optimisation methods do differ significantly at lower tracking error levels. As noted earlier these two frontiers become closer as the benchmark becomes more MV efficient!

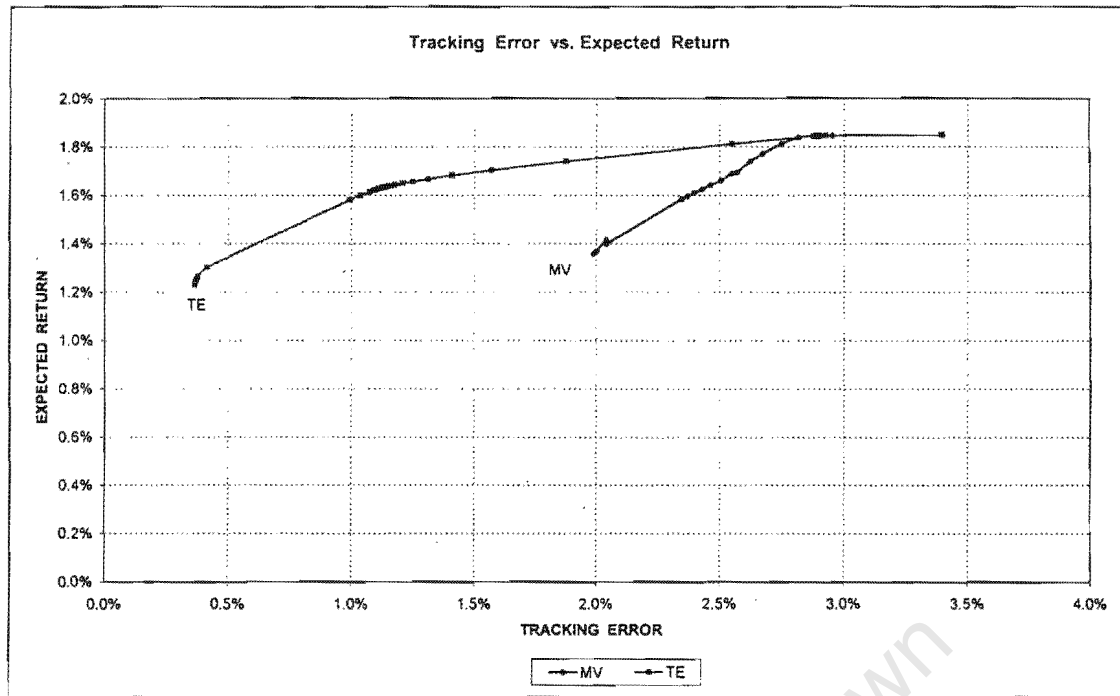


**Figure 11** Boundaries of the MVTE space plotted in tracking-error - return space based on the JSE data which spans the period 1992-1996

The optimal portfolios that lie within MV and MTE boundaries are obtained by letting  $\lambda$  and  $\psi$  run through all their values between 0 and  $\infty$  and at each step as  $\lambda$  and  $\psi$  change the MVTE objective function is optimised. At each step the optimal proportions to be invested in each one of the seven indices are determined. It is worth noting that the MV efficient frontier plotted in Expected Return and Expected Tracking-Error (not Expected (absolute) Risk) space can have a convex portion<sup>18</sup>, and vice versa.

<sup>18</sup> Depending on the choice of the benchmark, the expected tracking error may be increasing or decreasing along the MV frontier when plotted in expected tracking-error and expected return space - Chow (1995).

6.1 The inclusion of cash in the opportunity set

Figure 12 displays the effect of the inclusion of cash in the opportunity set on the efficient surface.

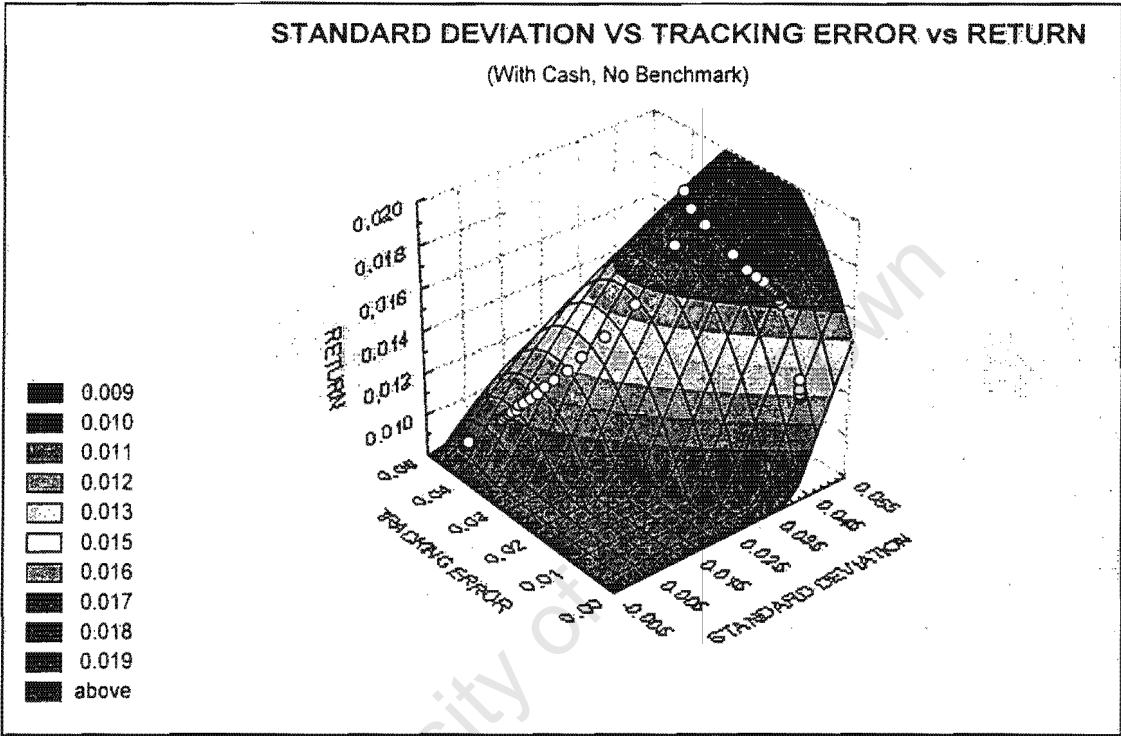


Figure 12 Efficient surface of the MVTE based on the JSE data which spans the period 1992-1996

From Figure 12 we see that the inclusion of the risk-free asset has caused the MV boundary to "bend back" on the TE axis. This insight was highlighted in the previous section. This "bending back" portion of the MV efficient frontier is the CML segment (which contains cash). This clearly suggests that including cash is highly inefficient in mean-TE space (for our choice of the benchmark, the JSE Overall Index).

## 6.2 *Insights for Practitioners*

- All the optimal portfolios together form an efficient *surface* in three dimensional mean-variance-tracking-error space. The mean-variance and mean-tracking-error efficient frontiers form two of the boundaries of this efficient surface.
- Just as cash affects the shape of the mean-variance efficient frontier in mean-tracking-error space so similarly the inclusion of cash in the opportunity set affects the shape of the efficient surface.
- If cash is included in the opportunity set then the efficient surface will span a much broader area.

University of Cape Town

## 7. SUMMARY COMPARISON BETWEEN MV, TE AND MVTE PORTFOLIOS

We have selected two levels of expected returns and have summarised, in Table 5, the attributes of the optimal portfolios obtained using each of the three asset allocation techniques. When the expected return is set to a specific value there are an infinite number of optimal MVTE portfolios since  $\lambda$  and  $\psi$  could assume an infinite number of values. In Table 5 we have considered the MVTE optimal portfolio that has an equal tolerance for the variance and tracking error and which simultaneously achieves the predefined expected return. In the case where the expected return is defined as 1.4% the value for  $\lambda$  and for  $\psi$  which achieves equal variance and tracking error tolerance is 21. Similarly in the situation where the expected return is defined as 1.7% the value for  $\lambda$  and for  $\psi$  which ensures equal variance and tracking error tolerance is 7.

**Table 5** Attributes of portfolios at selected Expected Return levels based on the JSE data spanning 1992-1996

| CHARACTERISTICS          | TECHNIQUE |      |      |      |      |      |
|--------------------------|-----------|------|------|------|------|------|
|                          | MV        | TE   | MVTE | MV   | TE   | MVTE |
| <i>Expected Return</i>   | 1.4%      | 1.4% | 1.4% | 1.7% | 1.7% | 1.7% |
| Standard Deviation       | 4%        | 4.6% | 4.1% | 4.2% | 4.4% | 4.3% |
| Tracking Error Deviation | 2%        | 0.6% | 1.3% | 2.6% | 1.6% | 1.7% |
| <b>Allocation</b>        |           |      |      |      |      |      |
| Coal Index               | 10%       | 3%   | 7%   | 15%  | 10%  | 10%  |
| Diamond Index            | 0%        | 4%   | 1%   | 0%   | 0%   | 0%   |
| All Gold Index           | 3%        | 4%   | 6%   | 0%   | 0%   | 0%   |
| Metals & Minerals Index  | 0%        | 1%   | 1%   | 0%   | 0%   | 0%   |
| Mining Financial Index   | 0%        | 33%  | 10%  | 1%   | 28%  | 22%  |
| Financial Index          | 32%       | 30%  | 35%  | 67%  | 62%  | 62%  |
| Industrial Index         | 55%       | 25%  | 40%  | 17%  | 0%   | 6%   |

As mentioned previously, it is evident that the MV technique allocates funds to very few of the assets being considered whereas by introducing the tracking error measurement the TE and MVTE models certainly do allocate funds to more of the assets. Intuitively if the opportunity set spans the benchmark a portfolio with a tracking error of zero could theoretically be constructed. It would of course be the benchmark itself! Thus if the benchmark is diversified some of the efficient portfolios selected by this model would also be diversified.

Now considering an expected return of 1.4% the MV technique obviously generates the portfolio with the minimum possible standard deviation (risk), viz. 4%. The TE method has a larger standard deviation of 4.6%. The MVTE model is basically a compromise between the MV and TE models. Consequently its standard deviation of 4.1% lies between the standard deviations of the MV and TE models. Similarly the tracking error of the MVTE, which is 1.3%, is a compromise between the minimum possible tracking error (obtained by means of the TE method) which is 0.6% and the largest tracking error of 2% which is obtained by the application of the MV process. These relationships are consistent for all feasible expected returns.

Effectively the MVTE method is a compromise of the MV method and the MTE method. The weight that is placed on each of these two methods is a function of the risk tolerance  $1/\lambda$  and tracking error tolerance  $1/\psi$  respectively. The optimal solution obtained by the application of the MVTE technique is always comprised of assets, which are either obtained by the implementation of the MV or the TE optimisation techniques.

## 8. SUMMARY OF THE IMPLEMENTATION OF THE MVTE MODEL

Chow (1995) claims that the MVTE method will be more popular with practitioners since by its implementation portfolio managers can find a practically sensible solution to their asset allocation problem. The process to be followed in order to implement the MVTE method is:

1. estimate the return of each index
2. estimate the variance-covariance matrix of absolute returns and the variance-covariance matrix of relative returns of the opportunity set of assets
3. define the benchmark
4. adjust the risk tolerance and tracking error tolerance in order to find an acceptable portfolio

The adjustments made to the risk tolerance and tracking error tolerance result in portfolio changes that make intuitive sense. A decrease in tracking error tolerance moves the optimal portfolio closer to the benchmark, and a decrease in risk tolerance shifts the portfolio away from the risky indices. The optimal portfolio involves adjustments to parameters that are investor specific. Each investor may have a different level of risk tolerance and tracking error tolerance and consequently the optimal solution found by the application of the MVTE is **INVESTOR SPECIFIC**. The MVTE optimisation method does not manipulate the forecasts of asset returns and covariances in order to find an acceptable portfolio.

In the empirical research we examined seven assets. For an  $n$ -asset problem one needs  $n$ -expected returns and  $\sum_{i=1}^n i$  expected covariances. Consequently, in our example the

total number of estimates needed were the 7 estimates of expected returns and  $\sum_{i=1}^7 i = (1+2+3+4+5+6+7) = 28$  estimates of the covariances. In total we need  $28 + 7 = 35$  estimates. The portfolio manager must estimate the true expected return and the covariance set. Errors in the estimates lead the optimisation algorithm to select a solution that appears to be optimal but is actually suboptimal. An area that needs to be examined closely and which is closely linked to the optimisation models is the

procedure of forecasting. This is an issue of concern because investors find that small changes in inputs translate into large changes in the resulting optimal portfolios.

University of Cape Town

## 9. DOWNSIDE MEASURE OF THE TRACKING ERROR (DTE)

In Section 4 where the objective was to minimise tracking-error volatility one is essentially searching for portfolios which jointly minimise returns below and *above* some benchmark. Our objective here is to beat the benchmark and to consider the minimum likelihood of underperforming the benchmark! There is certainly a strong case for arguing that a manager's objective should be focussed on minimising *under performance* of the benchmark! Why penalise funds which outperform the benchmark? Here we define risk in a manner that minimises the chance of not beating the benchmark.

The TE optimisation technique effectively penalises a portfolio that outperforms the benchmark. We consequently propose an adjustment to the measure of the tracking error to take account of this concern.

If relative returns were symmetrical (about the benchmark) this should result in the same asset allocation between TE and DTE techniques. The extent to which these differ is a function of the skewness of returns e.g. if an asset's return's were skewed to the positive side they would be ignored by the TE (or MV) because the average distance would still be large, however DTE would favour them.

In Section 4 the tracking error was defined as the variance of the *relative return*. Denoting the return of asset,  $i$ , at time  $t$  as  $R_{it}$  and the return of the benchmark as  $R_{bt}$  then the *relative return* at time  $t$  for security  $i$  ( $RR_{it}$ ) is written as:

$$RR_{it} = R_{it} - R_{bt}$$

We propose an adjustment to this measure:

$$\begin{aligned} \text{Downside } RR_{it} &= 0 && \text{for } R_{it} > R_{bt} \\ &= R_{it} - R_{bt} && \text{otherwise} \end{aligned}$$

The downside tracking error will now be the variance of the *downside* relative return.

Let  $\Sigma''$  represent the variance-covariance matrix of the *downside relative returns* ( $RR_i$ 's) of the assets in the opportunity set. Effectively the objective of the Downside Tracking Error (DTE) can be formulated as:

$$\text{MAX } X'E - \psi X'\Sigma''X$$

As before this objective function is typically subjected to the following two constraints:

1.  $\Sigma X_i = 1$
2.  $X_i \geq 0$  for all  $i$

Figure 13 below plots the DTE in mean-TE (not DTE) space for comparison purposes with the TE efficient frontier using the second tier indices of the Johannesburg Stock Exchange as the opportunity set. Hence by construction the DTE frontier cannot dominate the TE frontiers in mean-TE space.

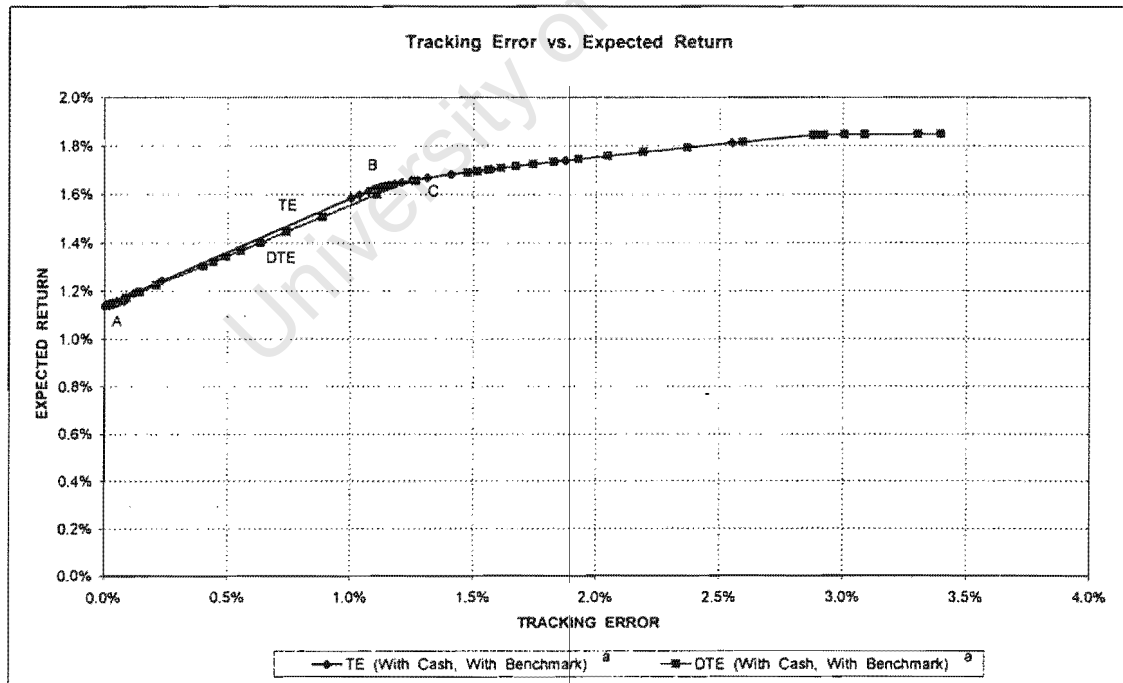


**Figure 13 Comparison between the TE and DTE in TE space based on the JSE data spanning the period 1992-1996**

<sup>a</sup> The expression "With Cash, With Benchmark" in parenthesis above indicates that the benchmark itself and the risk-free asset i.e. cash are included in the opportunity set of assets.

Interestingly the linear portion of the DTE efficient frontier extends marginally further than that of the TE efficient frontier. This suggests the DTE efficient portfolios approximate the benchmark (in composition) more closely than TE efficient portfolios in tracking error levels from 1% to 1.25% as indicated in Figure 13. That is the linear segment AC is marginally longer than AB. By construction the DTE efficient portfolios will be dominated by the TE efficient portfolios in TE space. The extent of domination is marginal in mean-TE space.

The efficient frontiers need to be compared in Mean-DTE space.

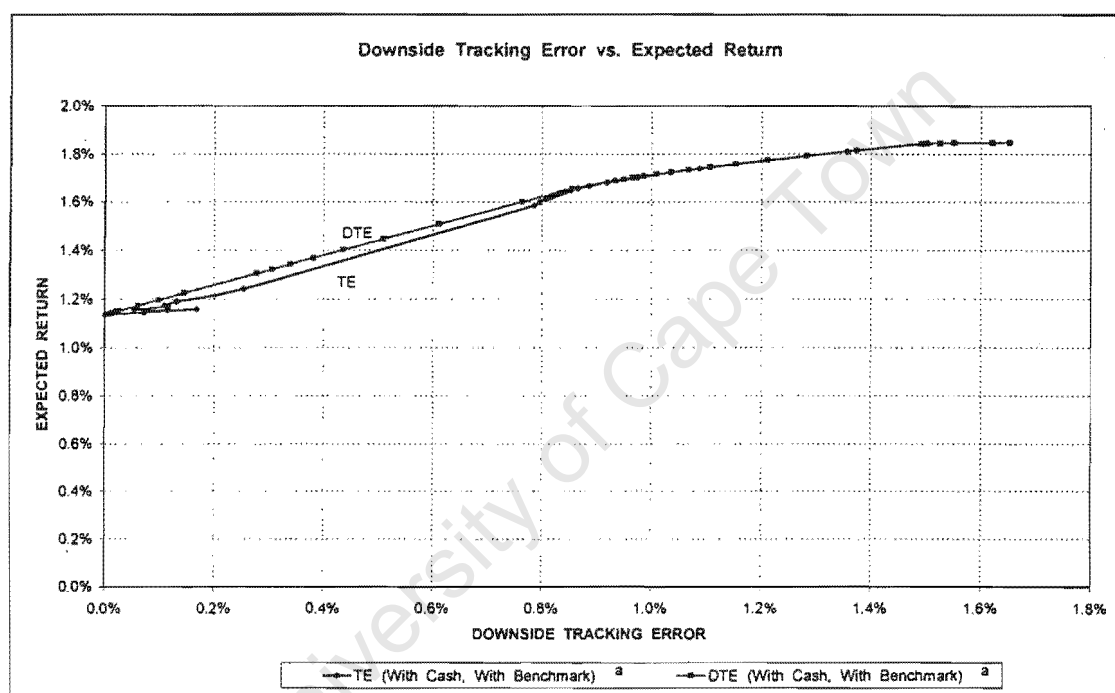


**Figure 14 Comparison between the TE and DTE in DTE space based on the data spanning the period 1992-1996. The compositions of these optimal DTE portfolios are listed in Table 6**

<sup>a</sup> The term in parenthesis "With Cash, With Benchmark" indicates that cash and the benchmark have been included in the opportunity set.

Interestingly the TE efficient frontier has a "kink" at a low level of DTE tolerance. The TE and DTE efficient frontiers are certainly "close" together in both mean-tracking error space as well as in mean-downside tracking error space for higher levels of TE's and DTE's. Of interest is relative differences in the compositions of the portfolios of these two frontiers. Table 4 and Table 6 display the compositions of the two efficient frontiers at selected levels of return.

**Table 6** Composition of the DTE efficient frontier based on the JSE data spanning 1992-1996

| DTE EFFICIENT FRONTIER: With Cash, Without Benchmark |         |         |         | PROPORTIONS |      |      |      |      |      |      |      |
|------------------------------------------------------|---------|---------|---------|-------------|------|------|------|------|------|------|------|
| RETURN                                               | STD     | DTE     | TE      | coal        | diam | gold | metm | mfin | fin  | Indu | Cash |
| 0.01366                                              | 0.04695 | 0.00676 | 0.00771 | 6%          | 5%   | 0%   | 1%   | 42%  | 19%  | 28%  | 0%   |
| 0.01369                                              | 0.04695 | 0.00676 | 0.00773 | 6%          | 5%   | 0%   | 1%   | 42%  | 19%  | 28%  | 0%   |
| 0.01371                                              | 0.04695 | 0.00676 | 0.00775 | 6%          | 5%   | 0%   | 0%   | 42%  | 19%  | 28%  | 0%   |
| 0.01375                                              | 0.04695 | 0.00676 | 0.00777 | 6%          | 5%   | 0%   | 0%   | 42%  | 19%  | 28%  | 0%   |
| 0.01387                                              | 0.04695 | 0.00676 | 0.00785 | 6%          | 5%   | 0%   | 0%   | 42%  | 21%  | 26%  | 0%   |
| 0.01403                                              | 0.04691 | 0.00678 | 0.00794 | 6%          | 4%   | 0%   | 0%   | 43%  | 22%  | 25%  | 0%   |
| 0.01412                                              | 0.04689 | 0.00680 | 0.00799 | 6%          | 4%   | 0%   | 0%   | 43%  | 23%  | 24%  | 0%   |
| 0.01426                                              | 0.04686 | 0.00683 | 0.00809 | 6%          | 4%   | 0%   | 0%   | 43%  | 25%  | 22%  | 0%   |
| 0.01454                                              | 0.04681 | 0.00691 | 0.00836 | 6%          | 3%   | 0%   | 0%   | 43%  | 28%  | 19%  | 0%   |
| 0.01540                                              | 0.04679 | 0.00736 | 0.00968 | 7%          | 2%   | 0%   | 0%   | 44%  | 39%  | 9%   | 0%   |
| 0.01559                                              | 0.04681 | 0.00750 | 0.01005 | 8%          | 2%   | 0%   | 0%   | 44%  | 41%  | 6%   | 0%   |
| 0.01583                                              | 0.04684 | 0.00768 | 0.01055 | 8%          | 1%   | 0%   | 0%   | 44%  | 44%  | 4%   | 0%   |
| 0.01613                                              | 0.04690 | 0.00794 | 0.01123 | 8%          | 1%   | 0%   | 0%   | 44%  | 47%  | 0%   | 0%   |
| 0.01620                                              | 0.04668 | 0.00801 | 0.01137 | 8%          | 0%   | 0%   | 0%   | 44%  | 48%  | 0%   | 0%   |
| 0.01628                                              | 0.04638 | 0.00810 | 0.01156 | 8%          | 0%   | 0%   | 0%   | 42%  | 49%  | 0%   | 0%   |
| 0.01638                                              | 0.04595 | 0.00824 | 0.01188 | 9%          | 0%   | 0%   | 0%   | 40%  | 51%  | 0%   | 0%   |
| 0.01655                                              | 0.04528 | 0.00854 | 0.01263 | 10%         | 0%   | 0%   | 0%   | 37%  | 53%  | 0%   | 0%   |
| 0.01690                                              | 0.04416 | 0.00934 | 0.01474 | 11%         | 0%   | 0%   | 0%   | 30%  | 59%  | 0%   | 0%   |
| 0.01695                                              | 0.04401 | 0.00948 | 0.01514 | 11%         | 0%   | 0%   | 0%   | 29%  | 60%  | 0%   | 0%   |
| 0.01701                                              | 0.04385 | 0.00965 | 0.01559 | 11%         | 0%   | 0%   | 0%   | 28%  | 61%  | 0%   | 0%   |
| 0.01708                                              | 0.04368 | 0.00985 | 0.01611 | 12%         | 0%   | 0%   | 0%   | 27%  | 62%  | 0%   | 0%   |
| 0.01715                                              | 0.04351 | 0.01008 | 0.01671 | 12%         | 0%   | 0%   | 0%   | 25%  | 63%  | 0%   | 0%   |
| 0.01724                                              | 0.04334 | 0.01035 | 0.01742 | 12%         | 0%   | 0%   | 0%   | 24%  | 64%  | 0%   | 0%   |
| 0.01734                                              | 0.04316 | 0.01067 | 0.01826 | 13%         | 0%   | 0%   | 0%   | 22%  | 66%  | 0%   | 0%   |
| 0.01745                                              | 0.04299 | 0.01106 | 0.01925 | 13%         | 0%   | 0%   | 0%   | 19%  | 68%  | 0%   | 0%   |
| 0.01758                                              | 0.04284 | 0.01152 | 0.02044 | 14%         | 0%   | 0%   | 0%   | 17%  | 70%  | 0%   | 0%   |
| 0.01774                                              | 0.04273 | 0.01210 | 0.02189 | 14%         | 0%   | 0%   | 0%   | 14%  | 72%  | 0%   | 0%   |
| 0.01792                                              | 0.04268 | 0.01281 | 0.02368 | 15%         | 0%   | 0%   | 0%   | 10%  | 75%  | 0%   | 0%   |
| 0.01815                                              | 0.04276 | 0.01372 | 0.02591 | 16%         | 0%   | 0%   | 0%   | 6%   | 79%  | 0%   | 0%   |
| 0.01844                                              | 0.04307 | 0.01490 | 0.02877 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   | 0%   |
| 0.01844                                              | 0.04308 | 0.01492 | 0.02882 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   | 0%   |
| 0.01844                                              | 0.04309 | 0.01492 | 0.02884 | 17%         | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   | 0%   |
| 0.01844                                              | 0.04309 | 0.01492 | 0.02886 | 16%         | 0%   | 0%   | 0%   | 0%   | 84%  | 0%   | 0%   |
| 0.01844                                              | 0.04311 | 0.01493 | 0.02890 | 16%         | 0%   | 0%   | 0%   | 0%   | 84%  | 0%   | 0%   |
| 0.01844                                              | 0.04314 | 0.01495 | 0.02897 | 15%         | 0%   | 0%   | 0%   | 0%   | 85%  | 0%   | 0%   |
| 0.01845                                              | 0.04322 | 0.01499 | 0.02916 | 14%         | 0%   | 0%   | 0%   | 0%   | 86%  | 0%   | 0%   |
| 0.01846                                              | 0.04370 | 0.01524 | 0.03004 | 10%         | 0%   | 0%   | 0%   | 0%   | 90%  | 0%   | 0%   |
| 0.01846                                              | 0.04419 | 0.01549 | 0.03086 | 7%          | 0%   | 0%   | 0%   | 0%   | 93%  | 0%   | 0%   |
| 0.01848                                              | 0.04557 | 0.01620 | 0.03301 | 2%          | 0%   | 0%   | 0%   | 0%   | 98%  | 0%   | 0%   |
| 0.01848                                              | 0.04619 | 0.01651 | 0.03394 | 0%          | 0%   | 0%   | 0%   | 0%   | 100% | 0%   | 0%   |
| 0.01848                                              | 0.04619 | 0.01651 | 0.03394 | 0%          | 0%   | 0%   | 0%   | 0%   | 100% | 0%   | 0%   |

From these tables it is evident that only at lower relative risk levels the TE efficient portfolios tend to be slightly more diversified - for higher relative risk levels the compositions are similar.

### **9.1 *Insights for Practitioners***

- The tracking-error and downside tracking-error measures are very similar in nature when the securities' returns are distributed symmetrically. The consequent optimal portfolios obtained when using either one of these risk measures are very similar in composition.
- Downside tracking-error might well be a better measure of risk to use for securities such as derivatives whose return distributions are certainly not symmetrical.

