

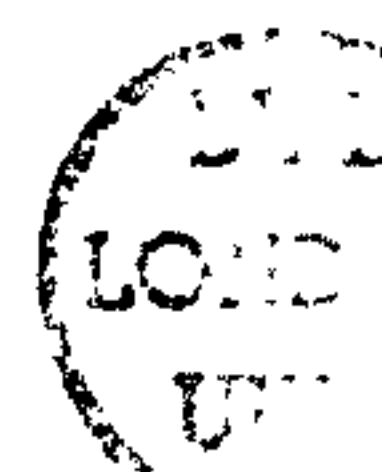
AN INTERIOR POINT ALGORITHM FOR MINIMAX AND ITS APPLICATION TO ROBUST
ASSET-LIABILITY MANAGEMENT

by

Enitan A. Obasanjo



Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Department of Computing
University of London
and Diploma of Imperial College,
2006.



To my parents and siblings, for their guidance, support and love, and in particular to my mother, Mrs. Oluremi Obasanjo, for her role as an inspiring force in my life.

PAGE

NUMBERING

AS ORIGINAL

Abstract

This PhD thesis is concerned with the development of computational tools and mathematical approaches for application to the study of computationally intensive portfolio optimisation problems - specifically, portfolio optimisation under uncertainty and with consideration to higher order moments. The thesis is structured as four sections. Part 1 provides an introduction to the theoretical framework and computational methods employed in the solution of the problems tackled in this thesis. Part 2 presents a primal-dual interior point method designed to exploit the structure of constrained nonlinear discrete minimax problems. The algorithm uses two merit functions and a switch between the merit functions to ensure progress towards a local optimum. Convergence properties are established and numerical results which indicate the algorithm's efficacy are provided. In Part 3, a multi-period asset liability management (ALM) portfolio problem is augmented with higher order moment measures, such as skewness and down-side risk measures. The minimax method is then applied to the ALM problems for robust optimisation with uncertain data. The minimax robust decision provides an optimal worst-case strategy which is guaranteed to improve if the worst-case does not occur. In Part 4, we discuss the results and highlight some interesting areas of future work. The numerical results from the application of the minimax methods for robust asset liability management, with consideration to higher order moments, are provided.

Alternative asset classes, such as derivatives and hedge funds, which exhibit non-normal return distributions, are becoming increasingly popular. The study and development of computational tools and mathematical approaches which cater for the uncertainty inherent in asset allocation models and deviations from the mean-variance framework is therefore of

importance, making this study a pertinent one.

Acknowledgements

This research was carried out at the Department of Computing, Imperial College London. Above all, I would like to gratefully acknowledge the enthusiastic supervision of Professor Berc Rustem, without whose steady encouragement the completion of this thesis could never have been accomplished. Colleagues in the Computational Management Science Research Group, past and present, are thanked for their stimulating discussions, assistance with research difficulties and general advice; in particular I would like to acknowledge the help of Dr. Ioannis Akrotirianakis, Dr. Nalan Gulpinar and Dr. Stan Zakovic. I am also grateful to Dr. Rodrigo Pinto, Dr. Tetsuya Noguchi, and thankful to Panos Parpas, George Tzallas-Regas and all my friends in the department for their support and friendship. Special thanks also go to the staff in the Department of Computing for their help with various technical difficulties and administrative matters.

I would also like to thank the staff at Palace Court for their care and attention in the many years I stayed there and worked with them.

Finally, I am forever indebted to my parents, siblings and Terence for their understanding, endless patience and encouragement when it was most required.

Table of Contents

I	Background	15
1	Introduction	16
1.1	The Optimal Decision-making Framework	16
1.1.1	Portfolio Theory and Optimisation	17
1.1.2	Higher Order Moments	19
1.2	General Nonlinear Programming Problems	22
1.3	Optimisation Under Uncertainty	24
1.3.1	The Minimax Robust Optimisation Methodology	26
1.3.2	The Minimax Interior Point Method	27
1.4	The Asset Liability Management Problem	28
1.4.1	Robust Asset Liability Management	29
1.5	Overview of Thesis	30
II	An Algorithm for Robust Optimisation	31
2	The Minimax Interior Point Algorithm	34
2.1	Description of the Algorithm	34
2.1.1	Description of the Minimax Problem	34
2.1.2	The Primal-Dual Method	35
2.1.3	The Newton System and Search Direction	38
2.1.4	Augmented Lagrangian Methods	39
2.1.5	The merit functions:	40
2.1.6	Penalty parameter selection rule	41
2.1.7	Line search strategies	42
2.1.8	Step-size rules:	43
2.1.9	Selection of barrier parameter:	45

3	Convergence Analysis and Numerical Experiments	48
3.1	Convergence Results	48
3.1.1	Convergence of the inner iterations	52
3.1.2	Convergence of the Outer Iterations	59
3.2	Numerical Results	64
III	ALM	72
4	The ALM Problem	76
4.1	The Asset Liability Management Problem Statement	76
4.2	Scenario Generation	79
4.3	Data and Summary Statistics	81
4.3.1	Empirical Analysis: Evidence of Non-Normality and Skewness . .	82
5	Robust Mean-Variance ALM	86
5.1	Introduction	86
5.2	The Mean-Variance ALM Model	86
5.3	Generation of Rival Scenario Trees	88
5.4	The Robust Mean-Variance ALM Model	89
5.5	Numerical Experiments	92
6	ALM with Higher Moments	96
6.1	Introduction	96
6.2	The Mean-Variance-Skewness (MVS) ALM Model	96
6.3	The Mean-Downside Risk (MDR) ALM Model	98
6.4	A Global Optimisation Heuristic	101
6.5	Numerical Experiments	102
6.5.1	The Mean-Variance-Skewness Relationship of the ALM Problem .	102
7	Robust ALM with Higher Moments	106
7.1	Introduction	106
7.2	The Robust Mean-Variance-Skewness ALM Model	106
7.3	Robust Mean-Downside Risk ALM Model	108
8	Simulation Results	115
8.1	Experimental Results	115

IV	Conclusions	121
9	Discussion	122
9.1	Summary	122
9.2	Future Work	127
9.3	Conclusion	129
V	Appendix	130