University of Cape Town

## 10. THE INFLUENCE OF THE BENCHMARK ON THE TE AND DTE MODELS

✓ 12.1

The objective of the TE model is to minimise the tracking error squared (the variance of the returns relative to the benchmark) at each level of return. Denoting the return of asset,  $i$ , at time  $t$  as  $R_{it}$  and the return of the benchmark as  $R_{bt}$  then the *relative return* at time  $t$  for asset  $i$  ( $RR_{it}$ ) is:

$$RR_{it} = R_{it} - R_{bt}$$

The relative return,  $RR_{it}$ , comprises of two variables: the return of asset  $i$  and the benchmark's return. Effectively, both of these variables will influence the objective function. An estimate frequently used for asset  $i$ 's return is some historical mean. It is however the portfolio manager himself who determines what to employ as the <sup>(True)</sup> benchmark. Consequently, if two portfolio managers have the same expectations on the returns of the assets in the opportunity set but they define their benchmarks differently, then the compositions of their optimal portfolios may certainly differ. ✓

### 10.1 Benchmarks for TE Models

The portfolio manager may well decide upon treating the benchmark as a constant return (e.g. 1% per month). One must bear in mind that the tracking error squared is a measure of the variance of the asset's returns about the predefined benchmark. Consequently the tracking error about any constant value will be the same no matter the magnitude of the constant (see Appendix 2 for the proof). Figure 15 displays the efficient frontiers of the TE models whose benchmarks were taken to be different constants viz. 1%, 2% and 5%, using the second tier indices of the Johannesburg Stock Exchange as the opportunity set.

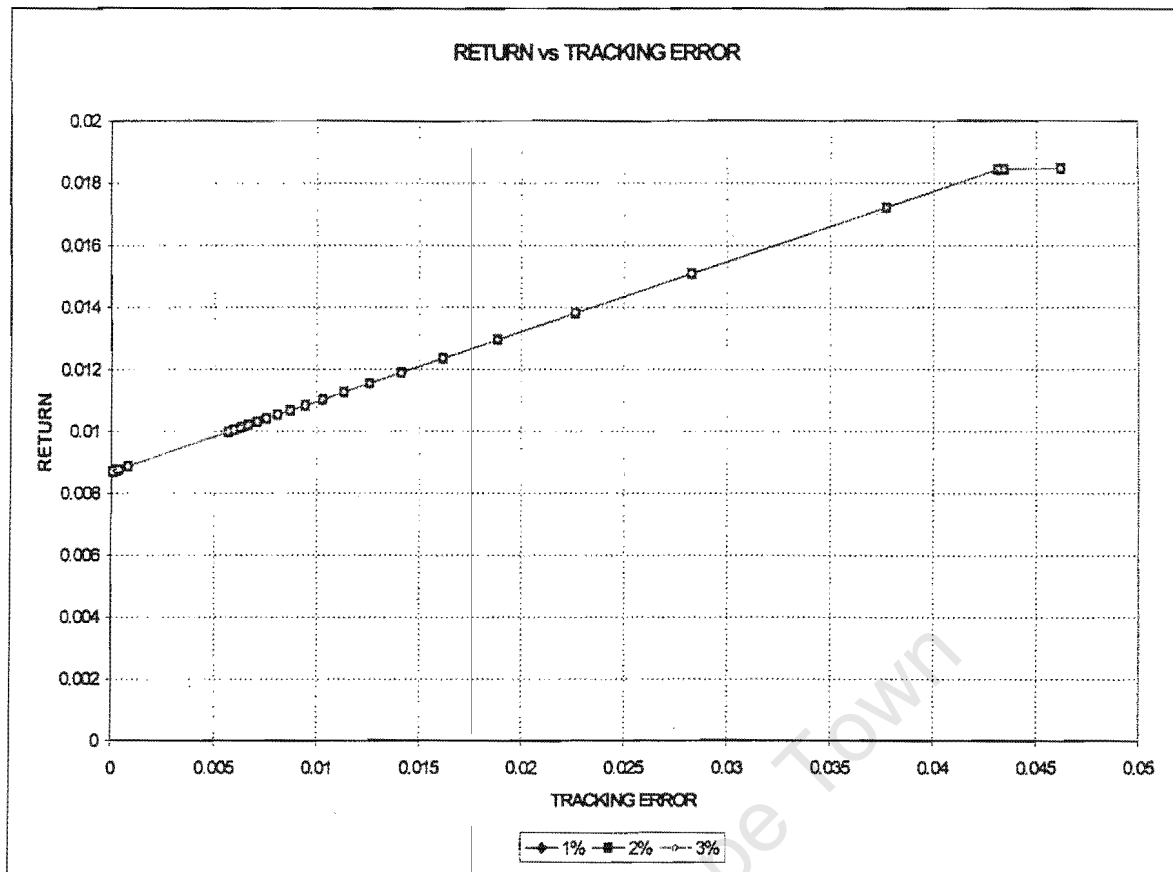


Figure 15 TE efficient frontiers based on the JSE data spanning 1992-1996

The respective TE efficient frontiers all coincide because the benchmarks taken for each of the TE models are constants (these constants do differ in magnitude). If one plotted the MV efficient frontier in return-tracking error space by treating the standard deviation co-ordinate as the tracking error co-ordinate, then the MV efficient frontier would also coincide with these TE efficient frontiers. This would be expected because effectively one is merely assigning a constant value of zero to the benchmark and consequently the variance and tracking error measures would be identical.

When the benchmark is truly a variable (not a constant) then the *locations* and *compositions* of TE efficient frontiers will certainly differ. Figure 16 displays the tracking error efficient frontier of a model whose benchmark was taken to be a constant, and the tracking error efficient frontier of the model whose benchmark was taken to be the JSE-Overall Index.

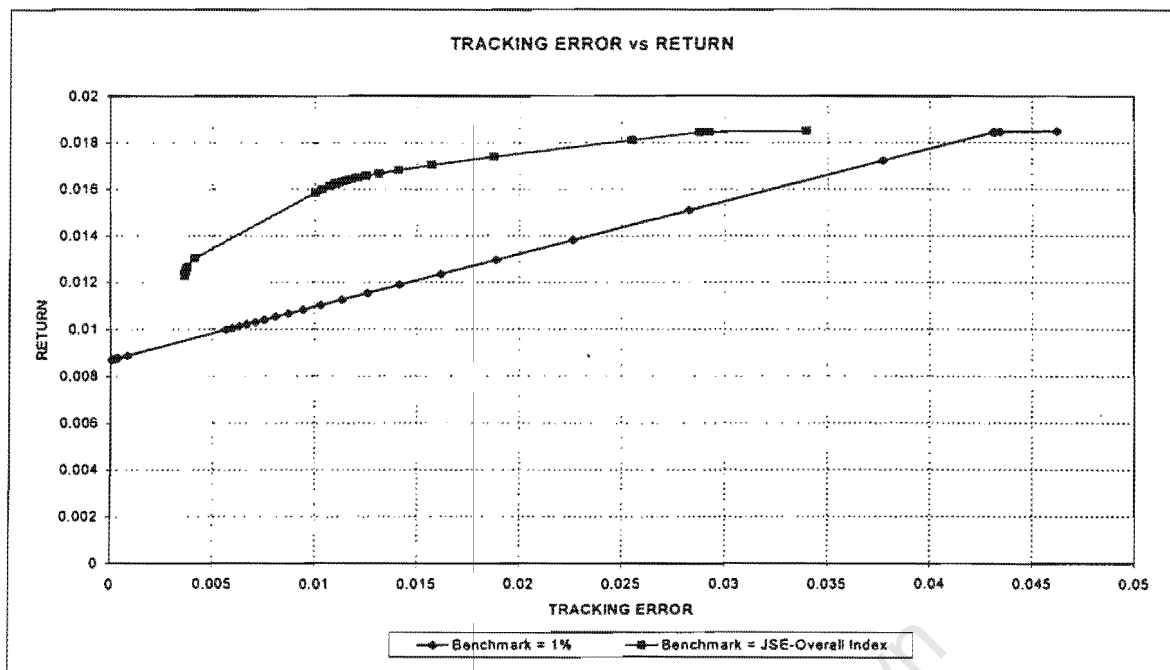


Figure 16 TE efficient frontiers based on the JSE data spanning the period 1992-1996

The opportunity set consisted of the 7 second-tier indices of the JSE together with a risk-free asset. From Figure 16 it is evident that the minimum TE for the TE model which used a constant return value as a benchmark, was zero. This was achieved by investing all the funds into the risk-free asset. The risk-free rate of return is a constant value and consequently it has zero variance relative to a constant benchmark hence its TE is zero.

Table 7 reinforces the notion that the *compositions* of the two TE efficient portfolios, which are plotted in Figure 16, differ.

Table 7 Comparative TE portfolio compositions at specified expected return levels based on the JSE data spanning the period 1992-1996

| CHARACTERISTICS           | BENCHMARK |               |      |               |
|---------------------------|-----------|---------------|------|---------------|
|                           | 1%        | OVERALL INDEX | 1%   | OVERALL INDEX |
| Expected Return           | 1.4%      | 1.4%          | 1.6% | 1.6%          |
| Standard Deviation        | 2.3%      | 4.6%          | 3.2% | 4.6%          |
| Tracking Error Deviation  | 2.3%      | 0.6%          | 3.2% | 1%            |
| Allocation                |           |               |      |               |
| Coal Index                | 9%        | 3%            | 13%  | 4%            |
| Diamond Index             | 0%        | 4%            | 0%   | 1%            |
| All Gold Index            | 0%        | 4%            | 0%   | 0%            |
| Metals & Minerals Index   | 0%        | 1%            | 0%   | 0%            |
| Mining Financial Index    | 0%        | 32%           | 0%   | 44%           |
| Financial Index           | 45%       | 31%           | 62%  | 49%           |
| Industrial Index          | 0%        | 25%           | 0%   | 2%            |
| Risk-free Asset ( $R_f$ ) | 46%       | 0%            | 25%  | 0%            |

## 10.2 Benchmarks for DTE Models

The benchmark is also a factor that partly determines the optimal DTE portfolio. The objective of the DTE model is to minimise the variance of the downside relative return (Downside  $RR_{it}$ ), which is defined as:

$$\begin{aligned} \text{Downside } RR_{it} &= 0 && \text{for } R_{it} > R_{bt} \\ &= R_{it} - R_{bt} && \text{otherwise} \end{aligned}$$

An estimate for the downside tracking error squared is:

$$\frac{\sum_{R_i < R_b}^n (R_i - R_b)^2}{n-1}$$

where  $n$  is the number of return observations.

Figure 17 displays the effects of different benchmarks on the DTE efficient frontier. The benchmarks considered were 4 different constant values and the JSE-Overall Index.

⇒ stochastic bm are better!

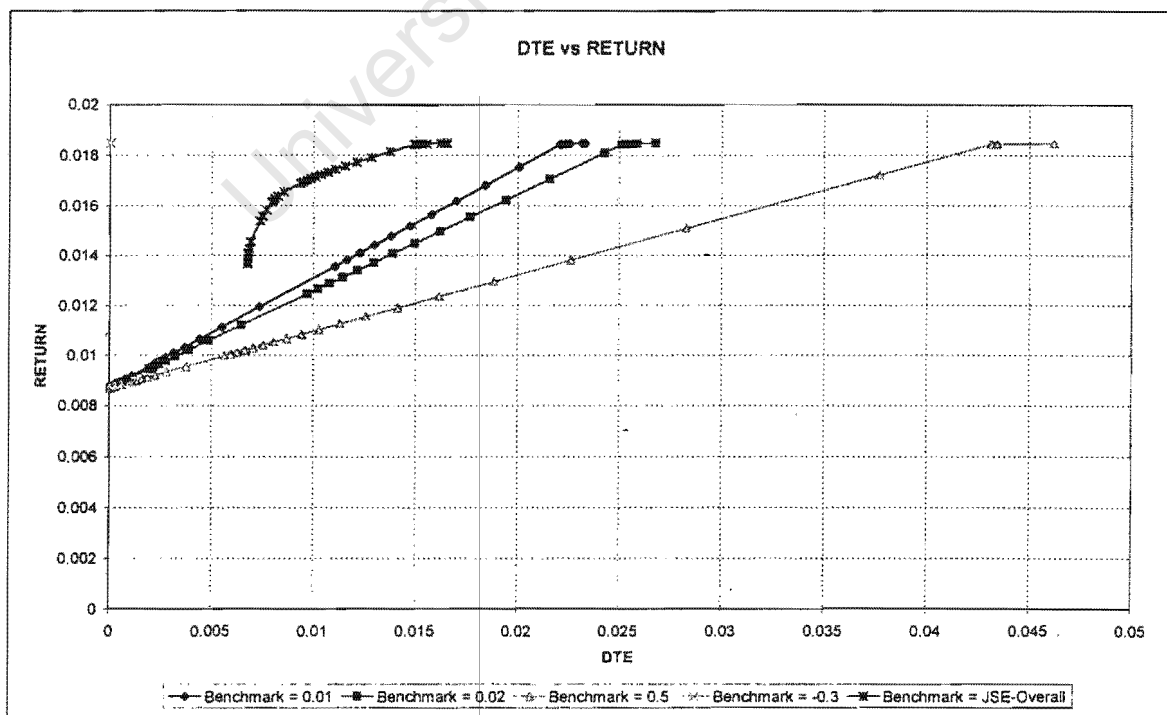


Figure 17 The effects of various benchmarks on the DTE efficient frontiers which are based on the JSE data spanning the period 1992-1996

It is noticeable that as the constant benchmarks increase in value so the DTE efficient frontiers shift to the right. This occurs because more of the distribution of returns falls below the specified constant benchmark as this constant value for the benchmark increases. For example, considering a constant benchmark of 0.01 an estimate of its

DTE would be  $\sum_{R_i < R_b} \frac{(R_i - 0.01)^2}{n-1}$  and the estimate for the DTE of a benchmark of 0.02 would be  $\sum_{R_i < R_b} \frac{(R_i - 0.02)^2}{n-1}$ . Now at any level of return  $R_{it} - R_{bt}$ .

$\sum_{R_i < R_b} \frac{(R_i - 0.01)^2}{n-1} \leq \sum_{R_i < R_b} \frac{(R_i - 0.02)^2}{n-1}$ . Effectively even benchmarks that differ only by

a constant value will produce efficient frontiers that differ in *location*, unlike in the TE models, where these TE efficient frontiers coincide. Table 8 displays the compositions of the optimal DTE portfolios at selected levels of return.

**Table 8 Comparative DTE portfolio compositions at specified expected return levels based on the JSE data spanning 1992-1996**

| CHARACTERISTICS           | BENCHMARK |      |      |               |      |      |      |               |
|---------------------------|-----------|------|------|---------------|------|------|------|---------------|
|                           | 1%        | 2%   | 50%  | Overall INDEX | 1%   | 2%   | 50%  | Overall INDEX |
| <i>Expected Return</i>    | 1.4%      | 1.4% | 1.4% | 1.4%          | 1.6% | 1.6% | 1.6% | 1.6%          |
| Downside Tracking Error   | 1.2%      | 1.4% | 2.3% | 0.7%          | 1.7% | 1.9% | 3.2% | 0.8%          |
| Tracking Error Deviation  | 2.3%      | 2.4% | 2.3% | 0.8%          | 3.2% | 3.2% | 3.2% | 1%            |
| Standard Deviation        | 2.3%      | 2.4% | 2.3% | 4.7%          | 3.2% | 3.2% | 3.2% | 4.7%          |
| <b>Allocation</b>         |           |      |      |               |      |      |      |               |
| Coal Index                | 11%       | 12%  | 9%   | 6%            | 15%  | 17%  | 13%  | 8%            |
| Diamond Index             | 0%        | 0%   | 0%   | 4%            | 0%   | 0%   | 0%   | 1%            |
| All Gold Index            | 0%        | 0%   | 0%   | 0%            | 0%   | 0%   | 0%   | 0%            |
| Metals & Minerals Index   | 0%        | 0%   | 0%   | 0%            | 0%   | 0%   | 0%   | 0%            |
| Mining Financial Index    | 0%        | 0%   | 0%   | 43%           | 0%   | 0%   | 0%   | 44%           |
| Financial Index           | 44%       | 42%  | 45%  | 22%           | 60%  | 58%  | 62%  | 46%           |
| Industrial Index          | 0%        | 0%   | 0%   | 25%           | 0%   | 0%   | 0%   | 2%            |
| Risk-free Asset ( $R_f$ ) | 46%       | 46%  | 46%  | 0%            | 25%  | 25%  | 25%  | 0%            |

because  $\sigma(R_{bt})$  is small!

The maximum monthly return for any of the indices measured over the period January 1992-December 1996 was 40%. Consequently if the benchmark is taken to be any constant value greater than 40% then the DTE measure would effectively be the same measure as the TE since each return would be less than the benchmark. It is evident that although the *locations* of the DTE efficient frontiers, whose benchmarks are constants, differ (as displayed in Figure 17), the *compositions* of these portfolios at selected levels of return differ only marginally. However, the *compositions* of DTE optimal portfolios do differ significantly if the benchmarks are not constant values.

This is evident when one compares the composition of the DTE optimal portfolio whose benchmark is a constant to the composition of the DTE optimal portfolio whose benchmark is the JSE-Overall Index. This again emphasises the importance and consequences in selecting benchmarks. The portfolio manager needs to determine → How? against which benchmark his portfolio's performance is being measured.

So why don't you hint of ways  
a portfolio manager should  
go about choosing a benchmark!

University of Cape Town

### 10.3 The Selection of the Benchmark

Portfolio managers are often judged by comparing their portfolios' performances relative to the performances of the portfolios of their opponents who are operating in the same industry. Consequently these portfolio managers wish to minimise their portfolios' risks relative to the average institutional portfolio, hence a benchmark often used in practice is the average institutional portfolio. Black and Litterman<sup>19</sup> — Why? (1991) proposed using the CAPM equilibrium as the benchmark. The CAPM states that the equilibrium expected excess return on a security is proportional to the covariance of the security's return with the return of the market portfolio. Mathematically the exact relationship is as follows:

$$E(R_i) - r_f = \beta(E(R_{mi}) - R_f)$$

Effectively by applying the CAPM equilibrium as a benchmark, one is minimising the risk of the portfolio relative to its expected CAPM equilibrium return. (Obviously)!

$$RR_{it} = R_{it} - R_{bt}$$

⇒ month on month new estimates of  $E(R_i)$  has to be made ⇒ Adjust the  $\beta$  continually.

Try this out!

<sup>19</sup> Black, F and Litterman, R, "Global Asset Allocation with Equities, Bonds and Currencies", *Fixed Income Research*, October 1991.

## 11. COMPARISON OF THE TE, DTE AND MV EFFICIENT FRONTIERS

The composition of a selection of optimal TE, DTE and MV portfolios are listed in Table 9 for set values of expected return.

**Table 9 Optimal TE, DTE and MV portfolios for a given return based on the JSE data spanning the period 1992-1996**

| CHARACTERISTICS                   | TECHNIQUE |       |      |      |      |      |
|-----------------------------------|-----------|-------|------|------|------|------|
|                                   | TE        | DTE   | MV   | TE   | DTE  | MV   |
| <i>Expected Return</i>            | 1.4%      | 1.4%  | 1.4% | 1.7% | 1.7% | 1.7% |
| Standard Deviation                | 4.6%      | 4.7%  | 2.4% | 4.4% | 4.4% | 3.7% |
| Tracking Error Deviation          | 0.6%      | 0.8%  | 3.1% | 1.6% | 1.6% | 2.8% |
| Downside Tracking Error Deviation | 0.6%      | 0.59% | 1.9% | 1%   | 0.9% | 1.6% |
| <b>Allocation</b>                 |           |       |      |      |      |      |
| Coal Index                        | 3%        | 6%    | 9%   | 10%  | 11%  | 15%  |
| Diamond Index                     | 4%        | 4%    | 0%   | 0%   | 0%   | 0%   |
| All Gold Index                    | 4%        | 0%    | 0%   | 0%   | 0%   | 0%   |
| Metals & Minerals Index           | 1%        | 0%    | 0%   | 0%   | 0%   | 0%   |
| Mining Financial Index            | 33%       | 43%   | 0%   | 28%  | 28%  | 0%   |
| Financial Index                   | 30%       | 22%   | 45%  | 62%  | 61%  | 71%  |
| Industrial Index                  | 25%       | 25%   | 0%   | 0%   | 0%   | 0%   |
| Risk-free Asset (Cash)            | 0%        | 0%    | 46%  | 0%   | 0%   | 14%  |

It is certainly noticeable that the compositions of the DTE optimal portfolios are effectively "adjustments" to the compositions of the TE optimal portfolios. The composition of a TE and a DTE optimal portfolio is certainly closer in resemblance than what the compositions of an MV optimal portfolio is when compared with either of the two. The similarity in the compositions of the TE and DTE optimal portfolios (and to a lesser extent the MV optimal portfolios) are even more evident as the TE tolerance and DTE tolerance (and MV tolerance) are relaxed. This procedure effectively places a greater "weight" on maximising the return relative to minimising the corresponding types of "risks" which are measured by the tracking error, downside tracking error and standard deviation respectively. This was certainly expected since the measures of the TE and the DTE are similar - but not identical.

We shall now compare the measures of TE and DTE directly.

The TE portfolio selection model's objective is to maximise return but simultaneously minimise the tracking error. If an asset offered an *abnormally* large return it would be

satisfying the objective of maximising return, but since it is *abnormal* it would simultaneously fall short of minimising the variance and consequently might well not be included in the optimal portfolio. The DTE objective function is slightly different. It only "penalises" an asset if it has a lower return than the benchmark. This penalty is accounted for in the DTE (which is an adjusted variance) of the asset. If the asset has a higher return than the benchmark then the tracking error is considered to be zero. Hence if an asset has an *abnormal* return it would satisfy the objective of maximising return and would not be penalised for its relatively larger tracking error. Consequently an optimal efficient downside tracking error portfolio might well differ slightly in composition to a traditional efficient tracking error portfolio.

Table 10 compares the rankings of the magnitudes of the assets' TE's and DTE's. The asset with the smallest TE receives a TE ranking of 1, similarly the asset with the smallest DTE receives a DTE ranking of 1.

**Table 10 Spearman's Rankings of the assets' TE's and DTE's obtained from the JSE data spanning 1992-1996**

| ASSET | TE    | TE RANK | DTE   | DTE RANK |
|-------|-------|---------|-------|----------|
| COAL  | .0888 | 6       | .0359 | 5        |
| DIAM  | .0782 | 5       | .0362 | 6        |
| GOLD  | .1107 | 7       | .0492 | 7        |
| METM  | .0778 | 4       | .0333 | 4        |
| MFIN  | .0669 | 3       | .0148 | 2        |
| FINI  | .0411 | 1       | .0147 | 1        |
| INDU  | .0416 | 2       | .0155 | 3        |

The correlations between these rankings can be computed by means of Spearman's Rank Correlation. This (non-parametric) correlation is 0.929 which is highly significant ( $p < 0.0025$ ). The TE and DTE are each a factor in the corresponding TE and DTE objective functions and due to a high correlation between them the TE and DTE objective functions will be similar and consequently the optimal portfolios will also be similar in composition.

## 12. MEAN-VARIANCE TRACKING ERROR DOWNSIDE TRACKING ERROR MODEL (MVTEDE)

Thus far we have suggested that portfolio managers are concerned about the expected return together with the absolute risk, tracking error and adjusted tracking error of their portfolios. Each portfolio manager might well have a different tolerance level for each of the measures of risk. Our final proposal is to combine each of these objectives into a single objective function. The overall objective is to:

$$\begin{aligned} \text{MAX Expected Return} - \frac{\text{Expected Variance}}{\text{Risk Tolerance}} - \frac{(\text{Expected Tracking Error})^2}{\text{Tracking Error Tolerance}} \\ - \frac{(\text{Expected Downside Tracking Error})^2}{\text{Downside Tracking Error Tolerance}} \end{aligned}$$

Formulating this objective function in matrix notation the objective is:

$$\text{MAX } X'E - \lambda X'\Sigma X - \psi X'\Sigma^* X - \alpha X'\Sigma'' X$$

The relative importance of the absolute risk, tracking error and the adjusted tracking error are controlled by the corresponding values of  $\lambda$ ,  $\psi$  and  $\alpha$ . This objective function can be decomposed into the MV, TE and DTE objective function as well as any combination of these three objective functions by adjusting the corresponding tolerance levels. For example, if the portfolio manager had no concern for the downside tracking error then downside tracking error tolerance would be infinite and  $\alpha$  would effectively assume the value zero and the objective function would decompose to the objective function of the MVTE model. Each portfolio manager might well have a different tolerance level for each of these measures of risk and effectively each of them would be consequently maximising a different objective function.

### 13. CONCLUSION

Recently much research has focussed on the objective function of asset allocation. The foundation of this research is the Markowitz mean-variance (MV) portfolio selection model. Many practitioners have been critical of the direct implementation of the Markowitz mean-variance model since this model often fails to provide satisfactory results in the form of realistic portfolios. The criticism has been levelled at the measurement of risk in this model. Franks (1992) and others believe that funds are often evaluated relative to a benchmark and consequently he proposed a tracking error (TE) portfolio selection model which incorporates this relative return measurement into the risk component of the objective function. The resultant portfolios achieved by applying the proposed TE portfolio selection model are certainly more diversified making them more appealing to portfolio managers. It is noticeable that portfolios comprising of a large portion of cash have relatively large tracking-errors and consequently the *optimal* portfolios obtained when using the tracking-error as a measure of risk comprise of *no* cash.

Chow (1995) recognised the significance of both models and he proposed a joint model that takes account of both the variance and tracking error of a portfolio in the objective function. The composition of the optimal portfolio in this mean-variance-tracking error (MVTE) model is certainly a compromise of the equivalent underlying MV and TE optimal portfolios. Consequently its diversification and perhaps its more realistic measurement of risk will certainly make it attractive to portfolio managers. We also considered a downside tracking error (DTE) portfolio selection model where the tracking error is curtailed by the sign of relative return i.e. whether its positive or negative. We conclude that DTE is only a meaningful measure of risk if the securities' returns are asymmetrically distributed.

## 14. APPENDIX

### 14.1 Appendix 1

In Section 4 we noted that the efficient frontier is linear if the benchmark is included in the opportunity set. We shall show why this is so in a simplified example that consists of only two assets in the opportunity set.

**SCENARIO:** There is a portfolio consisting of two assets viz. ASSET 1 and ASSET 2. ASSET 1 has a **Tracking Error (TE)** whereas ASSET 2 is the **Benchmark (B)** itself and consequently has no tracking error. We wish to show that the **TE efficient frontier is LINEAR!**

The following notation will be used for the proof:

- $X_i$  = Proportion of the funds invested in ASSET  $i$
- $R_i$  = Return on ASSET  $i$
- $R_p$  = Return on the PORTFOLIO
- $TE_{R_i}$  = Tracking Error of the variable  $R_i$
- $R_b$  = Return on the Benchmark

The portfolio consists of two assets viz. ASSET 1 and ASSET 2. Therefore if  $X_1$  is invested in ASSET 1 then  $X_2 = (1 - X_1)$  will have to be invested in ASSET 2.

The **Expected Return** of this portfolio is:

$$E(R_p) = X_1 \cdot E(R_1) + (1 - X_1) \cdot E(R_2)$$

The **Tracking Error squared** of this portfolio is:

$$\begin{aligned} TE_p^2 &= \text{Var}(R_p - R_b) \\ &= E[(R_p - R_b) - E(R_p - R_b)]^2 \end{aligned}$$

$$= E[(X_1(R_1 - R_B) + X_2(R_2 - R_B)) - E(X_1(R_1 - R_B) + X_2(R_2 - R_B))]^2$$

Rearranging

$$\begin{aligned} TE_p &= E[(X_1(R_1 - R_B) - E(X_1(R_1 - R_B))) + (X_2(R_2 - R_B) - E(X_2(R_2 - R_B)))]^2 \\ &= E[(X_1(R_1 - R_B) - X_1 E(R_1 - R_B)) + (X_2(R_2 - R_B) - X_2 E(R_2 - R_B))]^2 \\ &= E[X_1((R_1 - R_B) - E(R_1 - R_B)) + X_2((R_2 - R_B) - E(R_2 - R_B))]^2 \\ &= E\left[X_1^2((R_1 - R_B) - E(R_1 - R_B))^2 + X_2^2((R_2 - R_B) - E(R_2 - R_B))^2 \right. \\ &\quad \left. + 2X_1X_2((R_1 - R_B) - E(R_1 - R_B))((R_2 - R_B) - E(R_2 - R_B))\right] \\ &= X_1^2 E[(R_1 - R_B) - E(R_1 - R_B)]^2 + X_2^2 E[(R_2 - R_B) - E(R_2 - R_B)]^2 \\ &\quad + 2X_1X_2 E[(R_1 - R_B) - E(R_1 - R_B)][(R_2 - R_B) - E(R_2 - R_B)] \\ &= X_1^2 Var(R_1 - R_B) + X_2^2 Var(R_2 - R_B) + 2X_1X_2 Cov(R_1 - R_B; R_2 - R_B) \\ &= X_1^2 Var(R_1 - R_B) + X_2^2 Var(R_2 - R_B) + 2X_1X_2 \rho_{R_1 - R_B, R_2 - R_B} \sigma_{R_1 - R_B} \sigma_{R_2 - R_B} \\ &= X_1^2 TE^2(R_1) + X_2^2 TE^2(R_2) + 2X_1X_2 \rho_{R_1 - R_B, R_2 - R_B} TE(R_1)TE(R_2) \\ &= X_1^2 TE^2(R_1) + (1 - X_1)^2 TE^2(R_2) + 2X_1(1 - X_1) \rho_{R_1 - R_B, R_2 - R_B} TE(R_1)TE(R_2) \end{aligned}$$

If ASSET 2 is the BENCHMARK itself then its TRACKING ERROR is zero i.e.  $TE(R_2) = 0$ . The Tracking Error squared of the portfolio will now become:

$$\begin{aligned} TE_p^2 &= X_1^2 TE^2(R_1) + (1 - X_1)^2 (0) + 2X_1(1 - X_1) \rho_{R_1 - R_B, R_2 - R_B} TE(R_1)(0) \\ &= X_1^2 TE^2(R_1) \\ \therefore TE_p &= X_1 TE(R_1) \end{aligned}$$

Knowledge of the **mean** and the **tracking error** of a portfolio consisting of a single asset and the **benchmark** itself allows one to plot the efficient frontier. It is **LINEAR!** The proof is based on the slope of the efficient frontier being independent of  $X_1$ , the proportion of the portfolio invested in the asset that does have a tracking error.

**PROOF:**

Recall:

$$E(R_p) = X_1 \cdot E(R_1) + (1 - X_1) \cdot E(R_2)$$

and

$$TE_p = X_1 \cdot TE(R_1)$$

The *change in expected return* with respect to the *proportion* of the portfolio invested in ASSET 1 is:

$$\frac{dE(R_p)}{dX_1} = E(R_1) - E(R_2)$$

And the *change in the tracking error* with respect to the proportion of the portfolio invested in ASSET 1 is:

$$\frac{dTE_p}{dX_1} = TE(R_1)$$

Therefore the slope of the line is:

$$\frac{dE(R_p)}{dTE_p} = \frac{dE(R_p)/dX_1}{dTE_p/dX_1} = \frac{E(R_1) - E(R_2)}{TE(R_1)}$$

Consequently this frontier must be **LINEAR** because its slope does not change as  $X_1$ , the proportion of the portfolio invested in the asset with a tracking error, changes.

## 14.2 Appendix 2

We wish to show that the tracking error about a constant value e.g. a constant benchmark, is the same measure as the standard deviation.

**Proof:** The TE measured about a fixed benchmark is the same measure as the standard deviation

The following notation will be used

$RR_i$  = Relative Return of asset  $i$

$R_i$  = Return of asset  $i$

$R_b$  = Return of the benchmark

$c$  = A constant

$$RR_i = R_i - R_b$$

If  $R_b = c$  then  $RR_i = R_i - c$

$$TE_i^2 = \text{Var}(RR_i)$$

$$= \text{Var}(R_i - c)$$

$$= \text{Var}(R_i) + \text{Var}(c)$$

$$= \text{Var}(R_i)$$

$$\therefore TE_i = \text{STD}(R_i)$$

$$\Rightarrow x' \Sigma x = \sum_i \sum_j w_i w_j \sigma_{ij}$$

$$= \sum_i \sigma_i w_i^2$$

since  $\sigma_{ij} = 0 \forall i \neq j$

since  $\text{Cov}(b, RR_i) = 0$

**CHAPTER 4**

**INPUT ESTIMATION**

University of Cape Town

## 1. SYNOPSIS

Any portfolio manager would be interested in research aimed at attaining portfolios that tend to be more satisfactory to practitioners. In Chapter 3 the research focus was on the objective function of asset allocation. In this chapter we focus instead on the input estimates of the model formulation. We focus in particular on the central estimate, that of *return*. In Section 2 we give a brief background to the problem by first describing how expected returns have traditionally been estimated using historical means i.e. by the application of the Maximum Likelihood function. We argue that the Maximum Likelihood approach is not an acceptable model in the context of yielding realistic expected return estimates. Furthermore it is unable to capture personal expectations or views, which portfolio managers might well have regarding the expected returns of the assets in the opportunity set. In this chapter we consider an estimate that can well capture these personal expectations, that is, a Bayesian estimate.

Section 3 covers a comprehensive examination of the Bayesian estimate. Bayes proposes that the parameter we wish to estimate,  $\mu$ , the security's expected return, can actually be described by a probability density function ( $\pi(\mu)$ ). We also suggest that a sample of past returns will give us some information concerning the parameter,  $\mu$ , the security's expected return. We continue by arguing it is the Bayesian estimate that accounts for both of these sources of relevant information in its estimation of a security's expected return. The function from which the Bayesian estimate is extracted is termed a conditional probability density function ( $\pi(\mu|r_i)$ ). In summary this implies that a security's expected return,  $\mu$ , is influenced by its past returns, the  $r_i$ 's. In Sections 3.1 and 3.2 we describe the components of this function and we show how these components fit together when applying Bayes' Rule. In Section 3.3 we remind the reader that in the objective function we actually need to estimate both  $\mu$ , the security's expected return, and  $\sigma^2$ , the security's variance. In the earlier sections we have only examined  $\pi(\mu|r_i)$ , from which we have extracted an estimate for the security's expected return. In Section 3.3 we examine the conditional density function ( $\pi(\sigma^2|r_i)$ ) from which we shall estimate  $\sigma^2$ .

Thus far we have *stated* that the estimates for  $\mu$  and  $\sigma^2$  can be obtained from the functions  $\pi(\mu|r_i)$  and  $\pi(\sigma^2|r_i)$  respectively. In Section 3.4 we actually describe *how* the estimates  $\mu$  and  $\sigma^2$  are extracted from these two respective functions. As it turns out, this is achieved by minimising a Loss Function, which is a function of the conditional density function. The Bayesian estimate is the estimate that minimises this Loss Function.

The proposed Bayes' estimates are assessed empirically in Section 4. The Bayes' estimates and the consequent efficient frontiers are then compared to those obtained when using traditional sample estimates for the securities' expected returns and variances (which is effectively an application of the Maximum Likelihood Estimate, as described in Section 2).

## 2. INTRODUCTION

Portfolio managers should view the measurement of portfolios' returns and risks (variances) as central to their success. Their objective should be to select a portfolio that maximises its return at a given level of risk (variance), or similarly selecting a portfolio that minimises its risk at a given level of return. To put these objectives into practice estimates for returns, variances and covariances are needed. Many of the portfolio managers are quite content in using a covariance matrix, which is constructed from a sample of past returns, as an estimate of a portfolio's true covariance matrix. Equivalently the securities' expected returns are estimated by using the sample means of these past returns. If all portfolio managers use this optimisation technique and if these portfolio managers all use the same sample data then all of these portfolio managers would identify the same optimal portfolios.

This suggests that portfolio managers require little skill and are merely data managers. This could well be true when the market is in equilibrium and when all portfolio managers share the same expectations. In equilibrium the optimal portfolio is the market portfolio and all portfolio managers would select this market portfolio as their optimal portfolio. But in reality portfolio managers might well have different expectations with regard to the returns and risks of securities.

In the optimisation technique mentioned above, an estimate is needed for each security's expected return. Traditionally portfolio managers have used a simple sample mean as an estimate for a security's expected return. Based on the assumption that a security's returns are distributed normally portfolio managers have, perhaps quite unknowingly, selected a statistically "suitable" estimate for a security's expected return. We qualify this assertion below.

In order to justify our claim that this estimate is "suitable" we will describe how this estimate is actually obtained. We assume that we are presented with a series of a security's past returns viz.  $r_1, r_2, \dots, r_n$ . If all these returns are independent of one another then the joint likelihood function, which is the likelihood function that considers all the past returns simultaneously, can be expressed as the product of each

individual return's likelihood function. A security's likelihood function is identical to its probability density function except the variables are the known values, taken from the sample of past returns, and instead the parameters of the probability density function are considered to be the unknown variables. This likelihood function is given by:

$$L(\mu; r_1, r_2, \dots, r_n) = \prod_{i=1}^n L(\mu; r_i) \quad (1)$$

By the assumption that the security's return is distributed normally i.e.  $r_i \sim N(\mu, \sigma^2)$  where  $\mu$  is the expected return of the security and  $\sigma^2$  is the variance of the security's return this likelihood function can be expressed as:

$$L(\mu; r_1, r_2, \dots, r_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(r_i - \mu)^2}{\sigma^2}\right\}$$

The likelihood function  $L(\mu; r_1, r_2, \dots, r_n)$  has only one unknown variable viz.  $\mu$ . Clearly  $r_1, r_2, \dots, r_n$  are known since a sample of past returns are used as realisations of these values.  $\mu$  is unknown and it can be shown that the likelihood of the expected return,  $\mu$ , is maximised by letting  $\mu$  take on the value of the sample mean,  $\bar{r}$  (Rice<sup>1</sup>, 1995). Accordingly this estimate for the expected return,  $\mu$ , is termed the maximum likelihood estimate.

But in reality portfolio managers often don't share the same expectations regarding the market's expected performance and these personal or subjective expectations need to be considered when determining the optimal portfolio choice. The maximum likelihood estimation technique, which is what we described above, cannot accommodate these personal or subjective expectations. The reason is because in the likelihood function  $L(\mu; r_1, r_2, \dots, r_n)$ ,  $\mu$ , the expected return of the security, is treated as an unknown value, but this unknown value is considered to be a constant i.e. each portfolio manager believes that  $\mu$  will have the same (unknown) value.

Consequently we propose that a Bayesian estimate, which accommodates subjective expectations, be used as an estimate for a security's expected return rather than the traditional maximum likelihood estimate.

Section 3 gives a detailed discussion of the Bayesian estimate in the portfolio selection environment. In Section 3 we show that the Bayesian estimate allows a formulation which is in essence a compromise (weighting) between ones own personal expectations and the maximum likelihood estimate, which is the sample mean. If the reader is willing to accept this summary as quid pro quo or if the reader already has a thorough understanding of Bayesian estimators then the reader could well advance directly to Section 4, where we examine Bayesian estimates empirically.

---

<sup>1</sup> Rice, JA, "Mathematical Statistics and Data Analysis", 1995, pp. 256.

### 3. THE BAYESIAN ESTIMATOR TECHNIQUE

The unknown parameter we wish to estimate is  $\mu$ , a security's expected return. The maximum likelihood estimate focused upon this unknown parameter by using the likelihood function.  $\mu$ , the security's expected return, is the only *unknown* variable in the likelihood function. The likelihood function is also a function of a security's returns  $r_1, r_2, \dots, r_n$ , but these variables are treated as "known" by using the realisations of these returns obtained from a sample of past returns.

The Bayesian estimator also relies upon  $\mu$ , a security's expected return. This expected return is obtained using an approach which differs from the "traditional" maximum likelihood approach. In the Bayesian approach the expected return parameter  $\mu$  is described in terms of either a probability mass or probability density function,  $\pi(\mu)$ , rather than by a likelihood function. The Bayesian approach departs from the likelihood approach in that it considers  $\pi(\mu)$  as being subjective, and as such it may well differ from portfolio manager to portfolio manager, whereas the likelihood function is a function of *past returns* and will consequently be identical for each portfolio manager if they use the same sample.

A portfolio manager's own *subjective* expectations are captured by  $\pi(\mu)$ . The symbol " $\pi$ " is used rather than " $p$ " in order to indicate that this distribution is subjective. It will be shown that a Bayesian estimate is a function of  $\pi(\mu)$  and since  $\pi(\mu)$  is subjective the Bayesian estimate of a security's expected return would differ from portfolio manager to portfolio manager.

The Bayesian approach acknowledges that a sample of past returns contains information with respect to the security's expected return,  $\mu$ . Accordingly, the uncertainty of  $\mu$  can be adjusted to account for this additional information and this is achieved by means of the *conditional* probability function  $\pi(\mu | r_1, r_2, \dots, r_n)$ . This is the probability density function from which the Bayesian approach estimates the security's expected return.

We will first consider  $\pi(\mu | r_i)$  before  $\pi(\mu | r_1, r_2, \dots, r_n)$  and we will show that this conditional probability density function  $\pi(\mu | r_i)$  can be decomposed into three components.

### 3.1 Components of the Conditional Probability Density Function, $\pi(\mu | r_i)$

The three components of  $\pi(\mu | r_i)$  are:

$\pi(\mu)$  which is the subjective distribution of  $\mu$ .  $\pi(\mu)$  differs from portfolio manager to portfolio manager. This probability density function is termed the prior probability density function. This probability density function represents the uncertainty of the expected return  $\mu$  without any knowledge as to what the actual realisation of the security's return,  $R_i$ , is.

$L(r_i; \mu)$  which is the likelihood function of a past return.

$m(r_i)$  where  $m(r_i) = \int_{-\infty}^{\infty} \pi(\mu) L(r_i; \mu) d\mu$  (Winkler<sup>2</sup>, 1972) is termed the marginal density function of  $R_i$ , the security's return. It is a prediction of the expected likelihood sampling outcome (i.e.  $E(L(r_i; \mu))$ ) considering the current knowledge of  $\mu$  which is expressed in  $\pi(\mu)$ . The significance of this component is that it is totally independent of  $\mu$ , the security's expected return. The importance of this feature will be discussed later.

By the direct implementation of the Baye's rule (Winkler, 1972)

$$\pi(\mu | r_i) = \frac{\pi(\mu) L(r_i; \mu)}{m(r_i)} \quad (2)$$

We shall obtain an estimate of the security's expected return, *conditional* on the security's past return  $r_i$ , which contains information with respect to the security's

<sup>2</sup> Winkler, RL, "Introduction to Bayesian Inference and Decisions", 1972, pp. 144.

expected return, from this conditional probability function  $\pi(\mu | r_i)$ . The jargon *conditional* effectively implies that the security's expected return is influenced by its past returns.

This conditional probability function  $\pi(\mu | r_i)$  only considers a single past return, viz.  $r_i$ . A single past return may actually carry misinformation but we certainly believe that a *series* of past returns could well provide one with relevant information with regard to a security's expected return,  $\mu$ . Effectively we are interested in  $\pi(\mu | r_1, r_2, \dots, r_n)$  rather than simply in  $\pi(\mu | r_i)$ .

Consequently,

$$\pi(\mu | r_1, r_2, \dots, r_n) = \frac{\pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n)}{m(r_1, r_2, \dots, r_n)} \quad (\text{Winkler, 1972}) \quad (3)$$

where  $\Lambda(\mu; r_1, r_2, \dots, r_n)$  is the joint likelihood function of the past  $n$  returns. " $\Lambda$ " is the capital Greek letter L and is commonly used in texts to represent a *joint* likelihood function. In the simple case of a single past return the likelihood function is  $L(r_i | \mu)$ . When we consider a sequence of  $n$  *independent* random variables  $R_1, R_2, \dots, R_n$ , which are all drawn from the same distribution then the joint likelihood function,  $\Lambda(\mu; r_1, r_2, \dots, r_n)$ , is  $\prod_{i=1}^n L(r_i | \mu)$ , by the law of independence.

Considering  $\pi(\mu | r_1, r_2, \dots, r_n)$ , as stated in Equation ( 3 ), it is evident that only two out of the three components viz.  $\pi(\mu)$  and  $\Lambda(\mu; r_1, r_2, \dots, r_n)$  are functions of the security's expected return,  $\mu$ .  $m(r_1, r_2, \dots, r_n)$  is not a function of  $\mu$  and can consequently be absorbed into a proportionality constant when dealing with proportions. Thus, when considering proportionality, Equation ( 3 ), can be written as:

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n) \quad (4)$$

We shall now examine  $\pi(\mu)$  and  $\Lambda(\mu; r_1, r_2, \dots, r_n)$  separately and determine the effects which different forms of these functions have on  $\pi(\mu | r_1, r_2, \dots, r_n)$ .

### 3.1.1 Insights for Practitioners

- The practitioner is interested in the distribution of the security's expected return, based on the information contained in a sample of past returns. The Bayesian distribution describes this scenario. In order to formulate the Bayesian distribution one needs the distribution of the expected return, which the practitioner can describe him/herself, and the likelihood function of the security's past returns.

### 3.2 The Influence the Prior Distribution $\pi(\mu)$ and the Likelihood Function $\Lambda(\mu; r_1, r_2, \dots, r_n)$ have on the Conditional Probability Density Function $\pi(\mu | r_1, r_2, \dots, r_n)$

In Section 3.1 we established that

$$\pi(\mu | r_1, r_2, \dots, r_n) = \frac{\pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n)}{m(r_1, r_2, \dots, r_n)} \quad (5)$$

Only two of these components viz.  $\pi(\mu)$  and  $\Lambda(\mu; r_1, r_2, \dots, r_n)$  are functions of the security's expected return,  $\mu$ . Consequently we shall examine these two components and determine their effects upon  $\pi(\mu | r_1, r_2, \dots, r_n)$ , from which we shall ultimately obtain our estimate for the security's expected return, *conditional* on (affected by) the past returns.

Firstly, in Section 3.1 we established that

$$\Lambda(\mu; r_1, r_2, \dots, r_n) = \prod_{i=1}^n L(r_i | \mu) \quad (6)$$

If we assume that the return of the security at time  $i$  is distributed normally with parameters  $\mu$  and  $\sigma^2$  i.e.  $r_i \sim N(\mu, \sigma^2)$  then

$$\begin{aligned}
\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) &= \prod_{i=1}^n L(r_i | \mu, \sigma^2) \\
&= \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{1}{2}} \exp \left[ -\frac{(r_1 - \mu)^2}{2\sigma^2} \right] \dots \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{1}{2}} \exp \left[ -\frac{(r_n - \mu)^2}{2\sigma^2} \right] \\
&= \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^n (r_i - \mu)^2 \right]
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{i=1}^n (r_i - \mu)^2 &= \sum_{i=1}^n (r_i - \bar{r} + \bar{r} - \mu)^2 \\
&= \sum_{i=1}^n \left[ (r_i - \bar{r})^2 + 2(r_i - \bar{r})(\bar{r} - \mu) + (\bar{r} - \mu)^2 \right] \\
&= \sum_{i=1}^n (r_i - \bar{r})^2 + 2(\bar{r} - \mu) \sum_{i=1}^n (r_i - \bar{r}) + n(\bar{r} - \mu)^2 \\
&= (n-1)s^2 + n(\bar{r} - \mu)^2 \quad \text{since } \sum_{i=1}^n (r_i - \bar{r}) = 0
\end{aligned}$$

The summation term has been simplified and this was achieved by summing it out and introducing the mean,  $\bar{r}$ .

Therefore,

$$\begin{aligned}
\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) &= \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\bar{r} - \mu)^2 \right\} \right] \\
&= \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{n}{2}} \exp \left[ -\frac{(n-1)s^2}{2\sigma^2} \right] \exp \left[ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right] \quad (7)
\end{aligned}$$

We shall assume that  $\sigma^2$ , the variance of the security's returns is known, and is consequently not a variable (this assumption will be relaxed later). The only term of (7) which is a function of the security's expected return,  $\mu$ , is the term  $\exp \left[ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right]$ . Consequently,

$$\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) \propto \exp \left[ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right] \quad (8)$$

In Section 3.1 Equation (4) we established

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n)$$

Hence from Equation (8)

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \pi(\mu) \exp \left[ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right] \quad (\text{Winkler}^3, 1972) \quad (9)$$

We now consider  $\pi(\mu)$ .  $\pi(\mu)$  is the subjective prior distribution of the parameter  $\mu$ , which is the security's expected return. In simple terms,  $\pi(\mu)$  is the probability density function that describes a portfolio manager's uncertainty in the security's expected return.  $\pi(\mu)$ , may well differ from portfolio manager to portfolio manager and consequently  $\pi(\mu | r_1, r_2, \dots, r_n)$ , the conditional probability density function, from which the portfolio manager obtains an estimate for  $\mu$  will differ from portfolio manager to portfolio manager. We shall assume the prior information that the portfolio manager has regarding the parameter  $\mu$ , the security's expected return, can be represented by a normal distribution with mean  $\mu_0$  and variance  $\sigma_0^2$ .  $\mu_0$  represents the return which the portfolio manager personally and subjectively expects the security to achieve. These subjective expectations can be obtained from predictor models or the portfolio manager can simply state his expectation.  $\sigma_0^2$  is the variance of the security's expected return. It is used to represent the confidence that the portfolio manager has in his own expectation of the security's expected return. The more confident a portfolio manager is in his own expectation of the security's return the smaller the value he assigns to  $\sigma_0^2$ . We have assumed  $\mu \sim N(\mu_0, \sigma_0^2)$  hence

<sup>3</sup> Winkler, RL, "Introduction to Bayesian Inference and Decisions", 1972, pp. 169.

$$\pi(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right] \quad (10)$$

We can replace  $\pi(\mu)$  with Equation (10) in Equation (9) and consequently

$$\begin{aligned} \pi(\mu | r_1, r_2, \dots, r_n) &\propto \exp\left[-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right] \exp\left[-\frac{n(\bar{r} - \mu)^2}{2\sigma^2}\right] \\ &\propto \exp\left[-\frac{1}{2} \left\{ \frac{(\mu - \mu_0)^2}{\sigma_0^2} + \frac{n(\bar{r} - \mu)^2}{\sigma^2} \right\}\right] \end{aligned} \quad (11)$$

By completion of the square in  $\mu$ , the term in braces can be seen to be expressible as

$$\frac{(\mu - \mu_1)^2}{\sigma_1^2} + \text{terms not involving } \mu \quad (\text{Rice}^4, 1995)$$

where

$$\mu_1 = \frac{\sigma_1^2}{\sigma_0^2} \mu_0 + \sigma_1^2 \frac{n}{\sigma^2} \bar{r} \quad (12)$$

and

$$\sigma_1^{-2} = \sigma_0^{-2} + \frac{n}{\sigma^2} \quad (13)$$

From Equation (12) it is evident that the posterior mean,  $\mu_1$ , which is a measure of a security's expected return conditional on the sample of past returns,  $r_1, r_2, \dots, r_n$ , is a weighted average of the portfolio manager's own subjective expectation,  $\mu_0$ , in the

<sup>4</sup> Rice, JA, "Mathematical Statistics and Data Analysis", 1995, pp. 589.

security's return, and of the sample mean,  $\bar{r}$ . The sum of these two weights viz.  $\frac{\sigma_1^2}{\sigma_0^2}$  and  $\sigma_1^2 \frac{n}{\sigma^2}$  is 1 (see Appendix 1 for the proof).

The more certain a portfolio manager is in his/her own subjective expectations the greater is the weight  $\frac{\sigma_1^2}{\sigma_0^2}$ . As his/her certainty increases i.e. as  $\sigma_0^2 \rightarrow 0$  so the weight

$\frac{\sigma_1^2}{\sigma_0^2} \rightarrow 1$  (see Appendix 2 for the proof). Similarly, as the portfolio manager's certainty

in his/her own expectation increases so the weight  $\sigma_1^2 \frac{n}{\sigma^2}$  decreases and tends towards 0. Consequently, if a manager is totally certain in his/her own expectation then he/she will simply use his/her subjective expectation,  $\mu_0$ , as an estimate for the security's expected return,  $\mu_1$ , and he/she will not consider the security's past returns, i.e. he/she will not consider  $\bar{r}$ .

It is more likely that a portfolio manager will have some doubt in his/her own expectation of the security's return and as a result both his/her own expectation  $\mu_0$ , as well as the mean,  $\bar{r}$ , of a sample of the security's past returns,  $r_1, r_2, \dots, r_n$ , will contribute towards a measure of the security's expected return, which is conditional on these past returns.

In this Section there were two key assumptions that we made. Firstly, we assumed that a security's variance,  $\sigma^2$ , was known and secondly we assumed that the subjective prior distribution,  $\pi(\mu)$ , follows a normal distribution. We shall now examine the consequences when we relax these assumptions.

### 3.3 The Effect on the Conditional Probability Density $\pi(\mu|r_1, r_2, \dots, r_n)$ , when altering the Prior

In Section 3.2 we assumed that the variance of a security's return was known. We now consider the situation where this variance is unknown and it is treated as a variable. Consequently we are dealing with two unknown variables viz.  $\mu$ , the security's expected return, and  $\sigma^2$ , the security's variance. In Section 3.2 we represented the portfolio manager's own subjective views on the security's expected return,  $\mu$ , by the distribution  $\pi(\mu)$ . Similarly, when both  $\mu$  and  $\sigma^2$  are unknown we simultaneously capture the portfolio manager's views on the security's expected return,  $\mu$ , and the security's variance,  $\sigma^2$ , by the *joint* probability density function  $\pi(\mu, \sigma^2)$ . It is termed the *joint* probability density function because there are two unknown variables described by the distribution function rather than just a single variable.

In Section 3.1 Equation ( 4 ) it was established that

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n)$$

Similarly in the scenario of two unknown variables it can be shown

$$\pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n) \propto \pi(\mu, \sigma^2) \Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) \quad (14)$$

$\pi(\mu|r_1, r_2, \dots, r_n)$  can be obtained from Equation ( 14 ) by integrating Equation ( 14 ) with respect to  $\sigma^2$  and similarly  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  can be extracted from Equation ( 14 ) by integrating Equation ( 14 ) with respect to  $\mu$ .  $\pi(\mu|r_1, r_2, \dots, r_n)$  is the distribution of the security's expected return,  $\mu$ , when considering the information contained in a sample of past returns  $r_1, r_2, \dots, r_n$ , and similarly  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  is the distribution of the security's variance (uncertainty) conditional on a series of past returns. These are two important distributions since one can use them to obtain estimates of a security's expected return and a security's variance respectively and these two estimates are needed in order to construct an efficient frontier.

In Section 3.2 we assumed that a portfolio manager's own subjective expectations of a security's return could be represented by a normal distribution with mean  $\mu_0$  and variance  $\sigma_0^2$ . We further assumed that the variance (uncertainty) of a security's return was known. In that scenario it was established that the security's *conditional* expected return,  $\pi(\mu|r_1, r_2, \dots, r_n)$ , which takes the security's past returns into account, follows a normal distribution with parameters  $\mu_1$  and  $\sigma_1^2$ .  $\mu_1$  was shown to be a weighted average of a portfolio manager's personal belief in the expected return of the security and the sample mean of a series of the security's past returns. Similarly,  $\sigma_1^2$  is a combination of the portfolio manager's certainty (variance) in his own expectation and the sample's variance.

We now consider circumstances where we believe a normal distribution can't or won't best represent a portfolio manager's personal expectation in a security's return. Consequently, the security's expected return, which is conditional/influenced on/by its past returns and which is obtained from  $\pi(\mu|r_1, r_2, \dots, r_n)$  won't necessarily be normally distributed. We will also be considering scenarios where the security's uncertainty (variance) is not known for certain and consequently we will be also be dealing with the conditional distribution  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  which describes the return's uncertainty (variance).

If the portfolio manager is not totally sure of the security's uncertainty (variance), this uncertainty together with the portfolio manager's uncertainty in the security's expected return could be represented by the joint distribution  $\pi(\mu, \sigma^2)$ . The distribution  $\pi(\mu, \sigma^2)$  is a function of two parameters. If any information given to one about one of the two parameters does not affect the level of uncertainty of the other parameter then the two parameters are termed to be *independent* of one another. More explicitly, if the uncertainty in a security's expected return does not affect the uncertainty in a security's variance then the security's expected return and variance are termed to be *independent* of one another. If these two parameters are independent of one another then by the law of independence

$$\pi(\mu, \sigma^2) = \pi(\mu) \cdot \pi(\sigma^2) \quad (15)$$

In such a scenario the portfolio manager need not specify the single *joint* distribution  $\pi(\mu, \sigma^2)$  to represent his uncertainty in the security's expected return,  $\mu$ , and variance,  $\sigma^2$ , but rather he/she could specify separate distributions for  $\mu$  and  $\sigma^2$ . When these two separate distributions are multiplied together one then obtains the *joint* distribution,  $\pi(\mu, \sigma^2)$ , for  $\mu$  and  $\sigma^2$ .

We must emphasise that we are ultimately interested in  $\pi(\mu|r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  (which are the conditional distribution of the security's expected return,  $\mu$ , and the conditional distribution of the security's variance,  $\sigma^2$  respectively), when taking the information concerning these two parameters contained in a sample of past returns into account. We shall employ these two distributions to obtain estimates for a security's expected return and variance. These two estimates will be utilised for each security in an opportunity set from which an efficient frontier will be constructed. It is evident from Equation ( 14 ) that the two distributions  $\pi(\mu|r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  are dependent upon the likelihood function,  $\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n)$ , which we have assumed to be a normal distribution, and the *joint* distribution,  $\pi(\mu, \sigma^2)$ , which represents the portfolio manager's uncertainty in the security's expected return,  $\mu$ , and variance,  $\sigma^2$ , simultaneously. In Equation ( 15 ) it was established that if the two parameters  $\mu$  and  $\sigma^2$  are independent of one another then  $\pi(\mu, \sigma^2) = \pi(\mu).\pi(\sigma^2)$ . Consequently, if one knows the likelihood function  $\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n)$ , and the joint distribution function  $\pi(\mu, \sigma^2)$  (or,  $\pi(\mu)$  and  $\pi(\sigma^2)$  if  $\mu$  and  $\sigma^2$  are independent of one another) then one can determine the distributions  $\pi(\mu|r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$ .

Table 1 describes various scenarios and lists the family of distribution that we believe would best represent the portfolio manager's views on the security's expected return and variance in each of these scenarios. The table also lists the consequent conditional distributions  $\pi(\mu|r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  from which we shall obtain estimates for the security's expected return and the security's variance (risk) respectively. In the table we have merely stated the appropriate functions. Appendix 3 and Appendix 4 give detailed accounts of how these appropriate functions are attained, or they at least justify why we selected the functions listed in the table.