List of Figures

4.1	A two-period scenario tree, with 9 nodes and 6 scenarios	82
5.1	The minimax efficient frontier compared to the efficient frontiers computed using each of the three scenario trees.	93
5.2	Noninferiority of minimax: A clipping of the minimax frontier and the mean-risk profile of the minimax optimal portfolio holdings evaluated using the three scenario trees. Performance improves if the worst-case scenario is not realized.	94
5.3	The minimax efficient frontier compared to the mean-risk profiles obtained from portfolio holdings selected on the efficient frontiers for scenario trees 1, 2 and 3 when the worst-case tree is valid. Performance deteriorates for all other Scenario Trees if the worst-case scenario tree is valid	95
6.1	Mean-Variance Efficient frontier for the ALM problem. Guaranteed rate $g = 0\%$	104
6.2	Mean-Skewness “Efficient” frontier for the ALM problem. Guaranteed rate $g = 0\%$	104
6.3	Variance-Skewness “Efficient” frontier. Guaranteed rate $g = 0\%$	105
6.4	Mean-Skewness “Efficient” frontier for $g = 0\%$ and $g = 12\%$	105
7.1	Mean-Variance Efficient frontier for the ALM problem. Guaranteed rate $g = 12\%$	112
7.2	The normal and robust mean-skewness frontiers for the ALM problem. Guaranteed rate $g = 12\%$	112
7.3	The normal and robust variance-skewness “Efficient” frontier. Guaranteed rate $g = 12\%$	113
7.4	Mean-Variance-Skewness surface for $g=12\%$, $m = \text{mean}$, $v=\text{standard deviation}$, $s=\text{cubic root of skewness}$	114

8.1 Evaluation of the minimax mean-variance strategy over 11 years and comparison with mean-variance efficient frontier strategies for scenario trees 1,2 and 3.ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns 118

8.2 Evaluation of the minimax mean-variance-skewness strategy over 11 years and comparison with mean-variance-skewness efficient frontier strategies for scenario trees 1,2 and 3. ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns 119

8.3 Evaluation of the minimax mean-downside-risk strategy over 11 years and comparison with mean-downside-risk efficient frontier strategies for scenario trees 1,2 and 3. ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns . . . 120

8.4 Comparison of the results obtained from the application of the robust mean-variance, mean-variance-skewness and mean-downside-risk strategies in back-testing. 120

9.1 MV, MVS and MDR strategies with ScTree1 data (10 year annual returns) used in the back-testing. 123

9.2 MV, MVS and MDR strategies with ScTree2 data (10 year annual returns) used in the back-testing. 124

9.3 MV, MVS and MDR strategies with ScTree3 data (10 year annual returns) used in the back-testing. 124

9.4 Asset allocation of the minimax mean-variance strategy over 11 years . . . 125

9.5 Asset allocation of the minimax mean-variance-skewness strategy over 11 years 125

9.6 Asset allocation of the minimax mean-downside-risk strategy over 11 years 126

9.7 Example 1. 131

9.8 Example 2. 133

List of Tables

3.1	The Test Set	66
3.2	Numerical Results in comparison with LOQO and MINOS	68
3.3	Progress of the algorithm for chaconn1.	71
3.4	Progress of the algorithm for mifflin1.	71
4.1	Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1971 to 31 December 2003 (33 observations).	84
4.2	Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1971 to 31 December 1981 (11 observations).	84
4.3	Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1982 to 31 December 1992 (11 observations).	84
4.4	Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1993 to 31 December 2003 (11 observations).	84
4.5	Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1971-2003	85
4.6	Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1971-1981	85
4.7	Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1982-1992	85
4.8	Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1993-2003	85

9.1 Average iteration counts for the minimax algorithms on the back-testing problems, where MMX_1 is the minimax algorithm using the single merit function, and MMX_2 uses the switching merit function. 126

9.2 Realised returns for the five indices over the back-testing period. 126

9.3 Starting Points for the Test Problems. 134

Part I

Background

Chapter 1

Introduction

1.1 The Optimal Decision-making Framework

A formal incorporation of risk and uncertainty in decision-making theory was first accomplished by John Von Neumann and Oskar Morgenstern in [79], where they provided an axiomatic framework for the expected utility hypothesis in terms of agents' preferences over "lotteries", or ventures with random prospects. Let p be a simple probability measure on a set of outcomes X , and let x be an outcome in X . Then $p = (p(x_1), p(x_2), \dots, p(x_n))$ where $p(x_i)$ are probabilities of outcome $x_i \in X$ such that

$$p(x_i) \geq 0, \forall i = 1, \dots, n$$

and $\sum_{i=1}^n p(x_i) = 1$. Simple probability measures have finite support i.e., there are finite elements $x \in X$ for which $p(x) > 0$. We define $\Delta(X)$ as a set of simple probability measures on X . A particular lottery p is a point in $\Delta(X)$. The von Neumann-Morgenstern hypothesis assumes that probabilities are exogenous and cannot be influenced by the agent. The agent is faced with the problem of choosing the best lottery in $\Delta(X)$. A major contribution of von Neumann-Morgenstern's [79] was to show that if an agent has preferences defined over lotteries, then there is a utility function

$$U : \Delta(X) \rightarrow \mathbb{R}$$

that assigns a utility to every lottery $p \in \Delta(X)$ which represents these preferences. The von Morgenstern-Neumann theory therefore states that people exhibit preferences over lotteries or probability distributions rather than over the outcome.

The expected utility representation for simple probabilities derived from the von Neumann-Morgenstern theory can be expressed as:

$$U(p) = \sum_{x \in \text{Supp}(p)} p(x)u(x),$$

where $u : X \rightarrow R$ is an elementary utility function on the underlying outcomes X , and $\text{supp}(p)$ denotes the finite support of p .

1.1.1 Portfolio Theory and Optimisation

Portfolio theory is a quantitative approach to investing in which investments are described statistically. Asset returns are assumed to be random variables with an associated probability distribution or probability density function. The probability distribution for a (discrete) random variable is a list of probabilities associated with each of its possible values. The probability distribution for X gives the probability $p(x_i)$ that the random variable equals x_i for each value of $x_i \in X$. For continuous random variables, the probability density function is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval. The maximum expected return is sought for the investor's acceptable risk level. This process requires a framework for ranking uncertain outcomes. Modern portfolio theory has largely developed from the mean-variance framework that was formalised in the seminal work of Markowitz [51]. The mean-variance rule states that, given the assumption of jointly normal asset return distributions or an investor with a quadratic utility function, we can maximise the investor's expected utility simply by selecting assets that yield the optimal combinations of mean and variance of returns. Thus, given two portfolios, A and B, with expected returns $\bar{r}_A$ and $\bar{r}_B$ respectively, then A yields a higher utility than B if $\bar{r}_A \geq \bar{r}_B$ and $\sigma_A \leq \sigma_B$, where σ_A and σ_B are the variances of portfolio A and portfolio B respectively.

The underlying assumptions of the mean-variance framework lead to a loss of generality that causes it to be deficient in many instances. The mean-variance approach is inconsistent with second order stochastic dominance, and thus with the expected utility approach for portfolio selection, unless the asset return distributions are normal. A further discussion of stochastic dominance is provided in Section 1.1.2.

In addition, the classic Markowitz model is static, and does not explicitly consider the implications of being able to revise or re-balance the portfolio in the future. Over time, however, as the economic environment changes, the performance of a static mean-variance optimised portfolio may deteriorate. For long-term investment portfolios, such as held by insurance companies and pension funds, it is important to model the temporal nature of the decision-making problem.

One of the tools used to capture the temporal and uncertain nature of real-life decision making problems is the use of simulation methods or scenario generation techniques to produce realisations of time-varying stochastic variables. The advantage of multi-scenario analysis is that a range of scenarios can be evaluated based on the statistics available and/or the inclination of the investor as to the future direction of the uncertain factor. The information obtained from such analysis can then be used to protect portfolios from potentially negative changes or to make “optimal” decisions. Another shortcoming of the classic Markowitz model is the sensitivity it displays to small changes in the input, which can lead to unreasonably large variations in the optimised portfolio. In reality, traditional Mean-Variance optimisation (MVO) is only as good as the forecast or estimates used as model input, and the assumption that the inputs are the real values of the model parameters results in the portfolios defined by MVO being “error-maximised”.

Nonetheless, mean-variance methods remain the most widely used approach in optimal portfolio selection and asset allocation problems. The popularity of the mean-variance approach in portfolio selection and asset allocation problems has largely been due to its computational tractability. The relaxation of its assumptions was previously avoided due to computational limitations. However, with advances in technology, the adoption of more appropriate models for portfolio selection has become less impractical. Thus a flurry litera-

ture has been spawned in two main streams of research to extend portfolio models beyond the assumptions of the Markowitz model: portfolio optimisation with higher order moments and methods of robust portfolio selection.

1.1.2 Higher Order Moments

The application of a mean-variance variance framework is only truly correct when the asset return distributions are fully characterised by their first two moments, as is the case with normal distributions, or the investor has a quadratic utility function. There is substantial empirical evidence of asymmetry in the returns of a number of asset classes and of positive skewness preference in investors. Guidolin and Timmerman [33] find evidence of two distinct bull and bear states in the joint distribution of five major regions, and employ regime switching models which can capture outliers, fat-tails and skewness in asset allocation problems for an investor with a utility preference that incorporates skew and kurtosis preferences. They find significant differences in the optimal asset allocation suggested by their approach compared with the mean-variance approach. The relevance of higher moments, such as skewness, to an investor's decision making under uncertainty has been argued by many authors such as Samuelson [70], Arditti [3] and Rubinstein [62]. Evidence of non-normality of returns for individual securities and portfolios has also been found by several authors (e.g. Simkowitz and Beedles [73], Singleton and Wingender [74], Arditti and Levy [4]).

Higher order moments have been considered by augmenting portfolio problems with the skewness and/or kurtosis, e.g. Athayde [18] and Konno [46]. However, skewness is still frequently ignored in portfolio optimisation, largely due to computational difficulties associated with calculating the third moment. Also Chunchachinda *et al* (see [15]) and more recently Prakash *et al* (see [61]) and Sun and Yan (see [75]) have used the polynomial goal programming (PGP) approach to portfolio selection with skewness which was first described in Lai [49]. This approach, however, optimises skewness and return subject to unit variance. This collection of literature also ignores the non-convexity of the skewness function, which means the optimal portfolios are not necessarily globally optimal.

Stochastic Dominance and Lower Partial Moment Risk Measures

An alternative approach to mean-variance optimisation which accommodates higher order moments in portfolio optimisation is the selection of a risk measure characterised by a general set of utility functions and not constrained by a specific probability distribution, such as asymmetric risk measures. One such measure, proposed by Fishburn [22] is the n -degree Lower Partial Moment (LPM). Bawa [5] was the first to define lower partial moment as a general family of below target risk measures. Bawa ([5], [6]), Porter [60] and Fishburn [see Theorem 3][22], all demonstrate the strong relationship between LPM and stochastic dominance (SD). Stochastic dominance is an alternative framework to mean-variance efficiency for ranking uncertain outcomes. The basic stochastic dominance approach is to resolve risky choices under the weakest possible assumptions. SD assumes that individuals are expected utility maximisers and further assumptions are made regarding relative preferences for wealth and risk aversion. The stochastic dominance selection rules are:

- **First Order Stochastic Dominance (FSD):** is the selection rule for the entire class of utility function where the utility function is an increasing function of returns. The FSD selection rule is that a probability distribution F dominates a probability distribution G if and only if F never lies above and somewhere lies below G . The FSD can only order a small subset of the class of all distribution functions. It is the weakest assumption as it only assumes monotonicity i.e. investors prefer more money to less money and are non-satiated.
- **Second Order Stochastic Dominance (SSD):** is the selection rule for the class of increasing and risk-averse utility functions. That is, utility functions that are increasing with decreasing marginal utility everywhere. This selection rule is that a probability distribution F dominates G if and only if the integral of F never lies above and lies somewhat below the integral of G . The SSD rule orders a larger subclass of distributions than the FSD selection rule. SSD adds risk aversion to the non-satiety assumption of FSD. Expected utility is now smaller or equal to the utility of expected returns so concave utility functions are now assumed.

- Third Order Stochastic Dominance (TSD): is the selection rule for the class of increasing and risk-averse utility functions with positive third derivatives when the random variables are defined on a finite range. A positive third derivative is implied by decreasing absolute risk aversion. Decreasing absolute risk aversion (DARA) is generally considered to be the most restrictive class of utility functions acceptable to economists [5]. TSD assumes investors dislike negative skewness and like positive skewness as the third derivative of the utility is always positive. So, although investors may be risk-averse, they still have positive skewness preference. This is equivalent to an aversion to down-side risk.

Fishburn particularly demonstrates the relationship between the n -degree LPM and First, Second and Third Order Stochastic Dominance. The drawback of the stochastic dominance concept compared with mean-variance framework is that we do not yet have an algorithm to find SD efficient diversification strategies as we do for MV efficient strategies. The relationship between LPM and stochastic dominance implies LPM is consistent with expected utility theory (EUT). The LPM is thus offered as a form of utility based portfolio analysis without the restrictions and assumptions required by traditional mean-variance methods. Fishburn [22] shows that by varying the degree, n , of the n -degree LPM, the risk measure can accurately reflect investor utility towards risk and below target returns. Fishburn also proves that $n = 1$ separates risk-loving and risk-averse behaviour with respect to below target returns. He also argues that there is no compelling a priori reason for taking $n = 2$, and proves that different values of n can approximate a wide variety of attitudes towards the risk of falling below the target return, h .

In general, the investor's risk aversion is an increasing function of n . Since investors prefer positive skewness and dislike negative skewness, the LPM becomes a measure of risk, measuring the amount of negative skewness in the distribution. The higher the LPM value, the greater the negative skewness and the greater the risk of the investment. As n increases, the negative skewness of the security or portfolio remains the same, but the LPM (n, h) applies a heavier utility penalty (through a higher computed value of LPM) to the negative skewness. In Nawrocki [56], empirical results have found that as n increases the

skewness of the portfolio increases. Comparisons of LPM portfolio selection with mean-variance portfolio selection also indicated that LPM analysis provides for a greater number of investor utility functions. Since Fishburn [22] proves TSD for n -degree LPM with $n \geq 2$, this is consistent with an investor with DARA, i.e. a risk averse investor who exhibits a demand for a risky asset that increases with wealth. Fishburn also fits the LPM function to the utility curves for four middle managers in a large chemical company, presented by Green [32], and obtained estimates of n ranging from 2.3 to 4.6. In [5], Bawa notes that, for the class of DARA utility functions, an advantage of the mean-LPM $((n, h))$ over mean-variance and mean-variance-skewness models on theoretical grounds is that the latter two use neither necessary nor sufficient conditions for dominance. Bawa suggests that the mean-LPM (n, h) rule may be used as a reasonable approximation to the TSD and hence optimal selection rule for the entire class of distribution functions because it uses a necessary condition for dominance.