**Table 1 Distributions best representing the Portfolio Manager's expectations and the consequent conditional distribution functions**

| Scenario                                                                                                                                                                                                                                                                                                                                                                                                                                   | Subjective Distribution<br>of $\mu$<br>$\pi(\mu)$                                                                                                                                                                                                                                                    | Subjective Distribution<br>of $\sigma^2$<br>$\pi(\sigma^2)$                                                                                                                                                                                                                                                                                                                                                                          | Joint Distribution<br>$\pi(\mu, \sigma^2)$          | Posterior "Conditional"<br>Distributions                                                                                                                                                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Any situation where the portfolio manager has some subjective belief in a security's expected return, $\mu$ . He is, however, totally certain of the security's variance, $\sigma^2$ .                                                                                                                                                                                                                                                     | $\pi(\mu) \sim N(\mu_0, \sigma_0^2)$<br><br>The portfolio manager needs to stipulate a value for $\mu_0$ and $\sigma_0^2$ respectively.                                                                                                                                                              | We assume that $\sigma^2$ is known and is consequently not a variable. Hence it is not represented by a distribution.                                                                                                                                                                                                                                                                                                                | Not necessary.                                      | $\pi(\mu r_1, r_2, \dots, r_n) \sim N(\mu_1, \sigma_1^2)$<br><br>where<br>$\mu_1 = \frac{\sigma_1^2}{\sigma_0^2} \mu_0 + \sigma_1^2 \frac{n}{\sigma^2} \bar{r} \text{ and}$ $\sigma_1^{-2} = \sigma_0^{-2} + \frac{n}{\sigma^2}$ |
| The portfolio manager has <u>no</u> subjective views on the security's expected return, $\mu$ , or on the security's uncertainty (variance), $\sigma^2$ . The portfolio manager believes that the parameter, $\mu$ , the security's expected return, is independent of the security's variance, $\sigma^2$ i.e. any information given to one about one of the two parameters will not affect the level uncertainty in the other parameter. | $\pi(\mu) = k \quad -\infty < \mu < \infty$<br>$k \geq 0$<br><br>$\mu$ is distributed <i>uniformly</i> over the interval $-\infty$ to $\infty$ i.e. $\mu$ is equally likely to take on <i>any</i> value over the interval $-\infty$ to $\infty$ .<br><br>$k$ is any arbitrary constant. <sup>5</sup> | $\pi(\ln \sigma^2) = k \quad -\infty < \sigma^2 < \infty$<br>$k \geq 0$<br><br>Equivalently<br>$\pi(\sigma^2) = \frac{k}{\sigma^2}$<br><br>$\ln(\sigma^2)$ is distributed <i>uniformly</i> over the interval $-\infty$ to $\infty$ . $k$ is any arbitrary constant. $\sigma^2$ couldn't be considered to be distributed <i>uniformly</i> over the interval $-\infty$ to $\infty$ since $\sigma^2$ is never less than 0. <sup>6</sup> | $\pi(\mu, \sigma^2) = \pi(\mu) \cdot \pi(\sigma^2)$ | $\pi(\mu r_1, r_2, \dots, r_n) \sim t_{n-1}$ <sup>7</sup><br>$\pi(\sigma^2 r_1, r_2, \dots, r_n) \sim \text{Gamma}(\alpha, \beta)$ <sup>8</sup><br><br>where<br>$\alpha = \frac{n-1}{2} \text{ and } \beta = \frac{(n-1)s^2}{2}$ |

<sup>5</sup> See Appendix 3 for a detailed explanation.<sup>6</sup> See Appendix 3 for a detailed explanation.<sup>7</sup> See Appendix 3 for a detailed explanation.<sup>8</sup> See Appendix 3 for a detailed explanation.

| Scenario                                                                                                                                                                                                                                                                                                                    | Subjective Distribution<br>of $\mu$<br>$\pi(\mu)$       | Subjective Distribution<br>of $\sigma^2$<br>$\pi(\sigma^2)$ | Joint Distribution<br>$\pi(\mu, \sigma^2)$                                                                                                                                                             | Posterior "Conditional"<br>Distribution                                                                                                                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| The portfolio manager has no subjective views on the security's expected return, $\mu$ , or on the security's uncertainty (variance), $\sigma^2$ . However, the portfolio manager does believe that the two parameters $\mu$ and $\sigma^2$ do influence the outcome of one another i.e. they are dependent on one another. | Not necessary, it is captured in $\pi(\mu, \sigma^2)$ . | Not necessary, it is captured in $\pi(\mu, \sigma^2)$ .     | $\pi(\mu, \sigma^2) = \frac{k}{\sigma^2}$ <p>where</p> $-\infty < \mu < \infty$ $-\infty < \sigma^2 < \infty$ <p><math>k \geq 0</math> where <math>k</math> is any arbitrary constant.<sup>9</sup></p> | $\pi(\mu   r_1, r_2, \dots, r_n) \sim t_{n-1}$ <sup>10</sup><br>$\pi(\sigma^2   r_1, r_2, \dots, r_n) \sim \text{Gamma}(\alpha, \beta)$ <sup>11</sup><br><p>where</p> $\alpha = \frac{n-3}{2} \text{ and } \beta = \frac{(n-1)s^2}{2}$ |

<sup>9</sup> See Appendix 4 for a detailed explanation.

<sup>10</sup> See Appendix 4 for a detailed explanation.

<sup>11</sup> See Appendix 4 for a detailed explanation.

Appendix 3 and Appendix 4 give the justifications for the choice of the functions, which are simply listed in Table 1. The most important columns in the table are the first column, which describes a realistic scenario, and the last column, which describes the *consequent* conditional (i.e. influenced by a sample of past returns) distribution for the security's expected return,  $\mu$ , and similarly a consequent conditional distribution for  $\sigma^2$ , a security's variance.

From Table 1 we have established the conditional probability density functions for  $\mu$ , the security's expected return, and  $\sigma^2$ , a security's variance. In reality we do not actually know the true values for  $\mu$  and  $\sigma^2$  and we use their corresponding conditional probability density functions to estimate the respective values. We shall examine how we extract the estimates for the two parameters from these Posterior "Conditional" probability density functions in Section 3.4.

### 3.4 Obtaining Estimates for a Security's Expected Return and Variance

In order to construct an efficient frontier from an opportunity set of securities one needs an estimate of each security's expected return, an estimate of each security's variance, which is a measure of the security's risk and the covariance between each pair of securities. As described in Section 2, portfolio managers often simply take a sample of past returns and use the sample mean as an estimate of the security's expected return, and similarly use the sample variance as an estimate of the security's risk.

We propose extracting an estimate of a security's expected return from the distribution  $\pi(\mu|r_1, r_2, \dots, r_n)$ .  $\pi(\mu|r_1, r_2, \dots, r_n)$  is the distribution of a security's expected return,  $\mu$ , which also captures the information concerning a security's expected return, that is contained in a sample of past returns. Similarly we shall employ the distribution  $\pi(\sigma^2|r_1, r_2, \dots, r_n)$  to obtain an equivalent estimate of the security's risk (variance).

Considering a security's expected return,  $\mu$ , we ideally wish to obtain an estimate,  $\hat{\mu}$ , for this expected return that is equal to the actual expected return i.e. ideally  $\hat{\mu} =$

unknown  $\mu$ . Realistically one has to settle for  $\hat{\mu}$  being close to  $\mu$  but not equal to  $\mu$ , due to one's uncertainty. Effectively one will be "losing out" if  $\hat{\mu} \neq \mu$ . A *loss function*,  $L(\hat{\mu}, \mu)$ , is a measure of this loss. If the estimate  $\hat{\mu}$  equals  $\mu$  then there will be no loss, but if  $\hat{\mu} \neq \mu$  then there is a positive loss, no matter whether  $\hat{\mu} > \mu$ , or whether  $\hat{\mu} < \mu$ .

In summary,

$$\begin{aligned} L(\hat{\mu}, \mu) &= 0 && \text{for } \hat{\mu} = \mu. \\ &> 0 && \text{for } \hat{\mu} \neq \mu. \end{aligned}$$

$\mu$ , the security's expected return, is a random variable, consequently  $L(\hat{\mu}, \mu)$ , the loss function, is also a random variable.

We are interested in the loss we shall obtain when using the function  $\pi(\mu | r_1, r_2, \dots, r_n)$  to estimate the security's expected return. This expected loss,  $EL(\hat{\mu} | r_1, r_2, \dots, r_n)$ , can be calculated by performing the following computation

$$EL(\hat{\mu} | r_1, r_2, \dots, r_n) = \int L(\hat{\mu}, \mu) \pi(\mu | r_1, r_2, \dots, r_n) d\mu \quad (\text{Press}^{12}, 1989) \quad (16)$$

We are actually not too interested in the quantity of loss obtained when using an estimate,  $\hat{\mu}$ , for a security's expected return, but rather we wish to determine the estimate  $\hat{\mu}$  that will minimise this loss.

We need to "practically" define the loss function,  $L(\hat{\mu}, \mu)$ . If the loss function,  $L(\hat{\mu}, \mu)$ , is relatively smooth and symmetrical around  $\hat{\mu} \neq \mu$ , then it can be approximated by a quadratic loss function (Press, 1989), i.e.

$$L(\hat{\mu}, \mu) = k(\hat{\mu} - \mu)^2 \quad (17)$$

<sup>12</sup> Press, SJ, "Bayesian Statistics: Principles, Models and Applications", 1989, pp. 17, pp. 27.

This function has been labelled as the *squared error loss function*. Applying the squared error loss function Equation ( 16) can be rewritten as

$$EL(\hat{\mu} | r_1, r_2, \dots, r_n) = \int k(\hat{\mu} - \mu)^2 \pi(\mu | r_1, r_2, \dots, r_n) d\mu \quad (18)$$

This expected loss function,  $EL(\hat{\mu} | r_1, r_2, \dots, r_n)$ , is minimised by taking  $\hat{\mu}$  to be the mean of  $\pi(\mu | r_1, r_2, \dots, r_n)$  (Press, 1989 and Geweke<sup>13</sup>, 1996). We have been concentrating on obtaining the best estimate for a security's expected return i.e.  $\hat{\mu}$ . Following the same argument, the expected loss function,  $EL(\hat{\sigma}^2 | r_1, r_2, \dots, r_n)$  is minimised by taking  $\hat{\sigma}^2$  to be the mean of  $\pi(\sigma^2 | r_1, r_2, \dots, r_n)$ . This again emphasises the importance of the two probability density functions  $\pi(\mu | r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2 | r_1, r_2, \dots, r_n)$ .

In Table 1 we described different scenarios of a portfolio manager's beliefs in his own expectations of a security's return and risk. The last column of the table described the distributions  $\pi(\mu | r_1, r_2, \dots, r_n)$  and  $\pi(\sigma^2 | r_1, r_2, \dots, r_n)$  which captured both the portfolio manager's own expectations as well as the information contained in a series of a security's past returns. We have established that when using the squared error loss function the best estimate for  $\mu$ , the security's expected return, is the mean of  $\pi(\mu | r_1, r_2, \dots, r_n)$  and the best estimate for  $\sigma^2$  the security's risk (variance) is the mean of  $\pi(\sigma^2 | r_1, r_2, \dots, r_n)$ . These two estimates are listed in the last two columns of Table 2.

<sup>13</sup> Geweke, J, "Contemporary Bayesian Econometrics", 1996, pp. 2-17.

**Table 2 The best Bayesian Estimates for a Security's Expected Return and Variance in the situations described**

| Scenario                                                                                                                                                                                                                                                                                                                                                                                                                            | $\pi(\mu r_1, r_2, \dots, r_n)$                                                                                                                                                                                                     | $\pi(\sigma^2 r_1, r_2, \dots, r_n)$                                                                                                                             | $\hat{\mu}$                                                                                        | $\hat{\sigma}^2$                                                 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Any situation where the portfolio manager has some subjective belief in a security's expected return, $\mu$ . He is, however, totally certain of the security's variance, $\sigma^2$ .                                                                                                                                                                                                                                              | $\pi(\mu r_1, r_2, \dots, r_n) \sim N(\mu_1, \sigma_1^{-2})$<br>where<br>$\mu_1 = \frac{\sigma_1^{-2}}{\sigma_0^{-2}} \mu_0 + \sigma_1^{-2} \frac{n}{\sigma^2} \bar{r}$ and<br>$\sigma_1^{-2} = \sigma_0^{-2} + \frac{n}{\sigma^2}$ |                                                                                                                                                                  | $\hat{\mu} = \frac{\sigma_1^{-2}}{\sigma_0^{-2}} \mu_0 + \sigma_1^{-2} \frac{n}{\sigma^2} \bar{r}$ |                                                                  |
| The portfolio manager has no subjective views on the security's expected return, $\mu$ , or on the security's uncertainty (variance), $\sigma^2$ . The portfolio manager believes that the parameter, $\mu$ , the security's expected return, is independent of the security's variance, $\sigma^2$ i.e. any information given to one about one of the two parameters will not affect the level uncertainty in the other parameter. | $\pi(\mu r_1, r_2, \dots, r_n) \sim t_{n-1}$ <sup>14</sup>                                                                                                                                                                          | $\pi(\sigma^2 r_1, r_2, \dots, r_n) \sim \text{Gamma}(\alpha, \beta)$ <sup>15</sup><br>where<br>$\alpha = \frac{n-1}{2} \text{ and } \beta = \frac{(n-1)s^2}{2}$ | $\hat{\mu} = \bar{r}$ <sup>16</sup>                                                                | $\hat{\sigma}^2 = \frac{\beta}{\alpha}$ <sup>17</sup><br>$= s^2$ |

<sup>14</sup> See Appendix 3 for a detailed explanation.<sup>15</sup> See Appendix 3 for a detailed explanation.<sup>16</sup> See Appendix 3 for a detailed explanation.<sup>17</sup> See Appendix 3 for a detailed explanation.

| Scenario                                                                                                                                                                                                                                                                                                                    | $\pi(\mu r_1,r_2,\dots,r_n)$                            | $\pi(\sigma^2 r_1,r_2,\dots,r_n)$                                                                                                                     | $\hat{\mu}$                         | $\hat{\sigma}^2$                                                                 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|----------------------------------------------------------------------------------|
| The portfolio manager has no subjective views on the security's expected return, $\mu$ , or on the security's uncertainty (variance), $\sigma^2$ . However, the portfolio manager does believe that the two parameters $\mu$ and $\sigma^2$ do influence the outcome of one another i.e. they are dependent on one another. | $\pi(\mu r_1,r_2,\dots,r_n) \sim t_{n-1}$ <sup>18</sup> | $\pi(\sigma^2 r_1,r_2,\dots,r_n) \sim \text{Gamma}(\alpha,\beta)$ <sup>19</sup><br>where<br>$\alpha = \frac{n-3}{2}$ and $\beta = \frac{(n-1)s^2}{2}$ | $\hat{\mu} = \bar{r}$ <sup>20</sup> | $\hat{\sigma}^2 = \frac{\beta}{\alpha}$<br>$= \frac{n-1}{n-3} s^2$ <sup>21</sup> |

<sup>18</sup> See Appendix 4 for a detailed explanation.  
<sup>19</sup> See Appendix 4 for a detailed explanation.  
<sup>20</sup> See Appendix 4 for a detailed explanation.  
<sup>21</sup> See Appendix 4 for a detailed explanation.

In Table 2 we have simply listed the conditional probability density functions and the consequent Bayesian estimates for a security's expected return,  $\hat{\mu}$  and variance,  $\hat{\sigma}^2$ . Appendix 3 and Appendix 4 contain detailed explanations of these conditional probability density functions and the consequent Bayesian estimates.

### ***3.5 The Effect Bayesian Estimates have on Efficient Frontiers***

Markowitz (1959) proposed that an investor's choice among portfolios should be based solely on the expected return and uncertainty associated with that return. A portfolio's variance is a measure of this uncertainty. Consequently, when implementing the Markowitz model an estimate of each security's return and variance is needed. Sample estimates are frequently used for this purpose, but we have proposed using Bayesian estimates.

In accordance to the Markowitz model if two portfolios have the same risk (variance), the one with the larger expected return is preferred. Similarly if two portfolios have the same expected return, the one with less risk (variance) is preferred. Whence originates the definition of an "efficient" portfolio. A portfolio is considered to be "efficient" if there is no portfolio which has a greater expected return without incurring greater risk and similarly it is efficient if it is impossible to obtain a portfolio with smaller risk (variance) without decreasing the expected return. When one plots each of the efficient portfolios in expected return-risk space the resultant curve is termed the efficient frontier.

We wish to compare the efficient frontier generated from sample estimates of the portfolio's expected return,  $\mu$ , and variance,  $\sigma^2$ , to the efficient frontier generated using the proposed Bayesian estimates.

We shall first consider scenario 1, which was described in both Table 1 and in Table 2. In this scenario we are assuming that the portfolio manager has a subjective belief in the securities' expected returns. It was established that the Bayesian estimate of a security's expected return,  $\hat{\mu}$ , is a weighted average of the portfolio manager's

subjective expectation in the security's expected return,  $\mu_0$ , and the sample mean  $\bar{r}$ . In scenario 1 we also assumed that the portfolio manager knows the securities' variances. If he/she obtained these "known" values from a sample of past returns then these values would be identical to the sample variance estimates. Consequently, if the portfolio manager's expectations of the securities' returns are the same as the mean values obtained from the sample of past returns, the Bayesian estimates of the securities' expected returns and variances would coincide with the sample estimates. As a result the efficient frontier generated from the two different estimators will be identical.

But we have proposed using Bayesian estimates because we believe portfolio managers will certainly have personal expectations and these expectations may certainly differ from the sample estimates. Hence  $\mu_0 \neq \bar{r}$ . Accordingly the efficient frontier created using the Bayesian estimates will certainly differ in shape and in location to the efficient frontier formed from sample estimates (even if we were to assume that the securities' known variances were the same as the sample estimates of the variances). Hence the optimal portfolio generated from the application of each of the two different estimators may certainly differ. We have performed some empirical research that examines this issue.

## 4. EMPIRICAL RESEARCH

### 4.1 Outline of Empirical Research

In Section 3.4 we noted the Bayesian estimate for a security's expected return,  $\hat{\mu}$ , when the portfolio manager is totally certain of the security's variance,  $\sigma^2$ , to be:

$$\hat{\mu} = \frac{\sigma_1^2}{\sigma_0^2} \mu_0 + \sigma_1^2 \frac{n}{\sigma^2} \bar{r} \quad (19)$$

Effectively the Bayesian estimate  $\hat{\mu}$  is a weighted average of the portfolio manager's own subjective expectation,  $\mu_0$  and the sample mean,  $\bar{r}$ . In Appendix 1 we showed the sum of these two weights  $\frac{\sigma_1^2}{\sigma_0^2}$  and  $\sigma_1^2 \frac{n}{\sigma^2}$  to be 1. Past research has shown the sample variance of a security's past returns to be a good estimate of a security's expected variance  $\sigma^2$  (Divecha<sup>22</sup>, 1994).

We now wish to consider ways of obtaining  $\mu_0$ , the portfolio manager's own subjective expectation of the security's return. A truly subjective, and probably very error prone manner in obtaining  $\mu_0$ , would simply be for the portfolio manager to "guess" values for  $\mu_0$ . We rather propose using models to obtain values for  $\mu_0$ , the portfolio manager's subjective expectation in the security's return. Consequently the Bayesian estimate for the security's expected return,  $\hat{\mu}$  will be a weighted average of the expectation obtained from the proposed *model* and the sample mean,  $\bar{r}$ .

The models we propose using in order to gain values for  $\mu_0$ , the subjective expectation of a security's return are:

<sup>22</sup> Divecha, MK, "Emerging Markets and Risk Reduction", Global Asset Allocation, 1994, pp. 340.

1.  $\mu_0$  is set equal to the overall mean return of all the assets in the opportunity set (Michaud<sup>23</sup>, 1989).
2.  $\mu_0$  is set equal to the expected return of the portfolio with the minimum variance (which has been calculated using the sample means and sample variances of past returns of the securities in the opportunity set as estimates of these securities' respective expected returns and variances) (Michaud, 1989).
3.  $\mu_0$  is set equal to the expected return obtained from the application of the CAPM (Capital Asset Pricing Model).

Again we wish to emphasise that the Bayesian estimate of a security's expected return consists of two weights. The more certain the portfolio manager is in his/her own expectation,  $\mu_0$ , the greater the weight that is placed on this expectation. We propose two methods for obtaining these two weights:

1. We simply assume that the portfolio manager is very certain in his own expectation,  $\mu_0$ , (which is obtained from the models proposed. Consequently  $\mu_0$  will have a weight of 90% and  $\bar{r}$  will have a weight of 10%).
2. The weights are obtained as directed in Equation ( 19 ) where  $\frac{\sigma^2}{n}$  is the variance of a sample mean and  $\sigma_0^2$  is the standard error of the estimate obtained using one of the proposed models.

Effectively we shall be examining 6 Bayesian estimates for a security's expected return. We shall be comparing the 3 Bayesian estimates obtained using the assumed 90% weighting with one another, and then separately we shall compare the 3 Bayesian estimates obtained using the weightings extracted from the standard errors of the estimators obtained from the proposed models with one another.

In each scenario we shall compare the *optimal portfolios* obtained when using each of the proposed Bayesian models. We shall also display the efficient frontiers obtained

<sup>23</sup> Michaud, R, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?", *Financial Analysts Journal*, January-February 1989, pp. 31-42.

utilising each of these Bayesian models and we shall compare these efficient frontiers to the efficient frontier obtained when one simply uses the means and variances of a sample of past returns as estimates of the security's expected return and variance.

We employ two methods in obtaining the optimal portfolio:

1. The optimal portfolio is the portfolio whose premium (its return above the risk-free rate) per unit of risk (portfolio variance) is the maximum. Graphically it is the portfolio where the capital market line (CML) is a tangent to the efficient frontier.
2. The optimal portfolio is the portfolio with the maximum utility.

$$Utility = R_p - \lambda \sigma_p^2$$

where

$R_p$  = the portfolio's expected return

$\sigma_p^2$  = the portfolios expected variance

$$\lambda = \frac{1}{Risk\ Tolerance}$$

The portfolio manager needs to stipulate his risk tolerance ( $1/\lambda$ ) i.e. how much risk he is willing to take on for a unit gain in return. In our empirical research we will assume that the portfolio manager's risk tolerance is 0.1 ( $\lambda = 10$ ).

We have assumed that the opportunity set of assets consists of the 7-second tier JSE indices. The sample of past returns used in the estimates and in the proposed models comprises of the monthly returns of these 7 indices over the period January 1992 to December 1996.

## 4.2 Bayesian Estimates using a 90% Weight on the Portfolio Manager's Subjective Expectation

We have established that the Bayesian estimate is a weighting of the sample mean and of the portfolio manager's own subjective expectation. These subjective expectations can well be obtained from models. The models we shall be considering are the Overall Mean, the Minimum Variance and the CAPM models (as discussed in Section 4.1). In our first analysis we ourselves are going to stipulate these weights. We have stipulated a weight of 90% for the portfolio manager's subjective expectations, which are determined by the proposed models. We decided upon 90% as this ensures that the proposed model's "subjective" expectations become significant, otherwise one would simply achieve the optimal portfolios that are no different to the one attained when using the traditional sample means as the expectations for the securities' expected returns.

Table 3 and Table 4 list the details of the optimal portfolio obtained for each of the 3 Bayesian models when using the CML tangency optimisation technique, together with the optimal portfolio attained when one simply uses the "traditional" sample means as estimates for the securities' expected returns.

**Table 3** Details of the optimal portfolios when using the CML optimisation technique and a 90% weight on the portfolio manager's subjective expectation. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Bayesian Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|--------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                    |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Overall Mean       | 1.26%  | 4.22% | 16%                               | 0%   | 0%   | 0%   | 1%   | 74%  | 9%   |
| Minimum-Variance   | 1.39%  | 4.11% | 14%                               | 0%   | 1%   | 0%   | 0%   | 61%  | 24%  |
| CAPM               | 1.28%  | 4.37% | 9%                                | 0%   | 0%   | 0%   | 28%  | 63%  | 0%   |
| Traditional        | 1.84%  | 4.31% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |

Table 4 contains the components of the Bayesian estimates used for each asset in the opportunity set (i.e. for each second-tier index). Each Bayesian estimate  $\hat{\mu}$  is a weighting of the two components  $\mu_0$  and  $\bar{r}$ . In Table 4  $\hat{\mu} = 90\% \times \mu_0 + 10\% \times \bar{r}$ .

**Table 4 Components of the Bayesian Estimates. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).**

| Index | Bayesian Technique |         |           |                  |         |           |             |         |           | Traditional |           |
|-------|--------------------|---------|-----------|------------------|---------|-----------|-------------|---------|-----------|-------------|-----------|
|       | Overall Mean       |         |           | Minimum-Variance |         |           | CAPM        |         |           |             |           |
|       | $\hat{\mu}$        | $\mu_0$ | $\bar{r}$ | $\hat{\mu}$      | $\mu_0$ | $\bar{r}$ | $\hat{\mu}$ | $\mu_0$ | $\bar{r}$ | $\hat{\mu}$ | $\bar{r}$ |
| Coal  | 1.27%              | 1.21%   | 1.82%     | 1.40%            | 1.36%   | 1.82%     | 1.25%       | 1.19%   | 1.82%     | 1.82%       | 1.82%     |
| Diam  | 1.18%              | 1.21%   | 0.96%     | 1.32%            | 1.36%   | 0.96%     | 1.30%       | 1.34%   | 0.96%     | 0.96%       | 0.96%     |
| Gold  | 1.16%              | 1.21%   | 0.73%     | 1.29%            | 1.36%   | 0.73%     | 1.37%       | 1.44%   | 0.73%     | 0.73%       | 0.73%     |
| Metm  | 1.15%              | 1.21%   | 0.66%     | 1.29%            | 1.36%   | 0.66%     | 1.26%       | 1.32%   | 0.66%     | 0.66%       | 0.66%     |
| Mfin  | 1.22%              | 1.21%   | 1.33%     | 1.35%            | 1.36%   | 1.33%     | 1.36%       | 1.36%   | 1.33%     | 1.33%       | 1.33%     |
| Fina  | 1.27%              | 1.21%   | 1.85%     | 1.41%            | 1.36%   | 1.85%     | 1.24%       | 1.18%   | 1.85%     | 1.85%       | 1.85%     |
| Indu  | 1.20%              | 1.21%   | 1.09%     | 1.33%            | 1.36%   | 1.09%     | 1.18%       | 1.19%   | 1.09%     | 1.09%       | 1.09%     |

Using the Overall Mean Bayesian estimate each security's expected return is a weighting of the security's own sample mean and a common overall mean (1.21%). Since a heavy weight (90%) is placed on the overall mean, all the securities' Bayesian expected returns will be drawn quite significantly towards this common value. It is the securities whose sample means differ quite significantly from the overall mean that will be most affected by the Bayesian estimate. For example, when examining Table 4 it is evident that the expected return estimate for the Industrial Index (Indu) has increased from 1.09% to 1.20%. This will certainly make the Industrial (Indu) sector far more attractive investment than before, whereas the securities whose sample means were located near the overall mean of 1.21% will not be any more or any less attractive. The Financial (Fina) index has experienced the reverse to the Industrial (Indu) index. The Financial (Fina) index's sample mean (1.85%) is quite significantly larger than the overall mean (1.21%). Consequently it is drawn down quite substantially making it a relatively less attractive investment than before. As a result, when using the Bayesian Overall Mean estimates rather than the "traditional" sample mean estimates, the optimal proportion assigned to the Industrial (Indu) sector has increased from 0% to 9% whereas the optimal proportion assigned to the Financial (Fina) sector has declined from 83% to 74% (see Table 3). It is also evident in Table 3 that the optimal portfolio achieved using the Bayesian Overall Mean estimates is more diversified than one reached when using the "traditional" sample mean estimates, which will certainly make it a far more attractive to many portfolio managers.

Using the Minimum Variance Bayesian estimate each security's expected return is a weighting of the security's sample mean and the common mean value (1.36%) of the minimum variance portfolio. It is a similar estimate to the Bayesian Overall Mean estimate where each security's expected return is drawn towards a common value. Hence the argument concerning which securities are relatively more attractive when using the Bayesian estimates rather than the traditional sample mean estimates is the same as the one described in the paragraph above. As with the optimal portfolio attained when using the Bayesian Overall Mean estimates, the optimal portfolio accomplished when using the Bayesian Minimum Variance Bayesian estimates is far more diversified than that achieved when using traditional sample mean estimates (see Table 3).

The Bayesian CAPM estimate is slightly different to the previous two Bayesian estimates described. In each of the previous two cases the security's sample mean was drawn to a common value, whereas the Bayesian CAPM estimate is a compromise between the equivalent CAPM estimate and the "traditional" sample mean estimate. Table 5 below lists the details of the optimal portfolio achieved when using "traditional" sample mean estimates, CAPM estimates and Bayesian CAPM estimates.

**Table 5** Details of the optimal portfolio when using CAPM and "traditional" estimates in conjunction with the CML tangency optimisation technique. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Estimation Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|----------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                      |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Traditional          | 1.84%  | 4.31% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| CAPM                 | 1.24%  | 4.32% | 2%                                | 5%   | 6%   | 3%   | 17%  | 18%  | 49%  |
| Bayesian CAPM        | 1.28%  | 4.37% | 9%                                | 0%   | 0%   | 0%   | 26%  | 63%  | 0%   |

As expected the CAPM optimal portfolio is the market portfolio and it is certainly more diversified than the "traditional" optimal portfolio. The Bayesian CAPM optimal portfolio is a compromise between the other two optimal portfolios. For example, the Coal sector has a 9% weight in the Bayesian CAPM optimal portfolio, which is a compromise between the weight it has in the "traditional" optimal portfolio of 17%,

and the weight, it has in the CAPM optimal portfolio of 2%. The Bayesian CAPM optimal portfolio has evidence of being more diversified than the "traditional" optimal portfolio with funds allocated to the Mining and Financial (MFin) sector, unlike when using "traditional" estimates. The Bayesian CAPM optimal portfolio will continue to diversify and more closely resemble the CAPM optimal portfolio as the weight on its CAPM component increases. It will be identical to the CAPM optimal portfolio when its CAPM component has a weight of 100%.

We have proposed using the sample means as the core estimates of the securities' expected returns, and when one has subjective expectations, as captured by the proposed models (e.g. the CAPM), in the securities' expected returns, these core estimates are updated. A similar approach to this would be to rather use the CAPM estimates as the core or equilibrium estimates of the securities' expected returns, and then letting the portfolio manager specify his/her own views, if he/she has any, by stating the return he/she expects the security to achieve and secondly by specifying the confidence he/she has in each of these expectations. The consequent Bayesian estimate would be a weighting of the CAPM estimate and of the portfolio manager's own subjective expectation. The more certain the portfolio manager is in his/her subjective expectation the greater the weight placed on this subjective expectation. If the portfolio manager has no subjective expectations then the optimal portfolio will be the CAPM optimal portfolio. But, for example, if the portfolio manager was *very confident* that a specific asset was going to do far better than what the CAPM estimate predicts then this asset will certainly be more prevalent in the Bayesian optimal portfolio than what it is in the CAPM market portfolio.

We shan't examine this empirically since one needs to specify the portfolio manager's subjective expectations and his/her confidence in each of these expectations, which is somewhat arbitrary.

Figure 1 is a plot of the efficient frontiers obtained when using each of the proposed Bayesian estimator techniques together with the efficient frontier obtained when simply using sample means as estimates of the securities' expected returns. The

optimal portfolio is the portfolio where the CML is a tangent to the efficient frontier. Each circle in Figure 1 indicates the location of the optimal portfolio for each proposed estimator.

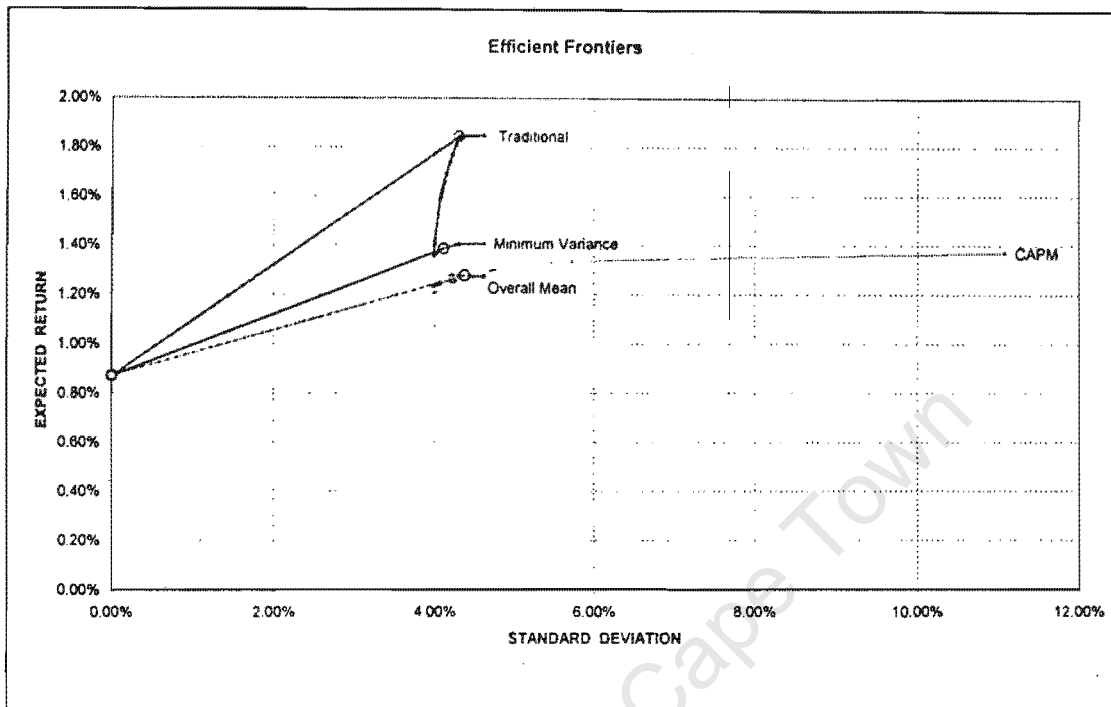


**Figure 1** Efficient Frontier plots of the optimal portfolios when using 90% weights together with the CML optimisation technique. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

From Figure 1 and Table 3 it is evident that the Bayesian CAPM optimal portfolio is the one with the greatest risk (standard deviation), whereas it is the Bayesian Minimum Variance optimal portfolio that has the lowest risk (standard deviation). We are simply using Figure 1 to display the levels of risk (standard deviation) for the respective optimal portfolios. In Figure 1 the "Traditional" efficient frontier dominates the other two efficient frontiers. From this one cannot conclude that the "traditional" efficient frontier is more "efficient" than the other two efficient frontiers because for each frontier the vertical axis (Expected Return) is actually a different measure. This "arbitrary" issue is discussed more extensively in Appendix 5.

We shall now consider the utility function optimisation technique. Table 6 lists the details of the optimal portfolio for each of the three Bayesian techniques, together with the "traditional" technique, when the risk tolerance is set at 0.1 (i.e.  $\lambda = 10$ ).

**Table 6** Details of the optimal portfolio when using a 90% weighting on the portfolio manager's subjective expectations together with the utility function optimisation technique. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Bayesian Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|--------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                    |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Overall Mean       | 1.23%  | 3.99% | 10%                               | 0%   | 3%   | 0%   | 0%   | 32%  | 55%  |
| Minimum-Variance   | 1.36%  | 3.99% | 10%                               | 0%   | 3%   | 0%   | 0%   | 32%  | 55%  |
| CAPM               | 1.22%  | 4.00% | 9%                                | 0%   | 4%   | 0%   | 0%   | 36%  | 51%  |
| Traditional        | 1.81%  | 4.27% | 17%                               | 0%   | 0%   | 0%   | 2%   | 78%  | 3%   |

Bayesian estimates certainly do produce different optimal portfolios to the one obtained when using "traditional" sample means as estimates for the securities' expected returns. This would be expected when one compares the slopes of the efficient frontiers with one another. The "traditional" efficient frontier has a significantly different shape to the Bayesian efficient frontiers, which are all very similar to one another (see Figure 1). The CAPM Bayesian optimal portfolio differs very marginally from the other two Bayesian optimal portfolios. If we displayed more decimal places it would be apparent that all the Bayesian optimal portfolios do in fact differ (insignificantly). It is interesting to note when using Bayesian estimates the portfolio manager would choose a more conservative portfolio i.e. the standard deviations of the optimal Bayesian portfolios are all close to 4% which is lower than the standard deviation of the "traditional" technique's optimal portfolio whose standard deviation is 4.27%.

So far we have used models to estimate the portfolio manager's subjective expectations in the returns of the securities. However, we have assumed that the portfolio manager would have the same certainty in each security's expected return. But if the portfolio manager is using models to determine his subjective expectations he could just as easily use these models to determine the certainty in each of these

expectations. This is achieved by using the standard errors of the estimates as indicators of these certainties.

### 4.3 Bayesian Estimates using the Standard Errors of the Estimates in Calculating the Weights placed on the Portfolio Manager's Subjective Expectations

Table 7 to Table 10 list the details of the optimal portfolios for each of the three Bayesian estimates when using the Capital Market Line (CML) tangency optimisation technique. The portfolio manager's personal expectations were estimated by each of the three proposed models. The Bayesian estimate in a security's expected return is a weighting of the proposed model's expectation and the sample mean. The weight placed on the model's expectation is determined by the standard error of the estimate. The tables also list the same details for the optimal portfolio achieved when using the "traditional" sample means as estimates for the securities' expected returns.

**Table 7 Details of the optimal portfolio when using standard error weightings in conjunction with the CML optimisation technique. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).**

| Bayesian Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|--------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                    |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Overall Mean       | 1.65%  | 4.34% | 12%                               | 0%   | 0%   | 0%   | 0%   | 88%  | 0%   |
| Minimum-Variance   | 1.55%  | 4.32% | 14%                               | 0%   | 0%   | 0%   | 0%   | 86%  | 0%   |
| CAPM               | 1.39%  | 4.30% | 20%                               | 0%   | 0%   | 0%   | 5%   | 75%  | 0%   |
| Traditional        | 1.84%  | 4.31% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |

Table 8 to Table 10 more specifically list the components of the Bayesian estimates used for each asset in the opportunity set. The column labelled "Subjective Estimate Weight" is the quantity  $\frac{\sigma_1^2}{\sigma_0^2}$  and the column labelled "Sample Estimate Weight" is the

quantity  $\sigma_1^2 \frac{n}{\sigma^2}$ . These weights were discussed in Section 4.1.

**Table 8 Components of the Overall Mean Bayesian Estimate. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).**

| Index | $\hat{\mu}$ | Subjective Estimate Weight | $\mu_0$ | Sample Estimate Weight | $\bar{r}$ |
|-------|-------------|----------------------------|---------|------------------------|-----------|
| Coal  | 1.47%       | 58%                        | 1.21%   | 42%                    | 1.82%     |
| Diam  | 1.09%       | 52%                        | 1.21%   | 48%                    | 0.96%     |
| Gold  | 1.06%       | 68%                        | 1.21%   | 32%                    | 0.73%     |
| Metm  | 0.94%       | 51%                        | 1.21%   | 49%                    | 0.66%     |
| Mfin  | 1.28%       | 44%                        | 1.21%   | 56%                    | 1.33%     |
| Fina  | 1.67%       | 27%                        | 1.21%   | 73%                    | 1.85%     |
| Indu  | 1.12%       | 23%                        | 1.21%   | 77%                    | 1.09%     |

**Table 9 Components of the Minimum Variance Bayesian Estimate. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).**

| Index | $\hat{\mu}$ | Subjective Estimate Weight | $\mu_0$ | Sample Estimate Weight | $\bar{r}$ |
|-------|-------------|----------------------------|---------|------------------------|-----------|
| Coal  | 1.44%       | 83%                        | 1.36%   | 17%                    | 1.82%     |
| Diam  | 1.27%       | 79%                        | 1.36%   | 21%                    | 0.96%     |
| Gold  | 1.28%       | 88%                        | 1.36%   | 12%                    | 0.73%     |
| Metm  | 1.21%       | 79%                        | 1.36%   | 21%                    | 0.66%     |
| Mfin  | 1.35%       | 74%                        | 1.36%   | 27%                    | 1.33%     |
| Fina  | 1.57%       | 57%                        | 1.36%   | 43%                    | 1.85%     |
| Indu  | 1.23%       | 52%                        | 1.36%   | 48%                    | 1.09%     |

**Table 10 Components of the CAPM Bayesian Estimate. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).**

| Index | $\hat{\mu}$ | Subjective Estimate Weight | $\mu_0$ | Sample Estimate Weight | $\bar{r}$ |
|-------|-------------|----------------------------|---------|------------------------|-----------|
| Coal  | 1.48%       | 54%                        | 1.19%   | 46%                    | 1.82%     |
| Diam  | 1.21%       | 66%                        | 1.34%   | 34%                    | 0.96%     |
| Gold  | 1.16%       | 61%                        | 1.44%   | 39%                    | 0.73%     |
| Metm  | 1.09%       | 65%                        | 1.32%   | 35%                    | 0.66%     |
| Mfin  | 1.36%       | 79%                        | 1.36%   | 21%                    | 1.33%     |
| Fina  | 1.37%       | 71%                        | 1.18%   | 29%                    | 1.85%     |
| Indu  | 1.17%       | 83%                        | 1.19%   | 17%                    | 1.09%     |

Using the Overall Mean Bayesian estimate each security's expected return is a weighting of the security's own sample mean and the common overall mean (1.21%).

The smaller the sample variance the greater the weight placed on the sample mean. Previously when simply using sample means as estimates of the securities' expected returns the optimal portfolio consisted of 17% of the funds invested in the Coal sector and 83% of the funds placed in the Financial (Fina) sector. Now using the Overall Mean Bayesian estimate the proportion of funds allocated to the Coal sector has declined to 12% and the proportion of funds allocated to the Financial (Fina) sector has increased to 88% (see Table 7). Each of these two indices has a sample mean which is greater than the common overall mean, hence when using the Bayesian estimate each of these two estimates is drawn *down* towards the common overall mean. However, the sample variance of the Financial (Fina) index is significantly smaller than the sample variance of the Coal Index and consequently the Financial (Fina) index has a greater weight (73%) placed on its sample mean compared to the weight (42%) placed on Coal index's sample mean. (see Table 8). Hence the Financial (Fina) sector now has a *relatively* better expected return than the Coal sector when using the Overall Mean Bayesian estimate, and since the estimate of the two securities' variances have remained unchanged there is an increase in the proportion of funds allocated to the Financial (Fina) sector relative to the Coal sector.

Using the Minimum Variance Bayesian estimate each security's expected return is a weighting of the security's own sample mean and the common mean value (1.36%) of the minimum variance portfolio. This is a similar estimate to the Bayesian Overall Mean estimate and consequently the outcomes are not surprisingly much different. Using the Bayesian Minimum Variance estimate the proportion of funds allocated to the Coal sector has declined from 17% (when using the "traditional" sample means as estimates of the securities' expected returns) to 14% and the proportion of funds allocated to the Financial (Fina) sector has increased from 83% to 86% (see Table 7). The argument for the reallocation of funds is along the same line as the argument outlined above with respect to the Bayesian Overall Mean estimate. A significant point which contributes to a slight difference in the optimal portfolio achieved when using Bayesian Minimum Variance estimates is that the Minimum Variance portfolio has a smaller variance (and standard error) than that of the Overall Mean estimate. Consequently a larger weight is placed on the return of the Minimum Variance

portfolio component of the Bayesian estimate than what is placed on the equivalent Overall Mean component of the proposed Bayesian estimate.

When applying the Bayesian CAPM estimates for the securities' expected returns each security's expected return is a weighting of its own sample mean and its CAPM expected return estimate. Unlike the previous two Bayesian estimates the CAPM Bayesian estimate is not drawn towards some *common* expected return (e.g. the overall mean). Consequently the optimal portfolio achieved using the Bayesian CAPM estimates differs from the other Bayesian optimal portfolios (see Table 7). With the other two Bayesian estimates it was the securities' sample variances that were determining the relative weights placed on each security's sample mean because the overall mean's variance (the minimum variance portfolio's variance in the Minimum Variance Bayesian estimate case) was common for all the securities. Now with the Bayesian CAPM estimates each security's sample variance and CAPM variance (the Mean Square Error) will differ from one another.

When simply using the sample means as estimates of the securities' expected returns the optimal portfolio comprises of 17% of the funds allocated to the Coal sector and 83% of the funds allocated to the Financial (Fina) sector. The previous two Bayesian estimates examined saw an increase in the proportion of funds allocated to the Financial (Fina) sector and a decrease in the proportion allocated to the Coal sector (see Table 7). Using the Bayesian CAPM estimates we actually observe the reverse. The Bayesian CAPM portfolio comprises of 75% Financial (Fina), 20% Coal and 5% Mining Financial (Mfin). The reason is as follows:

When simply using sample means as estimates for the securities' returns the Financial (Fina) index outperformed the Coal index by 0.03%, but according to the CAPM estimate, if the overall expected market return is 1.24% (which was the monthly average over the past 5 years) then the Coal index is actually expected to outperform the Financial (Fina) index by 0.01%. After applying the weights in accordance to the Bayesian CAPM estimate the Coal index is actually expected to outperform the Financial (Fina) index by 0.11% (i.e. Coal's Bayesian CAPM expected return is

1.48% and Financial's (Fina's) Bayesian expected return is 1.37% (see Table 10)). One must still be mindful that the Coal index has a greater variance than the Financial (Fina) index. Consequently the proportion of funds allocated to the Coal sector increases relative to the Financial (Fina) sector, but it does not totally replace the proportion assigned to the Financial sector because the Financial index still has a smaller variance than the Coal index.

The Bayesian CAPM estimate of a security's expected return is a compromise between the equivalent CAPM estimate and the "traditional" sample mean estimate. Consequently we may expect the Bayesian CAPM optimal portfolio to be a compromise between the CAPM optimal portfolio and the "traditional" optimal portfolio. By examining Table 10 we see that the CAPM estimate always receives slightly heavier weighting (between 54% for the Coal index and 83% for the Industrial (Indu) index) and from this we might well expect the Bayesian CAPM optimal portfolio to more closely resemble the CAPM optimal portfolio than the "traditional" optimal portfolio. Table 11 below lists the details of the optimal portfolios obtained when using each of these three respective expected return estimates.

**Table 11** Details of the optimal portfolio when using CAPM and "traditional" estimates in conjunction with the CML tangency optimisation technique. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Estimation Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|----------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                      |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Traditional          | 1.84%  | 4.31% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| CAPM                 | 1.24%  | 4.32% | 2%                                | 5%   | 6%   | 3%   | 17%  | 18%  | 49%  |
| Bayesian CAPM        | 1.39%  | 4.30% | 20%                               | 0%   | 0%   | 0%   | 5%   | 75%  | 0%   |

As expected the CAPM optimal portfolio is simply the market portfolio and is certainly far more diversified than the "traditional" optimal portfolio. The Bayesian CAPM optimal portfolio is also more diversified than the "traditional" optimal portfolio, which will certainly make it more attractive to many portfolio managers. The Bayesian CAPM optimal portfolio has a closer resemblance to the "traditional"

optimal portfolio than to the CAPM optimal portfolio even though the CAPM estimate receives heavier weighting than the sample mean estimate in the Bayesian CAPM estimate.

The Bayesian CAPM optimal portfolio is more like the "traditional" optimal portfolio than the CAPM optimal portfolio. The reason is because the *relative* magnitudes of the expected returns obtained from the Bayesian CAPM estimates are more similar to the *relative* magnitudes of the expected returns obtained from the "traditional" sample mean estimates than they are to the *relative* expected returns obtained from the CAPM estimates. Table 12 below examines the relative magnitudes of the estimates. This is achieved by ranking the magnitudes of the securities' expected returns and examining whether these ranks differ when using the three different estimates.

**Table 12** Rankings of the magnitudes of the securities' expected returns. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Index | Bayesian CAPM<br>Estimate | Rank | CAPM<br>Estimate | Rank | Traditional<br>Estimate | Rank |
|-------|---------------------------|------|------------------|------|-------------------------|------|
| Coal  | 1.48%                     | 1    | 1.19%            | 5.5  | 1.82%                   | 2    |
| Diam  | 1.21%                     | 4    | 1.34%            | 3    | 0.96%                   | 5    |
| Gold  | 1.16%                     | 6    | 1.44%            | 1    | 0.73%                   | 6    |
| Metm  | 1.09%                     | 7    | 1.32%            | 4    | 0.66%                   | 7    |
| Mfin  | 1.36%                     | 3    | 1.36%            | 2    | 1.33%                   | 3    |
| Fina  | 1.37%                     | 2    | 1.18%            | 7    | 1.85%                   | 1    |
| Indu  | 1.17%                     | 5    | 1.19%            | 5.5  | 1.09%                   | 4    |

It is evident that the rankings of the Bayesian CAPM estimates are closer to the rankings achieved using the "traditional" sample mean estimates than what they are to the rankings achieved using CAPM estimates. For example, Gold was ranked 6<sup>th</sup> when using either the Bayesian CAPM estimate or the "traditional" sample mean estimate but was ranked 1<sup>st</sup> when using the CAPM estimate. The rankings of the Bayesian CAPM estimates are significantly correlated with the rankings of the "traditional" sample mean estimates (Spearman's Rank Order Correlation = 0.929,  $p < .003$ ) whereas the rankings of the Bayesian CAPM estimates are not significantly correlated with the rankings of the CAPM estimates (Spearman's Rank Order Correlation = -0.468,  $p < .289$ ). This suggests that the Bayesian CAPM optimal

portfolio has a closer resemblance to the "traditional" optimal portfolio than it has to the CAPM optimal portfolio.

We shall now consider these same Bayesian estimates using the utility function optimisation technique instead. Table 13 lists the details of the optimal portfolios for each of the three Bayesian techniques together with the details of the optimal portfolio achieved if one used the "traditional" sample means as estimates of the securities' expected returns. The risk tolerance was set at 0.1 ( $\lambda = 10$ ). Table 14 is structured in a similar manner to Table 13 except in this scenario the risk tolerance is set higher at 4 ( $\lambda = 0.25$ ).

**Table 13** Details of the optimal portfolio when using Bayesian and "traditional" estimates in conjunction with the utility optimisation technique, when the risk tolerance is set at 0.1 i.e.  $\lambda = 10$ . The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Bayesian Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|--------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                    |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Overall Mean       | 1.52%  | 4.13% | 13%                               | 0%   | 2%   | 0%   | 0%   | 64%  | 21%  |
| Minimum-Variance   | 1.42%  | 4.04% | 11%                               | 0%   | 3%   | 1%   | 0%   | 50%  | 36%  |
| CAPM               | 1.29%  | 4.01% | 12%                               | 0%   | 2%   | 0%   | 0%   | 41%  | 45%  |
| Traditional        | 1.81%  | 4.27% | 17%                               | 0%   | 0%   | 0%   | 2%   | 78%  | 3%   |

**Table 14** Details of the optimal portfolio when using Bayesian and "traditional" estimates in conjunction with the utility optimisation technique, when the risk tolerance is set at 4 i.e.  $\lambda = 0.25$ . The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Bayesian Technique | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|--------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                    |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Overall Mean       | 1.67%  | 4.62% | 0%                                | 0%   | 0%   | 0%   | 0%   | 100% | 0%   |
| Minimum-Variance   | 1.57%  | 4.62% | 0%                                | 0%   | 0%   | 0%   | 0%   | 100% | 0%   |
| CAPM               | 1.42%  | 4.87% | 42%                               | 0%   | 0%   | 0%   | 0%   | 58%  | 0%   |
| Traditional        | 1.85%  | 4.62% | 12%                               | 0%   | 0%   | 0%   | 0%   | 88%  | 0%   |

Comparing Table 13 to Table 14 the following points are evident.

1. Considering the "traditional" estimates, which are simply sample means, as the risk tolerance increases so more of the funds are allocated to the Financial (Fina)

index. The reason being the Financial (Fina) index has the largest sample mean and hence when risk (variance) becomes negligent more of the funds are allocated to this asset, since the objective is then simply to maximise the portfolio's expected return. The same occurs when using Bayesian Overall Mean, and Bayesian Minimum Variance estimates.

2. However, when Bayesian CAPM estimates are used, as the risk tolerance increases more and more of the funds are allocated to the Coal index. In fact when the risk tolerance becomes infinite the optimal solution will be 100% of the funds allocated to the Coal index rather than to the Financial (Fina) index. The simple reason is the Coal index, not Financial (Fina) index has the highest Bayesian CAPM expected return.
3. The portfolios are certainly more diversified at lower levels of risk tolerance, and certainly more diversified than the optimal portfolio achieved when using the CML tangency optimisation technique.
4. The CAPM optimal portfolio has the lowest risk (its standard deviation is 4.01%) at the lower of the two levels of risk tolerance, but at the higher of the two levels of risk tolerance it has the highest risk (its standard deviation = 4.87%).

The portfolio manager's risk tolerance is a big determinant in the optimal portfolio.

All the empirical research has assumed that the portfolio manager has subjective expectations in the securities' returns. Practically they can be obtained from a model or simply stated by the portfolio manager. We shall now consider the scenario where the portfolio manager has no definite personal views in the securities' expected returns.

#### **4.4 The Effects of the "Biased" Bayesian Variance Estimate on the Efficient Frontier and Optimal Portfolios**

We now consider the situation that was described as scenario 3 in Table 1 and in Table 2 respectively. In this situation it is assumed that the portfolio manager has no subjective views on the securities' expected returns or on the securities' variances. It would be natural to assume that the estimates and efficient frontier produced using either Bayesian estimates or simple sample estimates would be the same. However, Bayesian estimates are biased relative to simple sample estimates. Examining Table 2 the Bayesian estimate of a security's variance has a "bias" factor of  $\frac{n-1}{n-3}$  i.e.

$$\hat{\sigma}^2 = \frac{n-1}{n-3} s^2.$$

Similarly it can be shown that the Bayesian estimate of the covariance between any two securities has the same "bias" factor of  $\frac{n-1}{n-3}$  (Jobson and Korkie<sup>24</sup>, 1980). A consequence of the Bayesian estimate is that it will have no effect on the composition of the efficient frontier when using the Capital Market Line (CML) tangency optimisation technique, but it will have an influence on the location of the efficient frontier. This is certainly evident when we examine the Bayesian estimate in the context of the Markowitz model, which is what we shall do below.

The following notation will be used for this discussion:

1.  $E_p$  = Expected Return on the portfolio
2.  $\sigma_p^2$  = Variance of the portfolio
3.  $X_i$  = Proportion of funds invested in security i
4.  $N$  = Total Number of securities considered
5.  $V$  = Variance-Covariance Matrix of the  $N$  securities
6.  $1/\lambda$  = Risk Tolerance

Sharpe (1970) considered the following objective function in order to find the set of efficient portfolios:

$$\text{Minimise } -\lambda E_p + \sigma_p^2 \quad \text{for all } \lambda \geq 0$$

Subject to the constraints:

1.  $\sum X_i = 1$                       that is the entire portfolio must be invested
2.  $X_i \geq 0$                       for all  $i = 1, 2, \dots, N$ . That is no security may be held in negative quantities

$\sigma_p^2 = X' V X$ . Consequently, when considering the objective function, one wishes to minimise  $X' V X$ , which is a measure of the portfolio's variance. The variance-covariance matrix can be estimated using a sample covariance estimate, denoted  $\hat{V}$ , or a Bayesian estimate, denoted  $\hat{\hat{V}}$ . Note that  $\hat{\hat{V}}$  is effectively a "biased" estimate of  $\hat{V}$  since

$$\hat{\hat{V}} = \frac{n-1}{n-3} \hat{V} \quad (\text{Jobson and Korkie, 1980}) \quad (20)$$

Our objective is to minimise the portfolio's variance at each level of expected return. If we use sample estimates for the covariance matrix our objective is to minimise  $X' \hat{V} X$ . Similarly, when using Bayesian estimates, our objective is to minimise  $X' \hat{\hat{V}} X$  or equivalently by replacing  $\hat{\hat{V}}$  with  $\frac{n-1}{n-3} \hat{V}$  our objective is to minimise

$$\frac{n-1}{n-3} X' \hat{V} X.$$

<sup>24</sup> Jobson, JP and Korkie, B, "Estimation for Markowitz Efficient Portfolios", *Journal of the American Statistical Association*, September 1980.

Minimising  $X' \hat{V} X$  is equivalent to minimising  $\frac{n-1}{n-3} X' \hat{V} X$  since the two values differ only by the constant proportional quantity of  $\frac{n-1}{n-3}$  and the solution  $X^*$  will be the same under either formulation.

Although the composition of the efficient frontier is unaffected by the Bayesian estimate its location certainly is affected and the location of the efficient frontier ultimately impacts on the optimal portfolio. This is clear when we compare the Bayesian variance estimate to the sample variance estimate:

$$\text{Let } S_{Bayes}^2 = \hat{V} \text{ and} \\ S_{Sample}^2 = \hat{V}$$

It has been established that

$$S_{Bayes}^2 = \frac{n-1}{n-3} S_{Sample}^2$$

As the sample size of returns increases so  $\frac{n-1}{n-3} \rightarrow 1$ , and consequently the sample “bias” is reduced. Effectively, the larger the sample sizes the smaller the estimation risk. However, one must be wary of merely increasing the sample size in order to reduce the estimation risk because there are also other risks involved. There is the risk that the distribution of  $\mu$ , the security's expected return, is non-stationary i.e. it changes over time. Effectively, increasing the sample size, that is going further back in time, will reduce the estimation risk, but simultaneously increase the risk of obtaining a non-stationary  $\mu$ .

The difference between the Bayesian and sample measures of risk (standard deviation) is:

$$\begin{aligned}
 \sqrt{S_{Bayesian}^2} - \sqrt{S_{Sample}^2} &= \sqrt{\frac{n-1}{n-3} S_{Sample}^2} - \sqrt{S_{Sample}^2} \\
 &= \sqrt{\frac{n-1}{n-3}} S_{Sample} - S_{Sample} \\
 &= S_{Sample} \left( \sqrt{\frac{n-1}{n-3}} - 1 \right)
 \end{aligned}$$

Thus in mean-standard deviation analysis, the Bayesian efficient frontier and the efficient frontier formed using sample estimates will be a distance of:

$$S_{Sample} \left( \sqrt{\frac{n-1}{n-3}} - 1 \right) \text{ apart, as shown in Figure 2.}$$

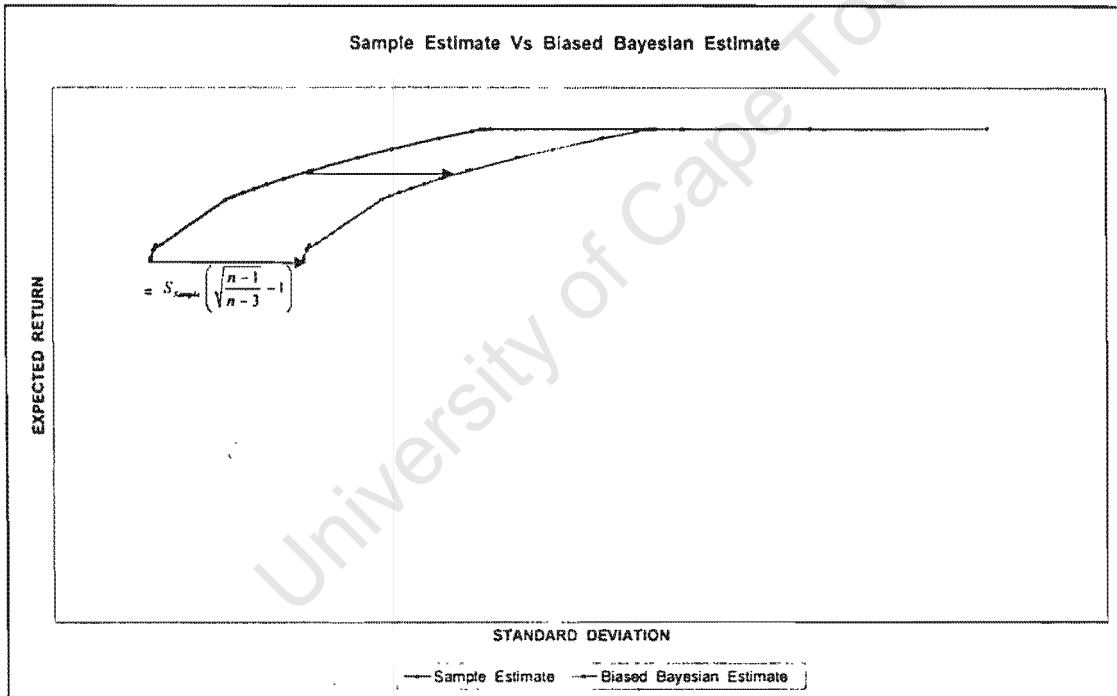


**Figure 2 The Effect of a Bayesian Estimate on the Efficient Frontier**

The Bayesian efficient frontier will shift horizontally away from the classical efficient frontier, which was created using sample estimates, into a more risky position. The size of this horizontal shift is a function of the sample's size. The larger the sample size, the closer the two efficient frontiers are from one another. The distance that the two efficient frontiers are apart from one another is also a function of the portfolios'

standard deviations. The smaller the portfolio's standard deviation, the closer the two efficient frontiers are from one another.