1.2 General Nonlinear Programming Problems

General constrained nonlinear programming problems can be expressed in the form:

$$\min_{x \in R^n} \{f(x) | g(x) = 0, x \geq 0.\} \quad (1.1)$$

where $x \in R^n$, and $f : R^n \rightarrow R, g : R^n \rightarrow R^q$ are smooth, twice differentiable functions.

The Lagrangian function associated with this problem is

$$L(x, \lambda, \mu) = f(x) + \lambda' g(x) - \mu' x,$$

where λ and μ are the Lagrange multipliers for the equality and non-negativity constraints respectively. The Karush-Kuhn Tucker (KKT) [48] optimality conditions for prob-

lem 1.1 are:

$$\begin{aligned}
 \nabla_x f(x) + \lambda' \nabla g(x) - \mu' x &= 0, \\
 g(x) &= 0, \\
 x &\geq 0, \\
 \mu &\geq 0.
 \end{aligned}
 \tag{1.2}$$

This defines a nonlinear system of equations. In interior point methods, the non-negativity constraints can be incorporated into the objective function with a logarithmic barrier term, transforming 1.1 above into

$$\min_{x \in R^n} \{ f(x) - \mu \sum_{i=1}^n \ln x_i \mid g(x) = 0. \}$$

The Lagrangian function for this problem is then given by:

$$L = f(x) - \mu \sum_{i=1}^n \ln x_i + \lambda' g(x).$$

The optimal solution to the problem satisfies the KKT conditions:

$$\begin{aligned}
 \nabla_x f(x) + \lambda' \nabla g(x) - \mu X^{-1} \mathbf{1} &= 0, \\
 g(x) &= 0, \\
 x &\geq 0, \\
 \mu &\geq 0.
 \end{aligned}
 \tag{1.3}$$

where $X = \text{diag}(x)$.

Algorithms for numerical optimisation generally proceed by solving for the equality conditions in the systems of equations described by (1.2) and (1.3) to obtain a search direction d_k , usually by Newton's method or a variant, and iteratively generating a sequence of points x_k , such that $x_{k+1} = x_k + \tau d_k$, where $\tau \in (0, 1]$.

1.3 Optimisation Under Uncertainty

Real-world optimisation problems are often characterised by uncertain or inexact data, and small perturbations in the data can sometimes render an “optimal” solution infeasible. The development of methodologies for optimisation with this type of data, known as robust optimisation, is a growing area of research. The data uncertainty can take many forms. One form of uncertainty is regarding problem data which represent the future, and cannot be known with certainty. This uncertainty is usually incorporated by modelling the possible evolution of the data in the optimisation problem, an approach known as stochastic optimisation. Stochastic optimisation seeks to find solutions that are robust with respect to all possible values of the problem parameters. Stochastic optimisation models are often solved by stochastic programming methods.

Another common form of uncertainty often dealt with in finance literature on robust optimisation refers to the choice of stochastic model used to describe the evolution of the input parameters for the optimisation problem. This form of uncertainty is often referred to as model or distributional uncertainty. In addition, the parameters of the stochastic models are often assumed to be known, but these are usually either estimated from available data or from subjective information, thereby bearing some estimation risk. This source of uncertainty is often referred to as parametric uncertainty.

Data uncertainty in optimisation has been addressed in several ways in the existing literature. Chopra [14] and Frost and Savarino [25] add constraints to the portfolio weights to reduce the sensitivity of the Markowitz model to perturbations of the input. Michaud [54] suggests repeatedly drawing from the means and variances/covariances of the probability distributions and solving the Markowitz problem for each random draw, then aggregating the resulting portfolios. Most other approaches involve either the use of robust estimators

such as shrinkage, stein or Bayesian estimators, or minimax/maximin methodology. Goldfarb and Iyengar [30] tackle estimation risk in mean-variance portfolio selection by introducing uncertainty structures corresponding to confidence regions for the least squares estimates of the market parameters and then selecting the optimal portfolio for the worst-case after reformulating the problem as a second-order cone programming problem. Tütüncü and Koenig [76] apply a similar approach of generating uncertainty structures to optimal asset allocation models. Victoria-Feser [78] addresses parametric uncertainty by using robust estimators of the mean and variance which are shown to be more stable than the classic MVO estimates. Both of these approaches tackle single period optimisation problems. Robust counterparts of multi-period portfolio problems with ellipsoidal uncertainty sets were presented by Ben-Tal *et. al* in [7], and compared with a stochastic programming approach. They also studied the effect of slight perturbation in the input data on a set of linear programming problems from the NETLIB library and found that the feasibility properties of their usual solutions could be severely affected. Victoria-Feser [78] addresses parametric uncertainty by using robust estimators of the mean and variance which are shown to be more stable than the classic MVO estimates. Both of these approaches tackle single period optimisation problems. Harvey *et al* [37] have tackled both higher order moments and robustness by using a Bayesian approach. They compare their results to the efficient frontier resampling approach to robust optimisation proposed by Michaud [54] which defines a confidence region and hence variability of the efficient frontier. More recently, Garlappi, Uppal and Wang [26] applied a minimax approach to robust portfolio optimisation under parametric and model uncertainty using a minimax approach. They also compared the performance of the minimax method against the use of two Bayesian estimators, and the minimax approach was found to provide better out-of-sample results. In addition, they noted the added benefit that the minimax approach is flexible enough to be combined with various types of estimators, such as Bayesian estimators, as well as the standard least-squares or maximum-likelihood estimators. However, their study was done under the assumption of normal asset return distributions. Similarly, robust optimal control methods have been developed. For example, model mis-specification has been dealt with for multi-period con-

tinuous time Merton models by maximin expected utility optimisation in a neighbourhood of a base or reference model [Gilboa and Schmedler [28], Hansen and Sargent [36]].

So far, robust optimisation methodologies have been applied to linear, conic, quadratic and semi-definite optimisation problems. Here, we extend the current body of research by applying a robust optimisation methodology, namely minimax, to nonlinear and not necessarily convex optimisation problems.

1.3.1 The Minimax Robust Optimisation Methodology

The minimax approach is an important principle in optimal parameter selection.

In the minimax robust framework, parameters are assumed to belong to a specified uncertainty set.