Although the location of the efficient frontier does indeed change, the optimal portfolio itself may or may not change - depending upon the optimisation technique used. We examine this issue using the monthly returns of the JSE second-tier indices over the period January 1992 - December 1996.

When applying the Capital Market Line (CML) tangency optimisation technique the optimal portfolio achieved using the sample variances as estimates of the securities' variances will be the same as the optimal portfolio achieved when we use the sample variances multiplied by the bias factor of  $\frac{n-1}{n-3}$  as estimates of the securities' variances. The details of these two optimal portfolios are listed in Table 15. All that differs is the standard deviation of these two optimal portfolios.

**Table 15** The optimal portfolios achieved when using sample variances as estimates of the securities' variances and those achieved when using the "biased" Bayesian sample variances as estimates. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Variance Estimate      | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|------------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                        |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Sample Variance        | 1.84%  | 4.31% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |
| Biased Sample Variance | 1.84%  | 4.38% | 17%                               | 0%   | 0%   | 0%   | 0%   | 83%  | 0%   |

Jobson and Korkie (1980) proved mathematically that the optimal portfolio reached using either of the two variance estimates would be identical when applying the CML tangency optimisation technique. The locations of these two optimal portfolios are displayed in Figure 3 (i.e. where the respective Capital Market Lines (CML's) are tangents to the respective efficient frontiers).

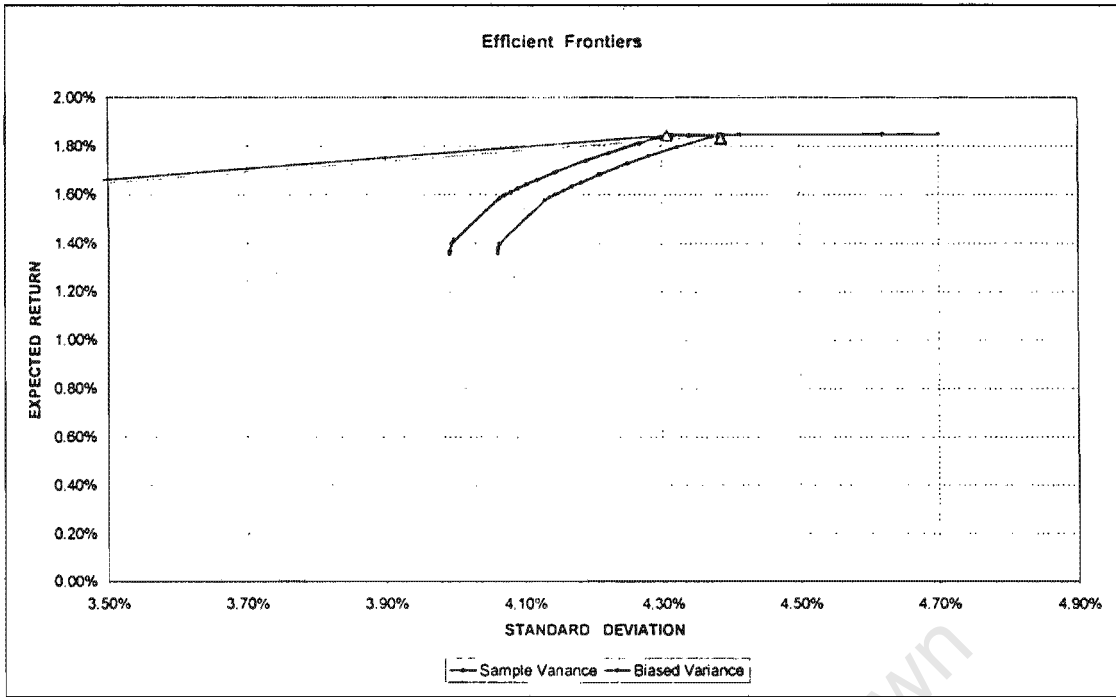


**Figure 3** The Efficient Frontier created when using sample variances as estimates of the securities' variances together with the Efficient Frontier created when using the "biased" Bayesian sample variances as estimates of the securities' variances. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

However Jobson and Korkie (1980) also indicated that when the utility function optimisation technique is used then the two optimal portfolios might well differ. This is evident in our example. Considering the utility function where the risk tolerance is 0.1 ( $\lambda = 0.1$ ) the two optimal portfolios would be as follows:

**Table 16** The optimal portfolios when using Utility Functions with a risk tolerance of 0.1. The opportunity set consists of the 7-second tier JSE indices (January 1992-December 1996).

| Variance Estimate      | Return | STD   | Proportion Invested in each Index |      |      |      |      |      |      |
|------------------------|--------|-------|-----------------------------------|------|------|------|------|------|------|
|                        |        |       | Coal                              | Diam | Gold | Metm | MFin | Fina | Indu |
| Sample Variance        | 1.81%  | 4.27% | 17%                               | 0%   | 0%   | 0%   | 2%   | 78%  | 3%   |
| Biased Sample Variance | 1.80%  | 4.32% | 16%                               | 0%   | 0%   | 0%   | 2%   | 77%  | 5%   |

The two optimal portfolios differ very *marginally*. The reason is because the bias factor of  $\frac{n-1}{n-3} = \frac{60-1}{60-3} = 1.04$  is almost insignificant. The optimal portfolios might well differ more significantly if we were to use fewer past observations.

#### 4.5 Practical Insights for Practitioners

- The Bayesian estimate of the covariance between any two securities has a "bias" coefficient of  $\frac{n-1}{n-3}$ . This "bias" coefficient influences the *location* of the efficient frontier and it may also, but not necessarily, impact upon the optimal portfolio.
- If the Capital Market Line (CML) tangency optimisation technique is applied then the optimal portfolio obtained using the "biased" sample variances and co-variances is *identical* to the one obtained when the equivalent unbiased sample variances and co-variances are used as measures of risk.
- When the utility function optimisation technique is used then the optimal portfolio obtained using the "biased" sample variances and co-variances as measures of risk differs in composition, although only marginally, to the optimal portfolio obtained when using the equivalent unbiased sample variances and co-variances as measures of risk.

## 6. APPENDIX

### 6.1 Appendix 1

In Section 3.2 we showed that the Bayesian Posterior mean,  $\mu_1$ , which is a measure of a security's expected return conditional on the sample of past returns to be the following:

$$\mu_1 = \frac{\sigma_1^2}{\sigma_0^2} \mu_0 + \sigma_1^2 \frac{n}{\sigma^2} \bar{r}$$

Effectively the Bayesian Posterior mean,  $\mu_1$ , is a weighted average of the portfolio manager's subjective expectation,  $\mu_0$ , and the sample mean,  $\bar{r}$ . We wish to show mathematically that the sum of these two weights  $\frac{\sigma_1^2}{\sigma_0^2}$  and  $\sigma_1^2 \frac{n}{\sigma^2}$  is 1.

**Required To Prove:**  $\frac{\sigma_1^2}{\sigma_0^2} + \sigma_1^2 \frac{n}{\sigma^2} = 1$

**PROOF:** 
$$\begin{aligned} \frac{\sigma_1^2}{\sigma_0^2} + \sigma_1^2 \frac{n}{\sigma^2} &= \frac{\sigma_1^2 \sigma^2 + \sigma_1^2 n \sigma_0^2}{\sigma_0^2 \sigma^2} \\ &= \frac{\sigma_1^2 (\sigma^2 + n \sigma_0^2)}{\sigma_0^2 \sigma^2} \end{aligned} \quad \text{(Appendix Equation 1)}$$

but in Section 3.2 Equation ( 13 ) it was established that

$$\begin{aligned} \sigma_1^{-2} &= \sigma_0^{-2} + \frac{n}{\sigma^2} \\ \Rightarrow \frac{1}{\sigma_1^2} &= \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2} \\ \Rightarrow \sigma_1^2 &= \frac{\sigma_0^2 \sigma^2}{\sigma^2 + n \sigma_0^2} \end{aligned}$$

Substitute for  $\sigma_1^2$  into (Appendix Equation 1).

$$\frac{\sigma_1^2}{\sigma_0^2} + \frac{\sigma_1^2 n}{\sigma^2} = \frac{\sigma_0^2 \sigma^2}{\sigma^2 + n \sigma_0^2} \times \frac{(\sigma^2 + n \sigma_0^2)}{\sigma_0^2 \sigma^2} \\ = 1$$

**QED**

University of Cape Town

## 6.2 Appendix 2

In Section 3.2 we showed that the Bayesian Posterior mean,  $\mu_1$ , which is a measure of a security's expected return conditional on the sample of past returns to be the following:

$$\mu_1 = \frac{\sigma_1^2}{\sigma_0^2} \mu_0 + \sigma_1^2 \frac{n}{\sigma^2} \bar{r}$$

Effectively the Bayesian Posterior mean,  $\mu_1$ , is a weighted average of the portfolio manager's subjective expectation,  $\mu_0$ , and the sample mean,  $\bar{r}$ .

We wish to show that as the portfolio manager's certainty in his/her personal expectation  $\mu_0$  increases (i.e. as  $\sigma_0^2$  decreases) so consequently more weight is placed on this personal expectation. The weight placed on the portfolio manager's personal expectation is  $\frac{\sigma_1^2}{\sigma_0^2}$ . The sum of the two weights is 1, hence we wish to show that as his/her certainty increases (as  $\sigma_0^2 \rightarrow 0$ ) so the weight placed on his/her own expectation,  $\mu_0$ , tends towards the maximum possible value of 1.

**Required To Prove:** As  $\sigma_0^2 \rightarrow 0$  so  $\frac{\sigma_1^2}{\sigma_0^2} \rightarrow 1$

**PROOF:** Consider  $\frac{\sigma_1^2}{\sigma_0^2}$

In Appendix 1 it was shown

$$\begin{aligned} \sigma_1^2 &= \frac{\sigma_0^2 \sigma^2}{\sigma^2 + n\sigma_0^2} \\ &= \sigma_0^2 \left( \frac{\sigma^2}{\sigma^2 + n\sigma_0^2} \right) \end{aligned}$$

$$\begin{aligned}\therefore \frac{\sigma_1^2}{\sigma_0^2} &= \frac{\sigma_0^2 \left( \frac{\sigma^2}{\sigma^2 + n\sigma_0^2} \right)}{\sigma_0^2} \\ &= \frac{\sigma^2}{\sigma^2 + n\sigma_0^2}\end{aligned}\quad (\text{Appendix Equation 2})$$

Considering the denominator  $\sigma^2 + n\sigma_0^2$ , as  $\sigma_0^2 \rightarrow 0$  so  $\sigma^2 + n\sigma_0^2 \rightarrow \sigma^2$

$\therefore$  considering (Appendix Equation 2) as  $\sigma_0^2 \rightarrow 0$  so  $\frac{\sigma^2}{\sigma^2 + n\sigma_0^2} \rightarrow \frac{\sigma^2}{\sigma^2} = 1$

**QED**

### 6.3 Appendix 3

In Table 1 and in Table 2 we described the scenario where the portfolio manager has *no* subjective views on the security's expected return,  $\mu$ , or on the security's uncertainty (variance),  $\sigma^2$ . An assumption we did make in this scenario was that the security's expected return,  $\mu$ , is *independent* of its variance,  $\sigma^2$ .

In Table 1 we simply *stated* the subjective distributions which we believe are adequate enough in describing  $\mu$ , the security's expected return, and  $\sigma^2$ , the security's variance. Here in Appendix 3 we shall justify why we selected these specific subjective distributions.

In Table 1 we also merely stated the Posterior "Conditional" Distributions. In this appendix we shall discuss how one obtains these posterior distributions. These posterior distributions are related to the two subjective distributions.

The final Bayesian estimates, which are simply stated in Table 2, are obtained from these posterior "conditional" distributions. We discuss how this is achieved below.

In Section 3.2 we assumed that the parameter  $\sigma^2$  was certainly known. Now we considers the circumstance where both parameters,  $\mu$  and  $\sigma^2$ , are unknown. In this situation it is more difficult to assess a joint prior distribution for  $(\mu, \sigma^2)$  which attempts to describe the uncertainty of these two parameters. Examining the situation where little or no information is known *a priori* about the parameters, "vague" or non-informative subjective distribution functions should be used. These non-informative subjective distribution functions viz.  $\pi(\mu)$  and  $\pi(\sigma^2)$ , are best seen as limiting cases in which the distribution of the unknown parameters are "smeared out" to give essentially equal weight to all possible values.

One way of suspending beliefs about the mean return  $\mu$  is to declare a subjective distribution function that is uniform over the range of this parameter's possible

values. This will lead to an improper subjective distribution function (which we simply stated in Table 1 without any explanation). For example:

$$\pi(\mu) = k \quad -\infty < \mu < \infty \quad \text{for any } k \geq 0$$

$\pi(\mu)$  is not strictly a density function, but it does envisage all real values of  $\mu$  as being equally likely. By this assumption  $\pi(\mu) \propto 1$ .

Considering  $\sigma^2$ , it is more difficult to suggest an “equally likely” improper subjective distribution function because  $\sigma^2$  is restricted to a range of values viz.  $0 < \sigma^2 < \infty$ . To unravel the restriction one might consider all values of  $\ln \sigma^2$  as being “equally likely” over an unrestricted range (which we simply stated in Table 1 without any explanation):

$$\pi(\ln \sigma^2) = k \quad -\infty < \ln \sigma^2 < \infty \quad \text{for any } k \geq 0$$

By the change of variable argument using

$$w = \ln \sigma^2 \quad (\text{Press}^{25}, 1989)$$

$$\begin{aligned} \pi_{\sigma^2}(\sigma^2) &= \pi_w(w) \left| \frac{dw}{d\sigma^2} \right| \\ &= \pi(\ln \sigma^2) \frac{1}{\sigma^2} \\ &= \frac{k}{\sigma^2} \end{aligned}$$

where  $k$  is a constant. Accordingly,  $\pi(\sigma^2) \propto \frac{1}{\sigma^2}$ .

In the Markowitz scenario we are examining the case of normal sampling with an unknown mean and an unknown variance and accordingly we could apply the following two principles:

1.  $\pi(\mu) \propto 1$  and

<sup>25</sup> Press, SJ, "Bayesian Statistics: Principles, Models and Applications", 1989, pp. 49.

$$2. \pi(\psi) \propto \frac{1}{\sigma^2}$$

A further practical assumption that we could use is that the two parameters are independent *a priori* (in the sense that any information given to one about one parameter will not affect the level of uncertainty of the other). Hence,

$$\begin{aligned} \pi(\mu, \psi) &\propto \pi(\mu) \cdot \pi(\sigma^2) && \text{for } -\infty < \mu < \infty \text{ and } 0 < \sigma^2 < \infty. \\ &\propto 1 \cdot \frac{1}{\sigma^2} \end{aligned}$$

In Section 3.1 Equation (4) we established

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \pi(\mu) \Lambda(\mu; r_1, r_2, \dots, r_n)$$

Similarly

$$\begin{aligned} \pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n) &\propto \pi(\mu, \sigma^2) \Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) \\ &\propto \pi(\mu) \pi(\sigma^2) \Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) \end{aligned}$$

Multiplying these functions together, with some re-arrangement of terms, yields the following:

$$\pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n) \propto \left[ \left( \frac{1}{\sigma^2} \right)^{\frac{(n-1)}{2}} \exp \left[ -\frac{(n-1)s^2}{2\sigma^2} \right] \right] \left[ (\sigma^2)^{-\frac{1}{2}} \exp \left[ -\frac{n}{2\sigma^2} (\mu - \bar{r})^2 \right] \right]$$

(Appendix Equation 3)

(Appendix Equation 4)

Without completing all the proportionality constants, we can observe the following:

1. Integrating  $\pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n)$  with respect to  $\mu$  will remove the second term (i.e. (Appendix Equation 4)) leaving the first term (Appendix Equation 3) behind. From this we observe that the marginal distribution for  $\sigma^2$ , the risk, is a Gamma

distribution with parameters:  $(n-1)/2$  and  $(n-1)s^2/2$  (Broemeling<sup>26</sup>, 1985) function (which we simply stated in Table 1 without any explanation). Effectively

$$\pi(\sigma^2 | r_1, r_2, \dots, r_n) \sim \text{Gamma}(\alpha, \beta)$$

where

$$\alpha = \frac{n-1}{2} \text{ and } \beta = \frac{(n-1)s^2}{2}$$

When using the squared error loss function (see Section 3.4) the expected loss function is minimised by taking

$$\begin{aligned} \hat{\sigma}^2 &= E(\sigma^2 | r_1, r_2, \dots, r_n) \\ &= \frac{\beta}{\alpha} \quad \text{i.e. the expected value of Gamma} \\ &= \frac{(n-1)s^2}{2} \div \frac{n-1}{2} \\ &= s^2 \end{aligned}$$

i.e. the Bayesian estimate of risk,  $\sigma^2$  is simply the sample variance,  $s^2$  (which is what we simply stated in Table 2 without any explanation).

2. For inferences regarding  $\mu$  one needs to integrate out the  $\sigma^2$  in  $\pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n)$ .

The marginal posterior of  $\mu$  is:

$$\begin{aligned} \pi(\mu | r_1, r_2, \dots, r_n) &= \int_0^\infty \pi(\mu, \sigma^2 | r_1, r_2, \dots, r_n) d\sigma^2 \\ &= \int_0^\infty \left( \frac{1}{\sigma^2} \right)^{\frac{(n-1)}{2}-1} \exp\left[-\frac{(n-1)s^2}{2\sigma^2}\right] (\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{n}{2\sigma^2}(\mu - \bar{r})^2\right] d\sigma^2 \end{aligned}$$

<sup>26</sup> Broemeling, L.D, "Bayesian Analysis of Linear Models", pp. 66, 1985.

$$\begin{aligned}
&= \int_0^{\infty} (\sigma^2)^{\frac{-n+1}{2}} (\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{(n-1)s^2}{2\sigma^2}\right] \exp\left[-\frac{n(\mu-\bar{r})^2}{2\sigma^2}\right] d\sigma^2 \\
&= \int_0^{\infty} (\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\left\{(n-1)s^2 + n(\mu-\bar{r})^2\right\} \frac{1}{\sigma^2}\right] d\sigma^2
\end{aligned}$$

(Appendix Equation 5)

Let  $y = \frac{1}{2} A \frac{1}{\sigma^2}$  where  $A = (n-1)s^2 + n(\mu-\bar{r})^2$  then

$$\begin{aligned}
\left| \frac{dy}{d\sigma^2} \right| &= \left| -\frac{1}{2} A (\sigma^2)^{-2} \right| \\
&= \left| \frac{1}{2} A (\sigma^2)^{-2} \right|
\end{aligned}$$

$$\text{Since } \int_0^{\infty} y^{\frac{n}{2}-1} \exp(-y) dy = \Gamma\left(\frac{n}{2}\right)$$

$$\pi(\mu | r_1, r_2, \dots, r_n) = \Gamma\left(\frac{n}{2}\right) A^{-\frac{n}{2}} \quad \text{from (Appendix Equation 5)}$$

The data  $r_1, r_2, \dots, r_n$ ,  $s^2$  and  $\bar{r}$  are constants thus if we set

$$t = \frac{\mu - \bar{r}}{s/\sqrt{n}} \quad \text{then}$$

$$A = (n-1)s^2 + s^2 t^2 \quad \text{and}$$

$$\left| \frac{\partial t}{\partial \mu} \right| = \frac{\sqrt{n}}{s} \quad \text{a constant thus}$$

$$\begin{aligned}
\pi(t | r_1, r_2, \dots, r_n) &\propto A^{-\frac{n}{2}} \\
&\propto [(n-1)s^2 + s^2 t^2]^{-\frac{n}{2}} \\
&\propto \left[ 1 + \frac{t^2}{n-1} \right]^{-\frac{n}{2}}
\end{aligned}$$

This posterior distribution is the probability density function for the Student's t with n-1 degrees of freedom. Effectively,

$$\pi(\mu \mid r_1, r_2, \dots, r_n) \sim t_{n-1}$$

Which is what we simply stated in Table 1 without any accompanying explanation. Applying the squared error loss function (see Section 3.4) the Bayesian estimate is:

$$E(\mu \mid r_1, r_2, \dots, r_n) = \bar{r} \quad (\text{which is what we stated in Table 2})$$

It is noticeable that the Bayesian estimate for  $\mu$ , the security's expected return, is the same as the estimated obtained when using standard sampling theory.

## 6.4 Appendix 4

In Table 1 and in Table 2 we described the scenario where the portfolio manager has *no* subjective views on the security's expected return,  $\mu$ , or on the security's uncertainty (variance),  $\sigma^2$ . An assumption we did make in this scenario was that the portfolio manager believes that the two parameters  $\mu$  and  $\sigma^2$  do influence the outcome of one another i.e. they are *dependent* on one another.

In Table 1 we simply *stated* the joint probability density function that describe the parameter  $\mu$  and  $\sigma^2$  simultaneously. Here in Appendix 4 we shall justify why we selected the specific joint probability density function listed in Table 1.

In Table 1 we also simply stated the posterior "conditional" distributions. In this appendix we shall discuss how one obtains these posterior probability density functions. These posterior distributions are related to the joint probability density function.

The final Bayesian estimates, which are simply stated in Table 3, are obtained from the posterior "conditional" distribution. We discuss how these are obtained below:

In Section 3.2 we assumed that the security's variance,  $\sigma^2$  was known for certain. We now consider the Bayesian estimate, which is based upon the assumption that the investor has no definite knowledge or beliefs in the security's variance  $\sigma^2$ . We shall also *not* assume that the security's expected return and variance are independent of one another.

For convenience we shall be using the symbol  $\tau$  where

$$\tau = \text{precision} = \frac{1}{\sigma^2}$$

In Section 3 we established

$$\begin{aligned}
\Lambda(\mu, \sigma^2; r_1, r_2, \dots, r_n) &= \left[ \frac{1}{2\pi\sigma^2} \right]^{\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \left\{ (n-1)s^2 + n(\bar{r} - \mu)^2 \right\} \right] \\
&\propto \left[ \frac{1}{\sigma^2} \right]^{\frac{n}{2}} \exp \left[ -\frac{(n-1)s^2}{2\sigma^2} \right] \exp \left[ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right] \\
&\propto \left[ \left( \frac{1}{\sigma^2} \right)^{\frac{n-1}{2}} \exp \left\{ -\frac{(n-1)s^2}{2\sigma^2} \right\} \right] \left[ \left( \frac{1}{\sigma^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{n(\bar{r} - \mu)^2}{2\sigma^2} \right\} \right]
\end{aligned}$$

Replacing  $\sigma^2$  with  $\frac{1}{\tau}$  we now have

$$\Lambda(\mu, \tau; r_1, r_2, \dots, r_n) \propto \left[ \tau^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau(n-1)s^2}{2} \right\} \right] \left[ \tau^{\frac{1}{2}} \exp \left\{ -\frac{n\tau(\bar{r} - \mu)^2}{2} \right\} \right]$$

now,

$$\pi(\mu, \tau | r_1, r_2, \dots, r_n) \propto \pi(\mu, \tau) \Lambda(\mu, \tau; r_1, r_2, \dots, r_n)$$

We have assumed that the joint distribution function  $\pi(\mu, \tau)$  is information less (which is what we simply stated in Table 1 without any explanation), i.e. the portfolio manager has no subjective beliefs in the security's expected return,  $\mu$ , nor in its precision,  $\tau$ , hence

$$\pi(\mu, \tau) \propto \tau^{-1} \quad (\text{Jeffreys}^{27}, 1939) \text{ thus}$$

$$\begin{aligned}
\pi(\mu, \tau | r_1, r_2, \dots, r_n) &\propto \tau^{-1} \left[ \tau^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau(n-1)s^2}{2} \right\} \right] \left[ \tau^{\frac{1}{2}} \exp \left\{ -\frac{n\tau(\bar{r} - \mu)^2}{2} \right\} \right] \\
&\propto \left[ \tau^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau(n-1)s^2}{2} \right\} \right] \left[ \tau^{\frac{1}{2}} \exp \left\{ -\frac{n\tau(\bar{r} - \mu)^2}{2} \right\} \right]
\end{aligned}$$

by integrating this function with respect to  $\mu$  we obtain:

<sup>27</sup> Jeffreys, H, "Theory of Probability", 1939.

$$\begin{aligned}\pi(\tau | r_1, r_2, \dots, r_n) &\propto \left[ \tau^{\frac{n-3}{2}} \exp\left\{-\frac{\tau(n-1)s^2}{2}\right\} \right] \\ &\propto \tau^\alpha \exp\{-\beta\tau\}\end{aligned}$$

where

$$\alpha = \frac{n-3}{2} \quad \text{and} \quad \beta = \frac{(n-1)s^2}{2}$$

Here  $\pi(\tau | r_1, r_2, \dots, r_n)$  follows a Gamma distribution with parameters  $\alpha+1$  and  $\beta$ . Equivalently  $\pi(\sigma^2 | r_1, r_2, \dots, r_n)$  follows a Gamma distribution with parameters  $\alpha$  and  $\beta$  (Broemeling, 1985) (which was stated in Table 1 without any accompanying explanation).

By using the squared error loss function (see Section 3.4) the Expected Loss function is minimised by taking:

$$\begin{aligned}\hat{\sigma}^2 &= E(\sigma^2 | r_1, r_2, \dots, r_n) \\ &= \frac{\beta}{\alpha} \quad \text{i.e. the expected value of a Gamma distribution} \\ &= \frac{n-1}{n-3} s^2 > s^2\end{aligned}$$

(This expectation was simply listed in Table 2 without any accompanying explanation)

Thus under the squared error loss the Bayesian estimate of the risk,  $\sigma^2$ , is not the familiar  $s^2$ . This Bayesian estimate can be considered “biased” in the standard sampling theory sense, i.e. the estimates obtained from repeated samplings from the same population (with the same parameter values) will not in general centre around the true value of this parameter being estimated. The reason is because the Bayesian estimate trades off bias error with variance, i.e. the bias is compensated for by a reduction in variance. In the non-informative subjective expectations case being considered, if one considered the following Loss Function:

$$L(\hat{\sigma}^2, \sigma^2) = \frac{(\hat{\sigma}^2 - \sigma^2)^2}{\sigma^2} \quad \text{then } EL(\hat{\sigma}^2 | \underline{r}) \text{ is minimal for } \hat{\sigma}^2 | \underline{r} \equiv s^2$$

which is then an “unbiased” estimate. This highlights the effect that the choice of the loss function has on the Bayesian estimate.

Considering the subjective distribution, conditional on the past returns, for the parameter  $\mu$ , i.e.  $\pi(\mu | r_1, r_2, \dots, r_n)$

$$\pi(\mu | r_1, r_2, \dots, r_n) \propto \left[ 1 + \frac{t^2}{n-1} \right]^{-\frac{n}{2}} \quad (\text{see Appendix 3})$$

This is the t-distribution with  $n-1$  degrees of freedom. Applying the squared error loss function (see Section 3.4) the Bayesian estimate is:

$$E(\mu | r_1, r_2, \dots, r_n) = \bar{r}$$

as stated in Table 2.

It is noticeable that the Bayesian estimate of  $\mu$ , the security's expected return, is the same as the estimate obtained when using standard sampling theory (when the portfolio manager has no subjective expectations).

## 6.5 Appendix 5

In Figure 1 Section 4.2 we used the efficient frontier to determine the levels of risk (standard deviation) for each of the respective optimal portfolios. We now wish to discuss the relative locations of these respective efficient frontiers. Figure 4 below is identical to Figure 1.

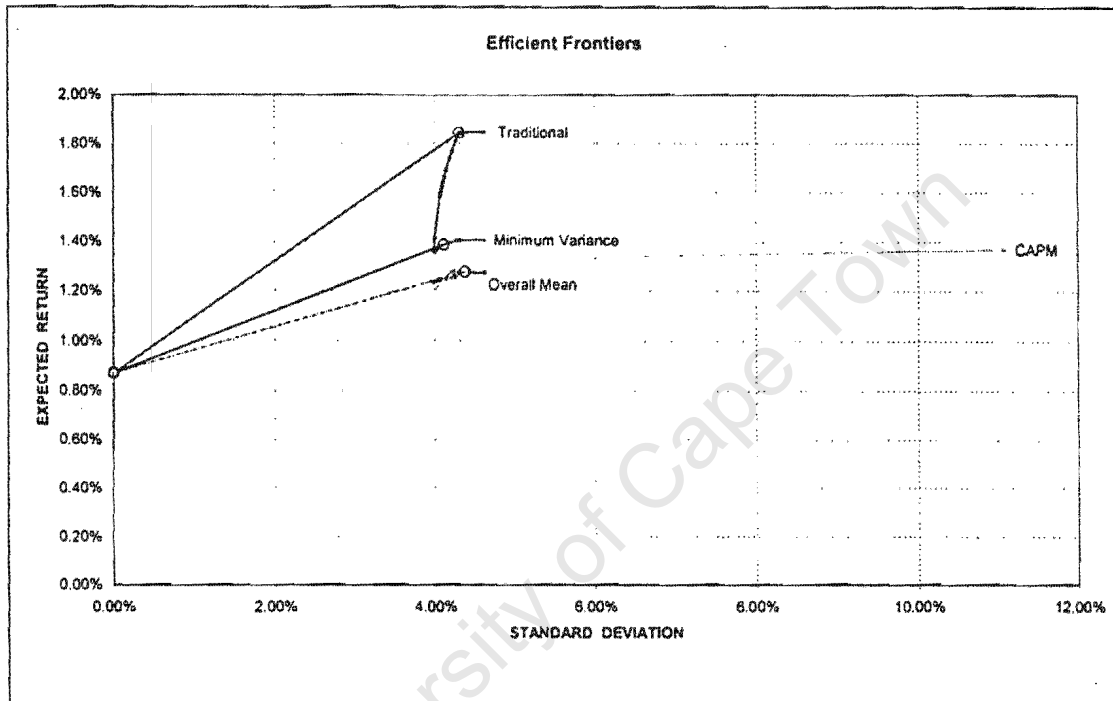


Figure 4 Efficient Frontier plots of the optimal portfolios when using 90% weights together with the CML optimisation technique

From Figure 4 it would appear that the Traditional efficient frontier dominates the other two efficient frontiers. One should not conclude that the Traditional optimal portfolios are more efficient than the CAPM and Bayesian CAPM optimal portfolios. The vertical axis in Figure 4 is not strictly correctly labelled because we are actually using four different measures of expected return. Figure 4 should actually be broken into four separate graphs. The vertical axes for these three graphs should be the Traditional Expected Return, CAPM Expected Return, Minimum Variance Expected Return and the Overall Mean Expected Return. The consequent graphs will look as follows:

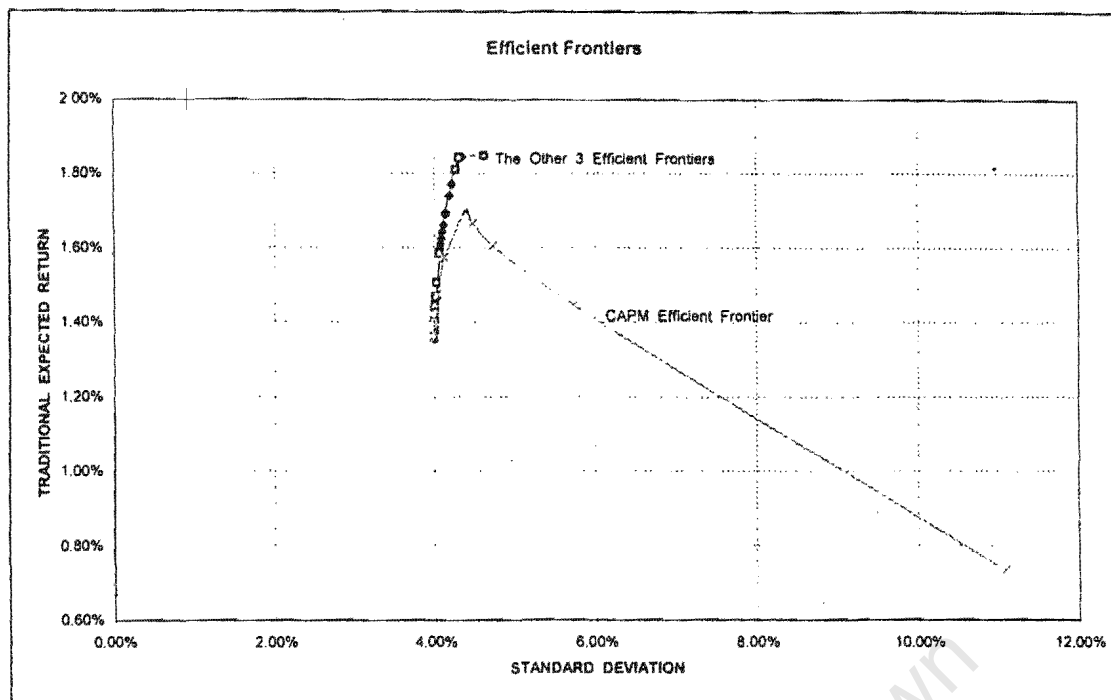


Figure 5 Efficient Frontiers in Traditional Expected Return-Standard Deviation space

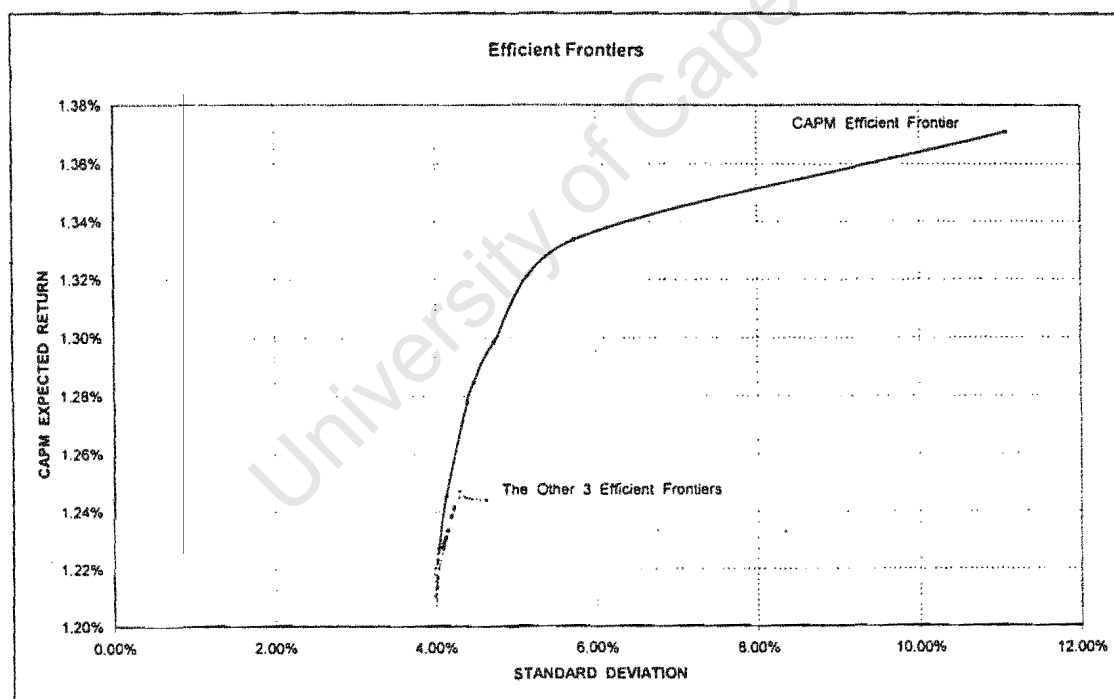


Figure 6 Efficient Frontiers in CAPM Expected Return-Standard Deviation space

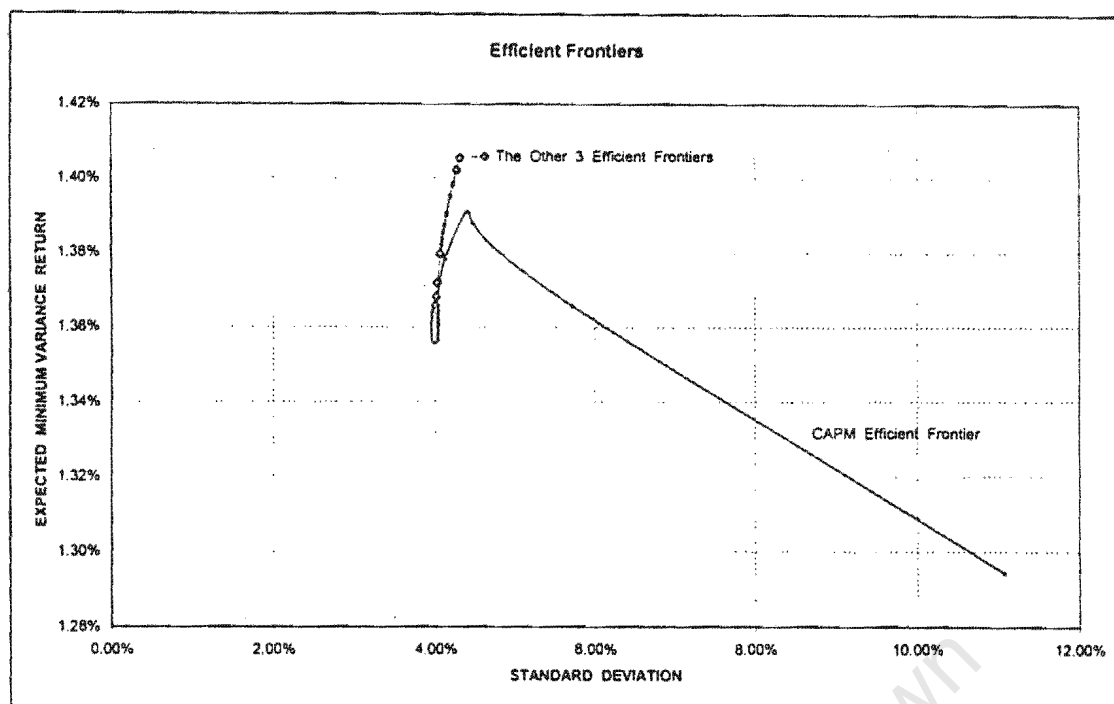


Figure 7 Efficient Frontiers in Expected Minimum Variance Return-Standard Deviation space

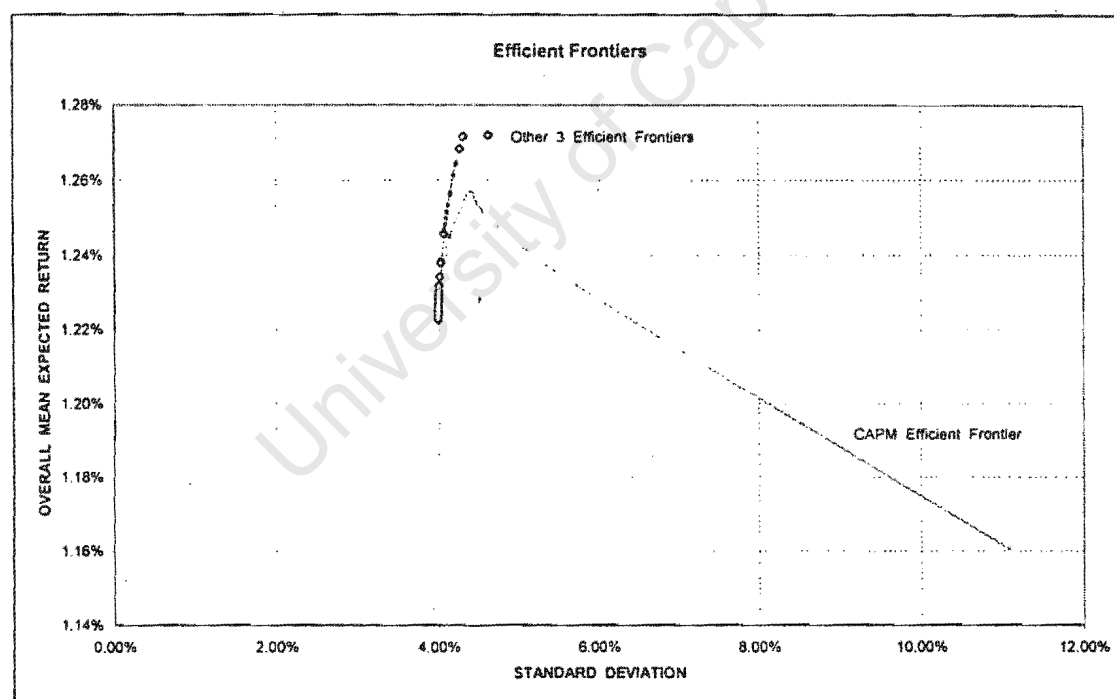


Figure 8 Efficient Frontiers in Overall Mean Expected Return-Standard Deviation space

The Traditional frontier dominates in Traditional Expected Return-Standard deviation space, the CAPM frontier dominates in CAPM Expected Return-Standard Deviation space and the Bayesian CAPM obviously dominates in Bayesian CAPM Expected

Return-Standard Deviation space. The dominance is so insignificant (certainly not noticeable in the graphs, except for the CAPM frontier in CAPM Expected Return-Standard Deviation space). This highlights the fact that optimal portfolios are not necessarily unique. There are obviously portfolios different in composition (achieved by using different portfolio selection models) but nevertheless very similar in expected return and variance. The optimal portfolio achieved using the Bayesian CAPM model is very inefficient in Figure 5, Figure 7 and in Figure 8.

University of Cape Town

## **CHAPTER 5**

### **BETA ESTIMATION**

University of Cape Town

## 1. INTRODUCTION

In Chapter 4 we observed that the application of Bayesian CAPM estimates for the securities' expected returns in the mean-variance optimisation algorithm results in *diversified* optimal portfolios. This diversification, which is often not achieved when using other estimates for the securities' expected returns, makes the Bayesian CAPM estimate attractive to portfolio managers. The Bayesian CAPM estimate is a weighting of a sample mean and a CAPM estimate. The CAPM estimate is consequently an important factor in Bayesian CAPM estimates, therefore we must ensure we obtain a reliable and accurate CAPM estimate. For this reason we have researched ways in obtaining reliable CAPM estimates.

An input into the CAPM estimate is the Beta estimate. The Beta is a measure of a stock's volatility relative to movements in the whole market. In this chapter we examine means of achieving a more *reliable* beta estimate. Consequently, this chapter is not examining the input estimations of the objective function directly but rather looking at a parameter of one of the input estimations we examined in Chapter 4. We include this chapter nevertheless as we feel it does make a contribution to the portfolio selection objective indirectly through the CAPM input estimate.

## 2. THE CAPITAL ASSET PRICING MODEL (CAPM)

It was Sharpe<sup>1</sup>, Lintner<sup>2</sup> and Treynor<sup>3</sup> (1964) who proposed the CAPM. The model proposes that in a competitive market the expected risk premium (the expected return above the risk-free rate) varies in direct proportion to beta,  $\beta$ , the measure of a security's market risk. We can write this relationship as:

**Expected risk premium on a security = beta × expected risk premium on market**

$$r_i - r_f = \beta_i(r_m - r_f) + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2 I) \quad \text{lost this}$$

where

$r_i$  = the return on security  $i$

$r_f$  = the risk-free rate

$r_m$  = the return on the overall market

$\beta_i$  = a measure of security  $i$ 's market risk

Making  $r_i$  the subject of the formula we obtain

$$r_i = r_f + \beta_i(r_m - r_f)$$

We can use this relationship to obtain an estimate for a security's expected return. In order to fulfil this we need to estimate security  $i$ 's beta ( $\beta$ ).

The beta of a security is determined by regressing the returns of the security against the returns of the overall market index. Similarly, in order to determine the beta of a sector index one regresses the returns of that sector index against the returns of the overall market index:

<sup>1</sup> Sharpe, WF, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, September 1964, pp. 425-442.

<sup>2</sup> Lintner, J, "The Validation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgeting", *Review of Economics and Statistics*, pp. 13-37, February 1965.

<sup>3</sup> Treynor's article has not been published.

$$r_{it} = \alpha_i + \beta_i r_{mt} + e_{it}$$

where

$r_{it}$  = the return on sector index  $i$  in the  $t$ -th period

$r_{mt}$  = the return on the overall market index in the  $t$ -th period

$\alpha_i$  = a parameter unique to sector index  $i$

$\beta_i$  = a parameter unique to sector index  $i$

$e_{it}$  = an error term unique to index  $i$  in the  $t$ -th period

This procedure is simple linear regression.

- (i) Choose for  $i$
- (ii) calculate
- (iii) calculate the correlation coefficient
- (iv) calculate the regression line
- (v) calculate the regression coefficient
- (vi) calculate the regression intercept
- (vii) calculate the regression equation
- (viii) calculate the regression equation
- (ix) calculate the regression equation
- (x) calculate the regression equation

### 3. OUTLIERS AND INFLUENTIAL POINTS

#### 3.1. Influential Points

An issue of concern is the existence of influential points and outliers. Such points will bias the estimate of  $\beta$ . A measure that can be used to determine the influential points

is  $\frac{V_{ii}}{1-V_{ii}}$  (Trosky<sup>4</sup>, 1998) where  $V_{ii}$  is the leverage of observation  $i$ .

$$V_{ii} = r_{m(i)} (r'_m r_m)^{-1} r'_{m(i)} = \text{hat matrix}$$

and

$$R_m = \begin{pmatrix} 1 & r_{m1} \\ 1 & r_{m2} \\ \dots & \dots \\ \dots & \dots \\ 1 & R_{mn} \end{pmatrix} = X$$

$\nearrow$  Constant.

$$r_{it} = \alpha_i + \beta_i (r_{mt}) + e_{it}$$

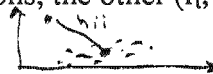
$\downarrow$   
X

$$X(X'X)^{-1}X' = \text{Hat Matrix} \checkmark$$

$$\text{leverage} = \frac{h_{ii}}{1-h_{ii}}$$

$r_{m(i)}$  is the  $i^{\text{th}}$  row of  $R_m$

The leverage,  $V_{ii}$ , is the measure of the distance between the point  $(r_i, r_{mi})$  and the mean of all the *remaining* pairs of observations, the other  $(r_i, r_{mi})$ 's. (distance to centre of convex hull)



The impact of an outlier, where the influence is measured by its leverage, is evident when examining it diagrammatically. Consider an extreme case where only one of the observations  $(r_i, r_{mi})$  has an extremely large leverage. Now if one had been able to identify this influential point, by means of its large leverage, and exclude it when determining the best least squares fit, the following regression line would have been calculated:

<sup>4</sup> Trosky's paper has not been published. It will be published in 1998. I had private consultations with Prof. Trosky.

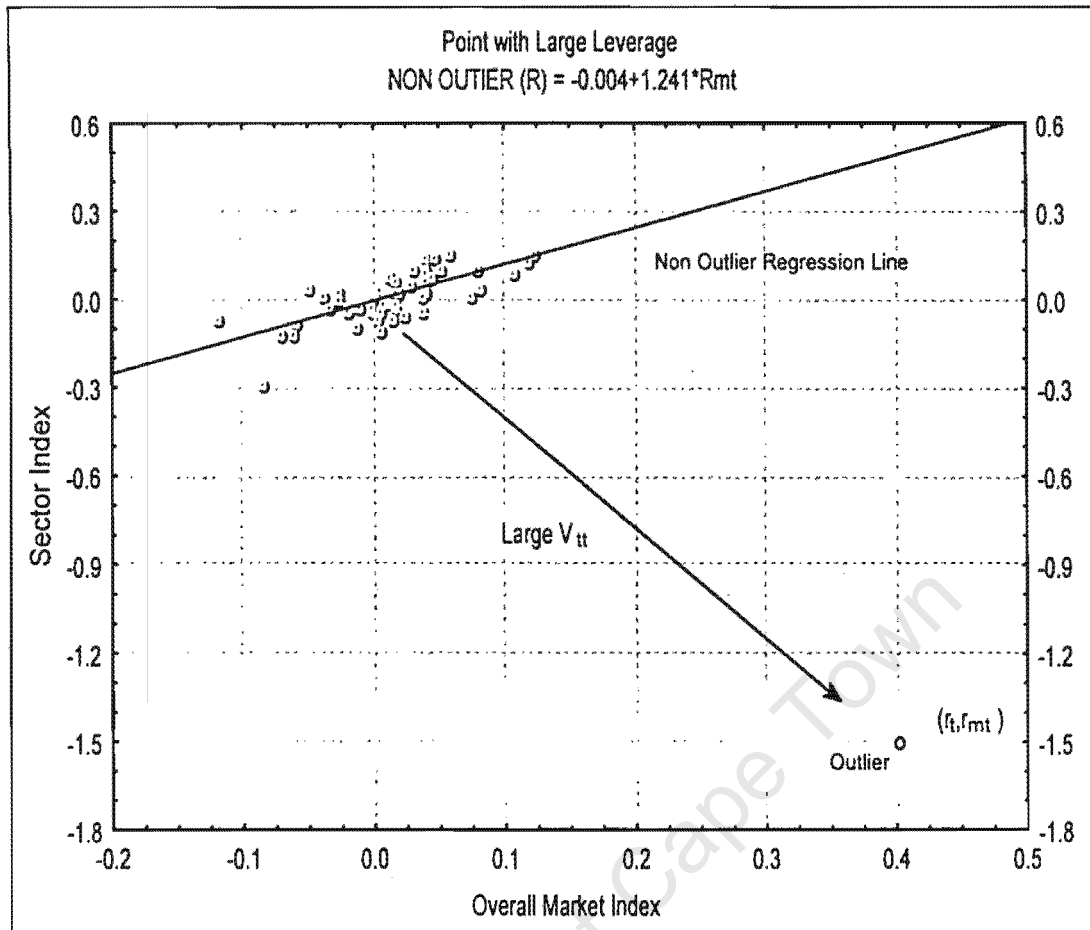


Figure 1 Regression line excluding the influential point

However, if the influential point had not been excluded when fitting the linear regression line a different regression line would have been produced:

*(Would actually go thru the point)!*

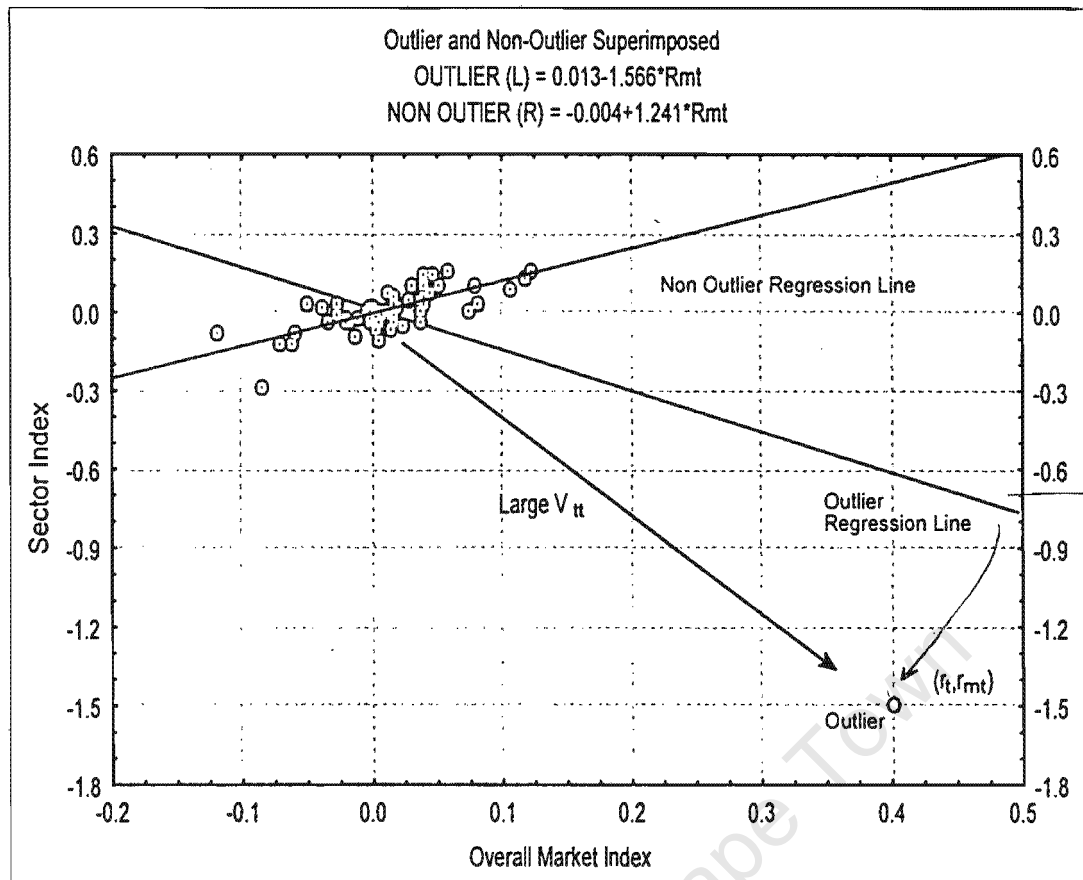


Figure 2 Regression line including the influential point

Hence, the beta produced by excluding the potential influential point is significantly different to the beta produced by including the potential influential point. This suggests that one should definitely attempt to identify potential influential points. If the influential points make a significant difference to the estimated beta then one should rather discard them and calculate the estimate for beta from only the remaining non-influential observations. *or down weight them and use weighted Regression.*

A "possibly" incorrect measure of the influence that an observation might have is its residual  $\hat{e}_i$ , where  $\hat{e}_i = r_i - \hat{r}_i$ . This is evident when one considers an "extreme" influential point. The residual of a highly influential point will actually be very small and will certainly not be as large as often expected. The reason is because the regression line is "pulled" towards the corresponding influential point  $(r_i, r_{mt})$  making  $\hat{e}_i$  small. Mathematically this can be shown as follows (Trosky, 1998):

*→ ? "cannot be correct" look @ figure 2.*

? = Actually true since if u only have 1 influential point, the regression line would pass thru the point but your sketch does not show this. Rnew should shift/lever more towards the outlier.

$$\hat{r} = r_m (r'_m r_m)^{-1} r'_m r = V r$$

where

$$V = r_m (r'_m r_m)^{-1} r'_m$$

therefore,

$$\hat{r}_i = V_{ii} r_i + \sum_{j=1}^n V_{ij} r_j$$

diag terms      off diag

now

$$0 \leq V_{ii} \leq 1 \quad \text{and} \quad \sum V_{ii} = p+1 = 2$$

(since  $0 \leq h_{ij} \leq 1$  is only possible region)

where  $p$  is the number of predictor variables i.e. 1

$$h_{ij} = h_{ij} + \sum_{i \neq j} h_{ij}^2$$

$V$  is  $N \times N$

$$\sum V_{ii} = \text{tr}(V) = \sum_{i=1}^n \lambda_i$$

$\lambda_i$  is eigenvalues.  
but how does this help??

$$\text{Average leverage} = \frac{\sum V_{ii}}{n}$$

If  $V_{ii} = 1$  then all  $V_{ij} = 0$  for  $j \neq i$  (Trosky, 1998). ✓

The most extreme leverage (influence) occurs when  $V_{ii} = 1$ . In this extreme case:

$$\begin{aligned} \hat{r}_i &= 1 \cdot r_i + 0 \\ &= r_i \\ \Rightarrow \hat{e}_i &= 0 \quad \checkmark \quad \text{same point.} \end{aligned}$$

$$\text{ie } h_{ii} > \frac{2k}{n}$$

$$k = p+1$$

If  $V_{ii}$ , the leverage, is large then one should eliminate that observation if it has a

substantial effect on the estimate of beta. We observe that  $\sum_{i=1}^n V_{ii} = p+1 = 2$ , therefore

the average  $V_{ii}$  is  $\frac{p+1}{n} = \frac{2}{n}$ . One could consider  $V_{ii}$  as being possibly significant if

$V_{ii} > 2 \times \frac{2}{n} = \frac{4}{n}$  (more accurately  $1.96 \times \frac{2}{n}$  if we use the 5% significance level). It is

$$\frac{2k}{n}$$

imperative that one notes that the leverage is only a measure of the influence of each observation taken separately and it assumes that the other observations are non-influential observations.

### 3.2. Student's $t$ Statistic

A further statistic that one could apply when determining possible outliers is the Student's  $t$  statistic:

$$t_i = \frac{\hat{e}_i}{\sqrt{s^2(u)(1-V_{ii})}} \sim t_{n-k-1} \quad (\text{Stefanski!!})$$

$t_i$  is calculated for each observation. It is based on examining the residuals and consequently has the shortcomings that were mentioned in Section 3.1. As with the leverage measure, it is considering each observation individually and can be used to determine whether a particular observation is a possible outlier on the assumption that the other observations are not outliers. The actual test statistic for determining whether there is at *least one* outlier present is  $T = \max |t_i|$ . This test would only be able to determine whether there is an outlier present or not. This test cannot be used to determine whether there are exactly  $x$  outliers because once  $x$  is greater than 1 the significance level becomes too large and hence insignificant (Trosky, 1998). ✓

(Should explain why!  $\alpha/(p)$ )  
 Since the  $\alpha/(p)$  point is  $F_{p, n-k-p}$   
 and this becomes very large, thus  
 one will reject the null hypothesis.  
 if there being  $p$  outliers. Should  
 rather use SVD to find/identify the  
 outliers!)

$$s^2(u) = \frac{1}{n-k-1} (\hat{\mathbf{e}}' \hat{\mathbf{e}} - \hat{\mathbf{e}}_t' \mathbf{M}_p^{-1} \hat{\mathbf{e}}_t)$$

### 3.3. Cook's Distance

A further statistic one could apply in order to identify "potential" outliers is Cook's<sup>5</sup> distance:

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_{(-i)})' r_m' r_m (\hat{\beta} - \hat{\beta}_{(-i)})}{(p+1)s^2} \sim F_{((p+1), (n-(p+1)))}^\alpha$$

A large  $D_i$  indicates the existence of an influential point or outlier. It combines the effect of the observation possibly being an influential point together with it possibly being an outlier. This is evident when one simplifies  $D_i$  (Cook 1977):

$$\begin{aligned} D_i &= \frac{\hat{e}_i^2}{(p+1)s_{(i)}^2(1-V_u)} \cdot \frac{V_u}{1-V_u} \quad \checkmark \\ &= \frac{Z_i^2}{p+1} \cdot \frac{V_u}{1-V_u} \\ &\quad (1) \quad (2) \end{aligned}$$

where

$$Z_i = \frac{\hat{e}_i}{\sqrt{s_{(i)}^2(1-V_u)}}$$

$Z_i$  is the Studentised residual and is used in the test for outliers, (see Section 3.2), and

$\frac{V_u}{1-V_u}$  measures the effect of the leverage and is an indicator of the influence that

each observation has. Cook's distance combines the outlier and influential effects.