The general form of the minimax problem is:

$$\max_{y \in Y} f(x^*, y) = \min_{x \in X} \max_{y \in Y} f(x, y), \quad (1.4)$$

If the set Y is a discrete set, the problem becomes a discrete minimax problem of the form:

$$\max_{y \in Y} f(x^*, y) = \min_{x \in X} \max_{i=1, \dots, m} f^i(x), \quad (1.5)$$

Consider the mathematical programming formulation of the mean-variance optimisation problem below:

$$\min_{x \in R^n} \{x' \Sigma x | r' x = m\} \quad (1.6)$$

The discrete minimax formulation of (1.6) where r and Σ are uncertain is:

$$\min_{v \in R^1, x \in R^n} \{ \alpha v - (1 - \alpha)m \mid v \geq x' \Sigma_i x, r'_j x = m, i = 1, \dots, N, j = 1, \dots, M \} \quad (1.7)$$

where $\alpha \geq 0$.

Minimax has also been used as an l_∞ risk measure for a conservative investor in contrast to variance, e.g. Young [82] and Cai *et al* [10]. Empirical studies of a minimax portfolio selection rule in comparison with a Markowitz mean-variance selection rule have also been reported e.g. Cai *et al* [11].

1.3.2 The Minimax Interior Point Method

A number of interior point algorithms for linear and nonlinear programming have been developed in the years following the introduction of Karmarkar's method in [39]. El-Bakry *et al* provided the formulation and theory of the Newton interior point method for nonlinear programming in [21], extending the general primal dual framework proposed for LP problems by Kojima, Mizuno and Yoshise [45]. Their algorithm used the l_2 norm of the residuals as merit function in the line-search procedure. This choice of merit function, however may lead the algorithm to converge to a saddle point or maximum when a minimum is being sought. Facilitated by advances in computing, the development of interior-point methods for large-scale nonlinear problems has been of great interest over the last decade. In light of the success of some interior-point methods to this class of problems (see e.g. [9]), this research aimed to develop an efficient and robust algorithm for large-scale, general nonlinear minimax problems using a primal-dual interior-point method. The resulting algorithm is presented in this thesis, and applied to an asset liability management model.

The algorithm is based on the primal-dual interior point method described in [2] and the minimax approach of [65], but differs mainly in the choice of merit function, step-size rule and computation of search direction. The nonlinear optimisation problems considered are assumed to be smooth and twice differentiable. Newton's method is used to solve

the primal-dual system of the problem, generating descent directions for the merit function. An adaptive penalty parameter is employed to ensure progress towards optimality and feasibility. Two merit functions are used to ensure convergence towards a solution of the optimisation problem.

The algorithm is designed to be used as a robust solution for optimisation under uncertainty for problems with a large number of variables and constraints. The algorithm code is linked with the mathematical programming language, AMPL, which provides first and second derivative information, but the approach is applicable to quasi-Newton methods. Global and local convergence results are established for the algorithm. The performance of the method is illustrated through some numerical examples, and is compared with the results obtained using MINOS and LOQO on the NEOS server 4.0 on problems drawn from the CUTE test set.

1.4 The Asset Liability Management Problem

The inclusion of liabilities in Markowitz optimisation, known as Asset Liability Management (ALM), has also been an area of recent research activity. ALM is intended to earn an adequate return while maintaining a comfortable surplus of assets beyond liabilities. Carino *et al.* [12] applied a stochastic optimisation approach to the Russell-Yasuda Kasai ALM problem, resulting in a strategy that yielded an extra income of \$79 million during the first two years of employing the model. Consigli and Dempster [16] presented the CALM model, where they sought the maximisation of the terminal wealth of a general ALM portfolio, over a long-term horizon with re-balancing at discrete-time periods, with uncertainty in the form of pension payments or borrowing costs. In Gondzio [31], the ALM model of a Dutch pension fund is considered using a stochastic programming approach. Zenios *et al.* [20] developed a model to maximise the expected utility of terminal wealth of a portfolio of mortgage-backed securities (MBS) funding a single known liability under interest-rate uncertainty, for an investor with a logarithmic utility function. Nielsen and Zenios [57] extended this approach to a model for funding Single Premium Deferred Annuities (SPDAs). Ziemba and Mulvey [83] provided a discussion of the factors to consider and steps to take

in the modelling of an ALM problem.

1.4.1 Robust Asset Liability Management

For a financial institution offering a product with a guaranteed rate of return, there is an added risk to the issuer of not being able to meet the target return engendered by the guarantee. The application of robust optimisation methods to higher-order moments in ALM will therefore

- provide a product issuer with an assurance of being able to meet its liabilities in the event of sizeable, possibly nonlinear perturbations of underlying assumptions,
- provide a more accurate picture of the risk-return trade off for a guaranteed investment portfolio, given evidence of non-normalities, thereby ensuring more appropriate asset allocation strategies are devised,
- ensure the use of appropriate risk-measures without unnecessary problem complexity.

The minimax optimisation strategy provides an optimal investment policy based on the worst-case outcomes of the uncertain factors. The minimax approach is referred to as robust, because it yields a guaranteed minimum performance of the objective, in relation to the scenarios specified. A number of multi-period stochastic ALM models have been implemented. The approaches for solving these problems have been either stochastic control or stochastic programming. The stochastic programming approach has the advantage of allowing for the temporal nature of the problem, with multiple decision stages where the portfolio can be re-structured as uncertainty is resolved. Multistage stochastic programming models of ALM problems have been successfully applied.

In this research, we adopt a minimax approach on a scenario-based multi-period model for robust asset liability management. The portfolio is required to fund a stream of stochastic liabilities over a given period. Our approach is concerned with the robustness of the portfolio strategy when there is model misspecification and parametric uncertainty. To be precise, we seek optimal portfolio strategies under consideration of the worst-case model

and parametric uncertainty. The robustness to parameter uncertainty is achieved by minimax optimisation over a set of parameter estimates generated from the historical data. The robust optimisation approach here is more general than the minimax methods described in [30]. No assumptions are made about the geometry of the uncertainty, we simply generate uncertainty sets from the historical data. The model robustness arises from the non-inferiority of the worst-case optimal minimax strategy. This is discussed further in [69]. The minimax robustness approach is suitable for situations which require protection against the risk of adopting the investment strategy based on the wrong data, as may be the case for pension funds, and insurance companies. In this study, we consider worst-case analysis for robust asset liability management of a guaranteed investment product.

1.5 Overview of Thesis

The rest of the thesis is organised as follows: Chapter 2 describes the interior point algorithm for minimax. In Chapter 3, convergence analyses of the algorithm is provided along with numerical results and comparison with LOQO and MINOS. Chapter 4 gives the background of the ALM problem. In Chapter 5, the robust mean-variance ALM model is described. In Chapter 6, we present the extension of the ALM models to incorporate higher order moments and in Chapter 7, we apply the robust optimisation approach to the ALM models for higher moments. In Chapter 8, we simulate the application of the strategies on the ALM problem over a period of 11 years using the MSCI indices for five regions as our asset universe and provide the results with comparisons of the strategies. Finally, in Chapter 9, we summarise our conclusions and discuss future work to develop the research presented in this thesis further.