<sup>5</sup> Cook, R.D, "Detection of Influential Observations in Linear Regression", *Technometrics*, pp. 15-18, 1977.

#### 4. OUTLIER SHIFT MODEL <sup>-very NB</sup>

By means of examining the leverages, the Students' *t* statistics and the Cook's distances one could possibly make a more informative decision in determining the outliers and influential points. However, these measures examine each observation *individually* and do not consider their combined influence. One approach in determining *all* possible outliers *simultaneously* is the application of the outlier shift model. The regression model that has been discussed thus far is:

$$\underline{r} = \underline{r}_m \underline{\beta}_m + \underline{e}$$

where

$$\underline{r} = \begin{pmatrix} r_1 \\ \dots \\ \dots \\ r_n \end{pmatrix}$$

$$\underline{r}_m = \begin{pmatrix} 1 & r_{m1} \\ \dots & \dots \\ \dots & \dots \\ 1 & r_{mn} \end{pmatrix}$$

$$\underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$\underline{e} = \begin{pmatrix} e_1 \\ \dots \\ \dots \\ e_n \end{pmatrix}$$

This model will be termed the basic model. The outlier shift model is merely an adjustment to this basic model. An additional variable is added to the basic model for

each suspicious outlier. If there are  $q$  suspicious outliers then  $q$  new variables are added to the basic model, viz.  $\lambda_1, \lambda_2, \dots, \lambda_q$ . The potential outliers can be determined by means of Cook's distances.  $\lambda_i$  is a binary variable. If one believes that the  $k^{\text{th}}$  observation is a possible outlier then one should add the variable  $\lambda_k$  to the basic model where:

$$\underline{\lambda}_k = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} \begin{matrix} \text{observation 1} \\ \text{observation 2} \\ \dots \\ \text{observation } k \\ \dots \\ \text{observation } n \end{matrix}$$

The new outlier shift model is as follows:

$$\begin{aligned} \underline{r} &= \underline{r}_m^* \beta^* + \underline{e} \\ &= \underline{r}_m^* \beta^{*(1)} + \underline{\lambda} \beta^{*(2)} + \underline{e} \\ (1) \quad & \quad (2) \end{aligned}$$

where (1) represents the variables contained in the basic model and (2) contains the new shift variables. In order to determine the significant regressors one can apply stepwise linear regression. The variables that remain in the model after the stepwise process are the significant variables. If  $\lambda_i$  is in the model once the stepwise process is complete then the  $i^{\text{th}}$  observation,  $r_i$ , is an outlier.  $r_i$  will therefore be an influential observation and should be removed. The adjusted beta for the security or index being considered can be calculated by considering only the non-influential observations.

## 5. EMPIRICAL RESEARCH

The returns of the JSE sector indices were regressed against the returns of the JSE overall index. 5 years, 1992-1996, of monthly returns were used in the analysis. Accordingly, there were 60 observations for each sector index.

### 5.1. Cook's Distances

Firstly, one wishes to determine the possible outlier/influential monthly returns. The method used for this task was Cook's distance. Cook's distance,  $D_t$ , was calculated for each observation. Hence, 60 Cook's distances were calculated for each sector index. Now  $D_t \sim F_{(p+1, n-(p+1))}$ , where  $p$  is the number of predictor variables. There is only one predictor variable, viz. The JSE-Overall Index, therefore,

$$D_t \sim F_{(2, 60-2)} = F_{2, 58}$$

$$\text{now } F_{crit} = F_{(p+1, n-(p+1))}^{\alpha}$$

Taking  $\alpha = 0.1$ ,

$$F_{crit} = F_{2, 58}^{0.1} = 0.106$$

If  $D_t > 0.106$  then, according to Cook's distance, observation  $t$  is a possible outlier.

The results obtained from each sector index considered separately were very similar. Considering the Mining-Diamond index the following results were obtained:

Table 1 lists the observations with the four largest Cook's distances.

**Table 1** Cook's Distances

| Observation   | Cook's Distance |
|---------------|-----------------|
| 1992-August   | .597            |
| 1995-January  | .234            |
| 1996-May      | .087            |
| 1993-November | .054            |

The first two observations in Table 1 have Cook's distances that are larger than  $F_{crit}$ , but the other two observations also have relatively large Cook's distances and will consequently also be considered as possible influential observations.

### 5.2. Outlier Shift Model

An outlier shift model is then created. The predictor variables will consist of the monthly return of the overall JSE index and four binary variables. These four binary variables will represent the four possible influential observations. The outlier shift model is as follows:

$$r_{diamond\ index} = \beta_0 + \beta_1 r_{overall\ index} + \lambda_1 r_{1992-August} + \lambda_2 r_{1995-January} + \lambda_3 r_{1996-May} + \lambda_4 r_{1993-November} + e$$

### 5.3. Forward Stepwise Regression

Table 2 is a summary of the forward stepwise regression. The variables are listed in the order in which they entered into the model.

**Table 2** Forward Stepwise Regression

| Variable      | Step In | p-level |
|---------------|---------|---------|
| Overall Index | 1       | .0000   |
| 1992-August   | 2       | .0004   |
| 1996-May      | 3       | .0737   |
| 1993-November | 4       | .1216   |
| 1995-January  | 5       | .1116   |

The model is highly significant at step 1 ( $p = .0000$ ) and step 2 ( $p = .0004$ ) and also certainly significant at step 3 ( $p = .0737$ ). At step 2 the variables Overall Index and

August 1992 had entered into the model. Table 3 lists the significance of each of these two highly significant predictor variables that entered into the outlier shift model at step 2.

Table 3 Significance of Predictor Variables

| Variable      | B     | St. Err. Of B | t <sub>57</sub> | p-level |
|---------------|-------|---------------|-----------------|---------|
| Intercept     | .0005 | .007          | .074            | .9410   |
| Overall Index | 1.087 | .146          | 7.451           | .0000   |
| 1992-August   | -.195 | .052          | -3.773          | .0004   |

This indicates that the return of August 1992 is a definite outlier as it is highly significant ( $p = .0004$ ) and should certainly be removed from the model before calculating beta ( $\beta$ ).

#### 5.4. Backward Stepwise Regression *reworkish*

In order to gain further support for considering the returns of August 1992 and possibly May 1996 as being outliers one can perform the backward stepwise regression on the outlier shift model. Table 4 is a summary of the backward stepwise regression. The variables are listed in the order in which they were removed from the model.

Table 4 Backward Stepwise Regression

| Variable      | Step Out | p-level |
|---------------|----------|---------|
| 1995-January  | 1        | .111    |
| 1993-November | 2        | .121    |
| 1996-May      | 3        | .073    |

The model is significant at step 3 ( $p = .073$ ), once the 3 binary variables, which are listed in Table 4, have been removed from the model. At step 3 the outlier shift model will only contain two predictor variables. The significance of each of these predictor variables is listed in Table 5.

Table 5 Significance of Predictor Variables

| Variable      | B     | St. Err. Of B | t <sub>57</sub> | p-level |
|---------------|-------|---------------|-----------------|---------|
| Intercept     | .0005 | .007          | .074            | .9410   |
| Overall Index | 1.087 | .146          | 7.451           | .0000   |
| 1992-August   | -.195 | .052          | -3.773          | .0004   |

Table 5 is identical to Table 3 since both models consist of the same two predictor variables. Consequently, since both the forward and backward stepwise procedures classify August 1992 to be an outlier, it should be removed from the basic model.

### 5.5. Consequence of Removing the Outlier

It has been established that the return of August 1992 is an outlier. Two data sets will now be compared. The first data set consists of the 60 pairs of monthly returns covering the period 1992-1996. The second data set consists of only 59 pairs of monthly returns. The Overall index return and the Mining-Diamond index return for August 1992 have been removed from the first data set to produce the second data set of 59 pairs of observations.

Table 6 contains the regression summary of the two data sets.

Table 6 Regression Summary

| Data Set            | B     | St. Err. of B | R <sup>2</sup> | t <sub>obs</sub> | p-level |
|---------------------|-------|---------------|----------------|------------------|---------|
| All 60 Observations | 1.241 | .155          | .524           | 7.993            | .000    |
| August 1992 Removed | 1.087 | .146          | .493           | 7.451            | .000    |

By removing the August 1992 return it is evident that the estimate of beta changes. In this example it is reduced from 1.241 to 1.087. 1.087 can be considered to be the "adjusted" beta. The adjusted beta will not necessarily be consistently greater than or consistently less than the unadjusted beta. It all depends upon the influence of the outlier. In the example being considered the month of August 1992 produced an abnormally large negative return on the Mining-Diamond Index. The return for the Mining-Diamond Index in August 1992 was  $-0.287$  and is substantially less than the mean monthly return for the Mining-Diamond Index of  $0.0096$ . In fact, the return for the Mining-Diamond Index in August 1992 certainly lies outside the 95% confidence

interval for the mean monthly Mining-Diamond index return. The 95% confidence interval is (-0.011; 0.298).

Graphically, the outlier is also exposed.

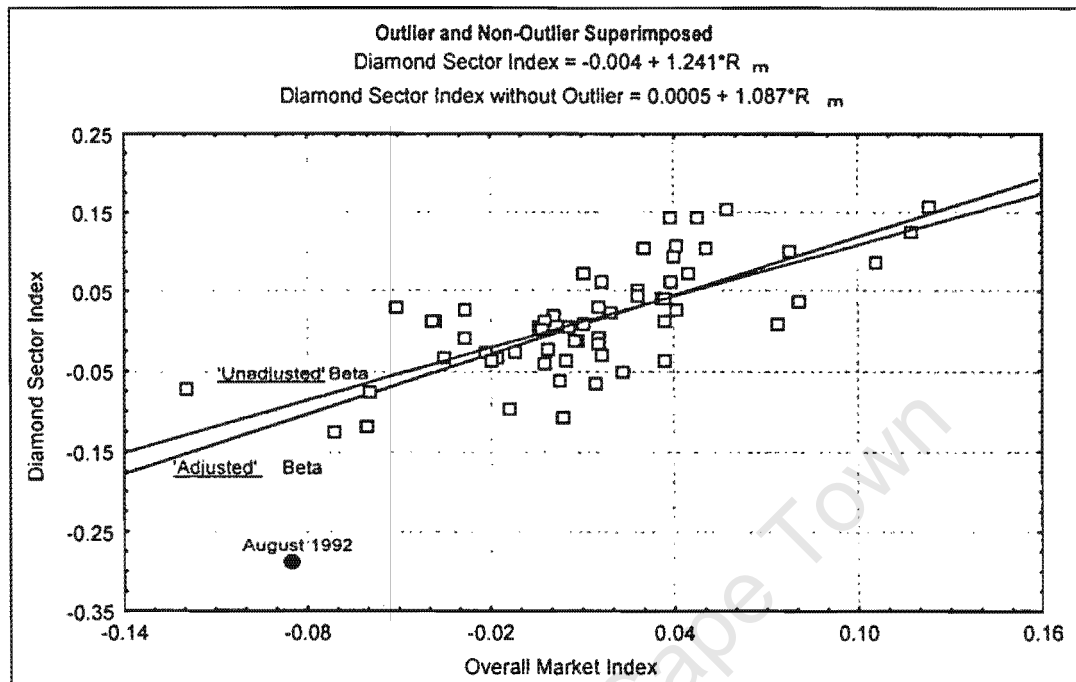


Figure 3 The effect of an outlier on the Beta estimate

The explained variance,  $R^2$ , of the adjusted model decreased by a marginal amount from .524 to .493. Table 7 combines the Anova table of the basic model and the Anova table of the outlier shift model.

Table 7 ANOVA Table

|                     | Effect     | Sums of Squares | df | Mean Squares | F     | p-level |
|---------------------|------------|-----------------|----|--------------|-------|---------|
| Basic Model         | Regression | .189            | 1  | .189         | 63.88 | .000    |
|                     | Residual   | .172            | 58 | .003         |       |         |
|                     | Total      | .361            | 59 |              |       |         |
| Outlier Shift Model | Regression | .133            | 1  | .134         | 55.52 | .000    |
|                     | Residual   | .137            | 57 | .002         |       |         |
|                     | Total      | .271            | 58 |              |       |         |

Both models are highly significant ( $p=.000$ ). All of the sums of squares have declined in the outlier shift model. Part of the reason for this general decline is because the outlier shift model has only 59 observations whereas the basic model has 60 observations. The outlier shift model has one observation less since the August 1992

returns have been removed. The reason for the decline in the  $R^2$  in the outlier shift model is because the SSR in this model has declined at a greater rate than the SSTO when compared with the comparative SSR and SSTO in the basic model. The  $R^2$  of the outlier shift model will not necessarily be less than the  $R^2$  of the basic model. In this case the  $R^2$  of the two models are very similar (.524 and .493) and thus there is no cause for concern.

A certain benefit of an outlier shift model is that its standard error of the beta estimate will be less than the standard error of the basic model's beta estimate. In this investigation the corresponding standard errors are .155 and .146 (see Table 6). This will give one greater confidence in the beta estimate obtained from the outlier shift model in comparison with the confidence one can obtain from the beta estimate from the basic model.

Effectively, the outlier shift model gives each of the outliers a weight of zero. In reality these observations are true realisations and consequently one might well be overreacting by giving them weights of zero. A similar approach to the outlier shift model is the robust method. The robust method, like the outlier shift model, recognises the existence of outliers, however, it does not discard them by giving them a weight of zero. Instead it gives the outliers a weight that is less than the weight of the non-outlier observations. The **extreme** outliers also receive a weight that is significantly less than the weight received by the **marginal** outliers.

One should bear in mind that one is examining sector indices, which might well be far more stable than individual securities. If one considered individual securities one might well encounter far more outliers in the equivalent 5 year period. Another factor to take note of is that in this example monthly returns are being examined. If weekly returns were examined, far more outliers might well be encountered, which would then certainly have an effect on the adjusted betas.

### **5.6. *Dates which Consistently Influenced the $\beta$ Estimate***

The previous discussion only considered the Mining-Diamond sector index. It was determined that the return of August 1992 was a significant outlier. Now the outlier shift model will be used to determine the significant outliers for each sector index. Again the forward stepwise regression technique will be applied. In the forward stepwise model an F-to-enter of 4.01 and an F-to-leave of 1 were specified. An F-to-enter value of 4.01 was used since the p-value of an  $F_{1,56}$ -variate (which is a model with 3 potential outliers) of 4.01 is 0.05 which is the conventional 5% significance level. The outliers are listed in Table 8.

Table 8 Outliers

| SECTOR                            | FORWARD<br>STEPWISE<br>OUTLIER |
|-----------------------------------|--------------------------------|
| <b>MINING PRODUCERS</b>           | May-1993                       |
| COAL                              | March-1994                     |
|                                   | April-1996                     |
| DIAMONDS                          | August-1992                    |
| ALL GOLD                          | May-1993                       |
| Rand & Others                     | May-1993                       |
|                                   | April-1993                     |
| Evander                           | December-1996                  |
|                                   | January-1996                   |
| Klerksdorp                        | May-1993                       |
|                                   | January-1996                   |
|                                   | December-1993                  |
| West Witwatersrand                | May-1993                       |
| <b>METALS &amp; MINERALS</b>      |                                |
| Manganese                         | January-1993                   |
| <b>MINING FINANCIAL</b>           |                                |
| Mining Holdings                   | October-1992                   |
| <b>FINANCIAL &amp; INDUSTRIAL</b> |                                |
| <b>FINANCIAL</b>                  |                                |
| Banks & Other Fin. Serv.          | November-1994                  |
| Insurance                         | August-1992                    |
| Investment Trusts                 | May-1992                       |
|                                   | January-1994                   |
| Property                          | June-1994                      |
| Property Trust                    | December-1993                  |
| Property Loan Stock               | May-1993                       |
|                                   | May-1996                       |
| <b>INDUSTRIAL</b>                 | May-1993                       |
| Beverage, Hotel & Leisure         | June-1994                      |
| Building & Construction           | May-1993                       |
| Chemicals & Oils                  | December-1993                  |
| Clothing, Footwear & Text         | May-1994                       |
| Electronics, etc.                 | October-1992                   |
| Electronics, etc.                 | May-1993                       |
| Furniture & Household             | May-1993                       |
| Furniture & Household             | May-1994                       |
| Motor                             | May-1993                       |
| Pharmaceutical, Medical           | August-1992                    |
| Pharmaceutical, Medical           | December-1993                  |
| Steel & Allied                    | March-1994                     |
| Stores                            | May-1996                       |

An issue worth investigating is whether a particular month had a significant number of outliers. Table 9 lists the months that had significant outliers based on the forward stepwise procedure. It also lists the months that were *considered* as possibly having significant outliers based on Cook's distances, as discussed in Section 3.3.

Table 9 Months with significant outliers

| MONTH         | # OF SECTORS. | #.OF SECTORS EXHIBITING OUTLIER BEHAVIOUR |                  |
|---------------|---------------|-------------------------------------------|------------------|
|               |               | COOK'S DISTANCES                          | FORWARD STEPWISE |
| May-1992      | 32            | 1                                         | 1                |
| August-1992   | 32            | 4                                         | 3                |
| October-1992  | 32            | 3                                         | 2                |
| December-1992 | 32            | 1                                         | 0                |
| January-1993  | 32            | 1                                         | 1                |
| April-1993    | 32            | 1                                         | 1                |
| May-1993      | 32            | 14                                        | 8                |
| November-1993 | 32            | 1                                         | 0                |
| December-1993 | 32            | 10                                        | 4                |
| January-1994  | 32            | 1                                         | 1                |
| March-1994    | 32            | 3                                         | 2                |
| May-1994      | 32            | 2                                         | 2                |
| June-1994     | 32            | 2                                         | 2                |
| November-1994 | 32            | 1                                         | 1                |
| January-1995  | 32            | 2                                         | 0                |
| January-1996  | 32            | 3                                         | 2                |
| April-1996    | 32            | 1                                         | 1                |
| May-1996      | 32            | 2                                         | 2                |
| December-1996 | 32            | 1                                         | 1                |

In order to test whether the number of significant outlier sector indices for each month were more than sampling theory suggests, a test<sup>6</sup> based on the binomial distribution was performed at the 5% significance level. This test indicates that with 32 sector indices being simultaneously tested at the 10% level, 6 or more would have to be rejected to constitute the existence of significant outliers in the month being examined. Inspection of Table 9 reveals that the Cook's distance measure and the forward stepwise model technique disclose the existence of a significant number of outliers in the highlighted months. These two months are May 1993 and December 1993. This basically informs one that in these two months it was not only one or two of the sectors which had outlier observations but instead a significant number of the sectors (more than 6) produced outliers. Repeating the analysis at the 2.5% significance level it is evident that the month of May 1993 is still significant.

<sup>6</sup> This is a binomial process with  $n = 32$  and  $p = 0.1$ , which can be approximated by the normal distribution with  $\mu = np$  and  $\sigma^2 = np(1-p)$ . At a 5% significance level the test statistic becomes:

$$np + 1.64\sqrt{np(1-p)} = 6.83$$

### **5.7    *Insights for Practitioners***

- One has greater confidence in the beta estimate when the outliers are removed.
- It is apparent that the securities' returns in the month of May 1993 were significantly abnormal as 25% of the sector indices had returns that were classified as outliers.

University of Cape Town

## 6. CONCLUSION

We have focused on achieving a reliable and accurate estimate for Beta,  $\beta$ , which is a parameter in the CAPM. This will consequently improve our confidence in the CAPM and Bayesian CAPM estimates (which we examined in Chapter 4). The CAPM related estimates result in far more diversified optimal portfolios than the optimal portfolios achieved when using sample means as the estimates for the securities' expected returns. This is evident from the results in Chapter 5 Sections 4.2 and 4.3. These diversified portfolios are certainly more attractive and practical to portfolio managers. This indicates that all factors relating to the mean-variance optimisation algorithm, including the models which produce expected returns for the securities, should be researched and improved if possible.

**CHAPTER 6**  
**CONCLUSION**

University of Cape Town

## 1. CONCLUSION

The Markowitz mean-variance portfolio selection model is one that receives much attention in journals. It receives both complements and criticisms. Many portfolio managers do believe that the Markowitz model does have the correct objective function viz. to maximise a portfolio's return while simultaneously minimising its risk. However, these same portfolio managers believe that the *traditional implementation* of the Markowitz model fails to provide realistic portfolios.

The foundation of the criticism is that the optimal portfolios are far from being *diversified*. When we implemented the "traditional" Markowitz model using our South African data set in Chapter 3 only four out of the seven indices were ever included in an optimal portfolio and one of these four indices was never more than 3% of the optimal portfolio.

We believe that acceptable and realistic portfolios can nevertheless still be obtained by the implementation of the Markowitz model. More *diversified* optimal solutions are obtained when tracking-error is used as the measure of risk rather than variance. In Chapter 3 when we used tracking-error as the measure of risk each of the assets in the South African opportunity set we considered was a member of at least one of the mean-tracking-error optimal portfolios. Tracking-error is justifiably used as a measure of risk since portfolio managers are concerned about the performance of the portfolio relative to a benchmark and tracking-error measures the aggregate deviations from this predefined benchmark.

One may still argue that from an investor's perspective variance might well be the best "theoretical" measure of risk. A compromise proposed in Chapter 3 and which, when using our data set, results in a *diversified* optimal portfolio is the mean-variance-tracking-error portfolio selection model. This model takes account of both *absolute* risk (measured by the portfolio's variance) and *relative* risk (measured by the portfolio's tracking-error). As when using the tracking-error as a measure of risk, each of the assets in our opportunity set is a member of an optimal portfolio achieved when applying the mean-variance-tracking-error portfolio selection model.

We paid much attention to measuring the risk of a portfolio and considered alternative measures of this risk. Simultaneously we believe that *realistic* portfolios will only be achieved when good forecasts for the securities' expected returns are used. In Chapter 4 we did not suggest some totally different estimate but we rather examined a compromise between the traditional sample mean estimate, which has been used in the past, and a subjective expectation. This subjective expectation could merely be stated or determined by some other model, for instance, the Capital Asset Pricing Model (CAPM). Such an estimate can be constructed by means of a Bayesian estimate. The empirical evidence in Chapter 4 revealed that when a Bayesian estimate is used then one certainly obtains a more diversified optimal portfolio, which makes it appealing to portfolio managers.

We have argued that the Markowitz Model itself captures the objectives of investors and portfolio managers but in order to achieve realistic optimal solutions the inputs of the Markowitz Model certainly need refinement.

## **BIBLIOGRAPHY**

University of Cape Town

Beninga, S, "Financial Modelling", 1997.

Berger, J, "Statistical Decision Theory and Bayesian Analysis," 1985.

Bowie, DC, and Bradfield, DJ, "Non-Normality and Robust Estimation of Beta Coefficients: Evidence from a Small Stock Market", this article is still to be submitted for publication.

Bradfield, DJ, "A Diagnostic Model for Improving the Efficiency of an Existing Portfolio", *European Journal of Operational Research*, 1994, pp. 284-293.

Find Journal  
Black, F, "Estimating Expected Return", *Financial Analysts Journal*, 1995, pp. 168-171.

Black, F and Litterman, R, "Global Asset Allocation with Equities, Bonds and Currencies", *Fixed Income Research*, October 1991.

Black, F and Litterman, R, "Global Portfolio Optimization", *Financial Analysts Journal*, September-October 1992, pp. 28-42.

Broemeling, L.D, "Bayesian Analysis of Linear Models", pp. 66, 1985.

Chatterjee, S and Jacques, WE, "An Outlier-Resistant Approach to Risk Estimation", *Financial Analysts Journal*, September-October 1994, pp. 69-75.

Chow, G, "Portfolio Selection Based on Return, Risk and Relative Performance", *Financial Analysts Journal*, March-April 1995, pp. 54-60.

Connor, G and Leland, H, "Cash Management and Index Tracking", *Financial Analysts Journal*, November-December 1995, pp. 75-80.

Cook, R.D, "Detection of Influential Observations in Linear Regression", *Technometrics*, pp. 15-18, 1977.

Divecha, MK, "Emerging Markets and Risk Reduction", *Global Asset Allocation*, 1994, pp. 340.

Fama, EF, "Efficient Capital Markets: II", *Journal of Finance*, December 1991.

Fischer, K and Statman, M, "The Mean-Variance Optimization Puzzle", *Financial Analysts Journal*, July/August 1997, pp. 41-50.

Franks, E.C, "Targeting Excess-of-Benchmark Returns", *The Journal of Portfolio Management*, Summer 1992, pp. 6-12.

Frost, PA and Savarino, JE, "An empirical Bayes approach to efficient portfolio selection", *Journal of Financial and Quantitative Analysis*, 1986, pp. 293-305.

Geweke, J, "Contemporary Bayesian Econometrics", 1996, pp. 2-17.

Green, R and Hollifield, B, "An Empirical Bayes Approach to Efficient Portfolio Selection", *Journal of Financial and Quantitative Analysis*, September 1992.

Harlow, W, "Asset Allocation in a Downside-Risk Framework", *Financial Analysts Journal*, September-October 1991, pp. 28-40.

Jeffreys, H, "Theory of Probability", 1939.

Jobson, JP and Korkie, B, "Estimation for Markowitz Efficient Portfolios", *Journal of the American Statistical Association*, September 1980, pp. 544-554.

Jorion, P, "Bayes-Stein Estimation for Portfolio Analysis", *Journal of Financial and Quantitative Analysis*, September 1986, pp.279-292.

Kandel, S and Stambaugh, R, "On the Predictability of Stock Returns: An Asset-Allocation Perspective", *The Journal of Finance*, June 1996, pp. 385-423.

Levy, H and Lerman, Z, "The Benefits of International Diversification in Bonds", *Financial Analysts Journal*, September-October, 1988.

Levy, M and Levy, L, "The Danger of Assuming Homogeneous Expectations", *Financial Analysts Journal*, May-June 1996, pp. 65-70.

Libby, R and Fishburn, P, "Behavioural Models of Risk Taking in Business Decisions: A Survey and Evaluation", *Journal of Accounting Research*, Autumn 1977.

Lindley, DV, "On a Measure of the Information Provided by an Experiment", *Annals of Mathematical Statistics* 22, 1956, pp. 986-1005.

Lintner, J, "The Validation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgeting", *Review of Economics and Statistics*, pp. 13-37, February 1965.

MacBeth, JD, "What's the Long-Term Expected Return to Your Portfolio?", *Financial Analysts Journal*, September-October 1995, pp.6-8.

Markowitz, HM, "Portfolio Selection: Efficient Diversification of Investment", 1959.

Markowitz, H.M, "Portfolio Selection", *Journal of Finance*, March 1952, pp. 77-91.

Michaud, R, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?", *Financial Analysts Journal*, January-February 1989, pp. 31-42.

Press, SJ, "Bayesian Statistics: Principles, Models and Applications", 1989, pp. 17, pp. 27, pp. 49.

Reichenstein, W and Rich, SP, "Predicting Long-Horizon Stock Returns: Evidence and Implications", *Financial Analysts Journal*, January-February 1994, pp. 73-76.

- Rice, JA, "Mathematical Statistics and Data Analysis", 1995, pp. 256, pp. 589.
- Roll, R, "A Mean/Variance Analysis of Tracking Error", *Journal of Portfolio Management*, Summer 1992, pp. 13-22.
- Rubenstein, M and Hayne, LE, "Replicating Options with Positions in Stocks and Cash", *Financial Analysts Journal*, July-August 1981, pp. 63-72.
- Shanken, J, "A Bayesian approach to testing portfolio efficiency", *Journal of Financial Economics*, 1987, pp. 195-215.
- Shanken, J, "Testing Portfolio Efficiency when the Zero-Beta Rate is Unknown: A Note", *The Journal of Finance*, March 1986, pp. 269-276.
- Sharpe, WF, "Portfolio Theory and Capital Markets", 1970.
- Sharpe, WF, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance*, September 1964, pp. 425-442.
- Sharpe, WF and Perold, AF, "Dynamic Strategies for Asset Allocation", *Financial Analysts Journal*, January-February 1988, pp. 16-27.
- Stein, C, "Inadmissibility of the Usual Estimate for the Mean of a Multivariate Distribution", *Proceedings of the 3<sup>rd</sup> Berkley Symposium on Probability and Statistics I*, 1955, pp. 197-206.
- Treynor's article has not been published.
- Trosky's paper has not been published. It will be published in 1998. I had private consultations with Prof. Trosky.
- Winkler, RL, "Introduction to Bayesian Inference and Decisions", 1972, pp. 144, pp. 169.