Part II

An Algorithm for Robust Optimisation

Notation and Convention for the Algorithm

- $v \in R^1$ scalar decision variable for the minimax objective function.
- $x \in R^n$ vector of the decision variables.
- $s \in R^m$ vector of slack variables.
- $\alpha \in R^m$ vector of the pooling variable for the minimax objective function.
- $\bar{x}$ vector of decision and slack variables.
- μ barrier parameter for the inequality constraints.
- λ multiplier vector for the equality constraints.
- w vector of all the variables (primal and dual) of the optimisation problem.
- w_p vector of the primal variables.
- w_d vector of the dual variables.
- $f(x)$ scalar objective function of $x \in R^n$.
- $\nabla_x f(x)$ or $\nabla f(x)$ gradient of f with respect to x .
- $g(x)$ vector valued function of equality constraints.
- $\nabla_x g(x)$ or $\nabla g(x)$ gradient of g with respect to x .
- $diag(x)$ the diagonal matrix with elements $\bar{x}^i$.
- $G(v, x, s)$ vector valued function of equality constraints formed from $g(x)$ and the minimax objective constraints.
- k inner iteration counter.
- l outer iteration councr.
- τ_{pk} step-length for the primal variables.

- τ_{dk} step-length for the dual variables.
- $L(.)$ Lagrangian function for the constrained optimisation problem with variables $(.)$.
- $L^a(.)$ augmented Lagrangian function for the constrained optimisation problem with variables $(.)$.
- c_k penalty parameter value at x_k
- μ^l barrier parameter value at iteration l .
- Q the Hessian or positive semi-definite approximation of the Hessian of the Lagrangian function.

Chapter 2

The Minimax Interior Point Algorithm

2.1 Description of the Algorithm

2.1.1 Description of the Minimax Problem

We consider the following equality constrained minimax problem:

$$\min_{x \in R^n} \max_{j \in 1, \dots, m} f^j(x) \text{ subject to } g(x) = 0, x \geq 0. \quad (2.1)$$

It has been shown in [65], from the original proof by Medanic and Andjelic in [52] and [53], that (2.1) is equivalent to the following min-max formulation with pooled objectives to form

$$\min_{x \in R^n} \max_{\alpha \in \Xi_+^m} \{\langle \alpha, f(x) \rangle \mid g(x) = 0, x \geq 0\}, \quad (2.2)$$

where $\alpha \in \Xi_+^m = \{\alpha \in R^m \mid \alpha \geq 0, \langle 1, \alpha \rangle = 1\}$, $1 \in R^m$ is a vector of all ones, and $f(x) = [f_1(x), \dots, f_m(x)]'$. This follows from the fact that the maximum of m numbers is equal to the maximum of their convex combination.

By introducing an additional variable, v , the discrete minimax problem can be repre-

sented as the nonlinear programme (NLP):

$$\min_{x \in R^n, v \in R^1} \{v | f(x) \leq 1v, g(x) = 0, x \geq 0\}. \quad (2.3)$$

To enable comparison with other algorithms for nonlinear programming, the minimax problems discussed in this paper are implemented using the NLP formulation above. However, (2.2) and (2.3) are used to represent (2.1) whenever convenient in the subsequent analysis.

2.1.2 The Primal-Dual Method

The use of slack variables, as was done in [77], yields the reformulation of minimax problem (2.3) as

$$\min_{x \in R^n, v \in R^1} \{v | f(x) - 1v + s = 0, g(x) = 0, x \geq 0, s \geq 0, \} \quad (2.4)$$

where $s \in R^m$, $f(x)$ and $g(x)$ are smooth, twice differentiable, non-linear vector functions, with $f : R^n \rightarrow R^m$, $g : R^n \rightarrow R^q$. The equality constraints can be combined to form a new constraint vector:

$$G(v, x, s) = \begin{pmatrix} f(x) + s - 1v \\ g(x) \end{pmatrix}. \quad (2.5)$$

For notational compactness, the decision variables x and slack variables s can be combined into a single variable:

$$\bar{x} = \begin{pmatrix} x \\ s \end{pmatrix}, \quad (2.6)$$

where $\bar{x} \in R^N$ and $N = n + m$.

Thus we can re-write equation 2.4 as:

$$\min_{\bar{x} \in R^N, v \in R^1} \{v | G(v, \bar{x}) = 0, \bar{x} \geq 0\} \quad (2.7)$$

The inequality constraints in (2.7) above can then be incorporated into a barrier term

$$B(\bar{x}) = \mu \left(\sum_{i=1}^N \ln \bar{x}^i \right),$$

leading to the problem:

$$\min_{\bar{x} \in \mathbb{R}^N, v \in \mathbb{R}^1} \{v - \mu \left(\sum_{i=1}^N \ln \bar{x}^i \right) | G(v, \bar{x}) = 0, \bar{x} \geq 0\} \quad (2.8)$$

with barrier parameter $\mu \geq 0$.

The Lagrangian function associated with the constrained optimisation problem in (2.8) is given by

$$L(v, \bar{x}, \lambda, \mu) = v + \langle \lambda, G(v, \bar{x}) \rangle - \mu \left(\sum_{i=1}^N \ln \bar{x}^i \right), \quad (2.9)$$

where $\lambda \in \mathbb{R}^p$, $p = q + m$, are the Lagrange multipliers of the equality constraints.

The Karush-Kuhn-Tucker (KKT) optimality conditions for this problem are

$$\begin{aligned} \nabla_v L(v, \bar{x}, \lambda, \mu) &= 1 + \langle \lambda, \nabla_v G(v, \bar{x}) \rangle = 0, \\ \nabla_{\bar{x}} L(v, \bar{x}, \lambda, \mu) &= \langle \lambda, \nabla_{\bar{x}} G(v, \bar{x}) \rangle - \mu \bar{X}^{-1} \mathbf{1} = 0, \\ \nabla_{\lambda} L(v, \bar{x}, \lambda, \mu) &= G(v, \bar{x}) = 0, \end{aligned} \quad (2.10)$$

where $\bar{X} = \text{diag}(\bar{x})$, is the diagonal matrix with elements $\bar{x}^i$, and $\nabla G(v, \bar{x})$ is the Jacobian matrix of $G(v, \bar{x})$.

By introducing the nonlinear transformation $z = \mu \bar{X}^{-1} \mathbf{1}$ we modify (2.10) to produce the standard primal-dual system:

$$E(v, \bar{x}, z, \lambda; \mu) = \begin{Bmatrix} 1 + \langle \lambda, \nabla_v G(v, \bar{x}) \rangle \\ \nabla_{\bar{x}} G(v, \bar{x}) \lambda - z \\ G(v, \bar{x}) \\ \bar{X} Z \mathbf{1} - \mu \mathbf{1} \end{Bmatrix} = 0, \quad (2.11)$$

where $Z = \text{diag}(z)$, with $\bar{x}, z > 0$.

Let us now consider the Lagrangian of 2.4:

$$L(v, \bar{x}, \lambda) = v + \langle \lambda, G(v, \bar{x}) \rangle - \langle \mu \bar{x}, \quad (2.12)$$

where $\lambda \in R^p$ and $\mu \in R^N$ are the Lagrange multipliers of the equality constraints, $G(x) = 0$, and non-negativity constraints, $\bar{x} \geq 0$, respectively. The KKT system of 2.12 is given by:

$$E(v, \bar{x}, z, \lambda) = \begin{Bmatrix} 1 + \langle \lambda, \nabla_v G(v, \bar{x}) \rangle \\ \nabla_{\bar{x}} G(v, \bar{x}) \lambda - z \\ G(v, \bar{x}) \\ \bar{x} Z 1 \end{Bmatrix} = 0, \quad (2.13)$$

with $\bar{x}, z \geq 0$. Equation 2.11 is an approximation to the KKT system of problem 2.4, which increases in accuracy as $\mu \rightarrow 0$. Thus, the parameter $\mu \rightarrow 0$ defines a central path, C , to the solution point through a sequence of outer iterations, l . The outer iteration consists of a decrease in the barrier parameter μ . For a fixed value of μ , a point $(v(\mu), \bar{x}(\mu), z(\mu), \lambda(\mu))$ which solves the perturbed KKT conditions 2.11 belongs to this central path.

The design of the algorithm used to solve problem (2.8) is based on the following assumptions:

- (A1) Smoothness: $G(v, \bar{x})$ is a smooth twice continuously differentiable function of v and $\bar{x}$.
- (A2) Regularity : The columns of the matrix $[\nabla G(v, \bar{x})]$ corresponding to the binding constraints are linearly independent.
- (A3) Complementarity: At a solution point, $w_* = (v_*, \bar{x}_*, z_*, \lambda_*)$, strict complementarity is satisfied, *i.e.* if $\bar{x}_*^i > 0$ then $z_*^i = 0$ and if $z_*^i > 0$, $\bar{x}_*^i = 0$ for $i = 1, 2, \dots, N$.
- (A4) Second order sufficiency: The second order sufficiency condition is satisfied at a solution point.

$$i = 1, 2, \dots, q, \text{ and } 1' y = 0, i \in I_{\bar{x}}^0 \text{ then } y' \nabla_{\bar{x}\bar{x}} L(v, \bar{x}, \lambda, \mu) y > 0.$$

- (A5) Feasibility: The feasible region is bounded *.

2.1.3 The Newton System and Search Direction

Primal-dual methods use the Newton or quasi-Newton method to solve approximately the perturbed KKT conditions for a fixed value of μ . The Newton system for (2.11) above is:

$$\begin{bmatrix} 0 & 0 & 0 & -1' \\ 0 & \nabla_{\bar{x}\bar{x}}^2 L(v, \bar{x}, \lambda, \mu) & -I & \nabla_{\bar{x}} G'(v, \bar{x}) \\ 0 & Z & \bar{X} & 0 \\ -1 & \nabla_{\bar{x}} G(v, \bar{x}) & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta v \\ \Delta \bar{x} \\ \Delta z \\ \Delta \lambda \end{bmatrix} = -E(v, \bar{x}, z, \lambda; \mu), \quad (2.14)$$

since

$$\nabla_v^2 L(v, \bar{x}, \lambda, \mu) = 0,$$

and

$$\nabla_{v\bar{x}} L(v, \bar{x}, \lambda, \mu) = \nabla_{\bar{x}v} L(v, \bar{x}, \lambda, \mu) = 0,$$

Let $w = (v, \bar{x}, \lambda, z)'$ represent the variables and $w_p = (v, \bar{x})'$ the primal variables of problem (2.7) respectively. The system (2.14) is then solved to give the search directions $\Delta w = (\Delta v, \Delta \bar{x}, \Delta z, \Delta \lambda)'$, and the algorithm proceeds iteratively from an initial point $w_0 = (v_0, \bar{x}_0, z_0, \lambda_0)'$ through a sequence of inner iterations k such that $w_{k+1} = w_k + \tau_k \Delta w_k$, where

$$\tau_k = \text{diag}(\tau_{pk} I_v, \tau_{pk} I_{\bar{x}}, \tau_{pk} I_z, \tau_{dk} I_\lambda),$$

with τ_{pk} being the primal step-size and τ_{dk} the dual step-size. The inner iterations seek to minimise a merit function, which combines optimality and feasibility of the problem into a single function. The choice of merit functions is discussed further below.

Using matrix-vector notation, (2.14) can be expressed as:

*This may be enforced, if required, in the usual way of placing upper and lower bounds on $\bar{x}$.

$$\nabla E \Delta w = -E(w). \quad (2.15)$$

Thus, the Newton system (2.14) can be re-written as:

$$-\langle 1, \Delta \lambda \rangle = -[1 - \langle \lambda, 1 \rangle] \quad (2.16)$$

$$Q \Delta \bar{x} + \langle \nabla_{\bar{x}} G(v, \bar{x}), \Delta \lambda \rangle - \Delta z = -[\nabla_{\bar{x}} G'(v, \bar{x}) \lambda - z] \quad (2.17)$$

$$Z \Delta \bar{x} + \bar{X} \Delta z = -(\bar{X} Z 1 - \mu 1) \quad (2.18)$$

$$-1 \Delta v + \nabla_{\bar{x}} G'(v, \bar{x}) \Delta \bar{x} = -G(v, \bar{x}) \quad (2.19)$$

where Q is the Hessian of the Lagrangian function, $L(v, \bar{x}, \lambda, \mu)$, or a positive definite approximation of the Hessian of the Lagrangian function.

We note that, from the definition of w_p , (2.19) can be expressed as:

$$\nabla_{w_p} G' \Delta w_p = -G(w_p) \quad (2.20)$$

where

$$\nabla_{w_p} G = \begin{bmatrix} \nabla_v G(v, \bar{x}) & \nabla_{\bar{x}} G(v, \bar{x}) \end{bmatrix}. \quad (2.21)$$

and $\Delta w_p = (\Delta v, \Delta \bar{x})'$.

2.1.4 Augmented Lagrangian Methods

By adding a quadratic penalty term to the Lagrangian function, we relax the convexity assumptions required to define a saddle point of the Lagrangian function as described in [64], and obtain the augmented Lagrangian function:

$$\begin{aligned} L^a(v, \bar{x}, \lambda; c, \mu) &= v + \frac{c}{2} \|G(v, \bar{x})\|_2^2 + \langle \lambda, G(v, \bar{x}) \rangle \\ &\quad - \mu \sum_{i=1}^N \ln \bar{x}^i \end{aligned} \quad (2.22)$$

where μ is the barrier parameter.

This function can be considered as the Lagrangian of problem

$$\min_{\bar{x} \in R^N, v \in R^1} \{v + \frac{c}{2} \|G(v, \bar{x})\|_2^2 - \mu (\sum_{i=1}^N \ln \bar{x}^i) | G(v, \bar{x}) = 0\}. \quad (2.23)$$

The KKT optimality conditions for (2.23) are the system of nonlinear equations:

$$\begin{aligned} \nabla_v L^a : 1 + c \nabla_v G'(v, \bar{x}) G(v, \bar{x}) + \langle \lambda, \nabla_v G(v, \bar{x}) \rangle &= 0, \\ \nabla_{\bar{x}} L^a : c \nabla_{\bar{x}} G'(v, \bar{x}) G(v, \bar{x}) + \nabla_{\bar{x}} G(v, \bar{x}) \lambda - z &= 0, \\ \nabla_{\lambda} L^a : G(v, \bar{x}) &= 0. \end{aligned} \quad (2.24)$$

It can be seen that a point $(v, \bar{x}, z, \lambda)$ which satisfies the KKT conditions (2.24) also satisfy the KKT conditions of (2.10) by substituting $\nabla_{\lambda} L^a$ into $\nabla_v L^a, \nabla_{\bar{x}} L^a$ above. Thus, (2.23) has the same minima as (2.8).

2.1.5 The merit functions:

The constraint vector at iteration k is defined as

$$G_k = G(v_k, \bar{x}_k).$$

Let $\epsilon_G > 0$ be a finite precision to which the equality constraints are satisfied, *i.e.* we have a worst-case feasibility precision

$$\|G_k\|_2^2 > \epsilon_G. \quad (2.25)$$

While (2.25) holds, the inner iterations, k , use the following merit function

$$Merit_1 : F(v_k, \bar{x}_k, c_k, \mu) = v_k + \frac{c_k}{2} \|G_k\|_2^2 - \mu \sum_{i=1}^n \ln \bar{x}_k^i \quad (2.26)$$

For any inner iteration, k , where $0 < \|G(v_k, \bar{x}_k)\|_2^2 \leq \epsilon_G$, the algorithm switches to the second

merit function,

$$Merit_2 : \|E(w_k; 0)\|_2^2, \quad (2.27)$$

for all consecutive inner iterations, where E is as defined in (2.11) with $\mu = 0$. Once the convergence of the inner iterations is achieved, the algorithm switches back to the original merit function.

2.1.6 Penalty parameter selection rule

The penalty parameter selection strategy is an important part of the algorithm. At each iteration for which (2.25) holds, a value of c is determined such that the descent property

$$\Delta v_k - \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1} - c_k \|G(v_k, \bar{x}_k)\|_2^2 + \frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq 0, \quad (2.28)$$

holds.

The above expression ensures the monotonic descent of $Merit_1$ function as discussed in *Theorem 1* below.

Since μ is fixed throughout the inner iterations, it can be deduced that the sign of the function on the left hand side of (2.28) is dependent on the value of c . For insufficiently large values of c , the descent property may not be satisfied. When this is the case, the penalty parameter is updated to a new value $c_{k+1} > c_k$ which guarantees the satisfaction of the descent condition. Lemma 2 shows that the algorithm chooses the value of the penalty parameter to satisfy the descent condition. In Lemmas 2 and 4, we show that descent is always guaranteed if (2.25) holds or if $G(v_k, \bar{x}_k) = 0$, and that the penalty parameter $c_k = c_k(\epsilon_G)$ remains finite, due to the use of two merit functions. This scheme is a variation of the *watchdog* technique, first suggested by Chamberlain *et al* [13], and used in the context of interior point methods by Gay *et al* in [27]. The convergence property of an algorithm based on the merit function (2.27) is discussed in El-Bakry *et al* [21]. The algorithm in [21] has two salient features that need to be noted. The first is the choice of penalty parameter selection rule for (2.26) which extends the sequential quadratic programming approach proposed in Rustem ([63], [66]) for nonlinear programming and in Rustem and Nguyen ([67]) for minimax. The

second, and perhaps more important, aspect is the objective and method of application of the second merit function, (2.27). We illustrate, through a numerical example in *Appendix I*, that the l_2 norm of the KKT conditions on its own is unsuitable for use as a merit function for non-convex optimization, but this is overcome with a combination of merit functions (2.26) and (2.27).

2.1.7 Line search strategies

The algorithm utilises an Armijo line search strategy to determine step-sizes in the search direction such that decrease of whichever merit function is in use is ensured. When $Merit_1$ is in use, the stepsize strategy in the algorithm determines the stepsize τ_{pk} as the largest value of $\tau = (\tau)^j, j = 0, 1, 2, \dots$, such that $v_{k+1}, \bar{x}_{k+1}$ given by $\bar{x}_{k+1} = \bar{x}_k + \tau_{pk}\Delta\bar{x}_k, v_{k+1} = \max_j f^j(\bar{x}_{k+1})$ satisfy the inequality

$$F(v_{k+1}, \bar{x}_{k+1}, c_{k+1}, \mu) - F(v_k, \bar{x}_k, c_k, \mu) \leq \rho \tau_{pk} \phi(\Delta v_k, \Delta \bar{x}_k, c_{k+1}, \mu), \quad (2.29)$$

where $\rho \in (0, \frac{1}{2})$ is a given scalar and

$$\phi(\Delta v_k, \Delta \bar{x}_k, c_{k+1}, \mu) = \Delta v_k - \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1} - c_{k+1} \langle G_k, G_k \rangle + \frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle. \quad (2.30)$$

When $Merit_2$ is in use, the stepsize strategy in the algorithm determines the stepsize τ_k as the largest value of $\tau = (\tau)^j, j = 0, 1, 2, \dots$, such that w_{k+1} given by

$$w_{k+1} = w_k + \tau_k \Delta w_k$$

satisfies the inequality

$$\|E(w_k + \tau_k \Delta w_k); \mu = 0\|_2^2 - \|E(w_k); \mu = 0\|_2^2 \leq \rho \tau_k \langle \nabla E'(w_k) E(w_k), \Delta w_k \rangle \quad (2.31)$$

2.1.8 Step-size rules:

Primal step-size rule

For the primal variables, (x, s) , represented by $\bar{x}$, we determine the maximum allowable step-sizes by the boundary of the feasible region:

$$\tau_{pk}^{max} = \min_{1 \leq i \leq (N)} \left\{ \frac{(-\bar{x}_k^i)}{(\Delta \bar{x}_k^i)} \mid (\Delta \bar{x}_k^i) < 0 \right\}.$$

This is indeed the maximum step allowed because τ_{pk}^{max} gives an infinitely large value to at least one term of the logarithmic barrier function $\sum_{i=1}^N \ln \bar{x}_{k+1}^i$. However, if the step-size is in the interval $[0, \tau_{pk}^{max})$ then the next primal iterate $\bar{x}_{k+1}$ is strictly feasible and none of the logarithmic terms becomes infinitely large. As an initial step, $\hat{\tau}_{pk}$, we choose a number close to τ_{pk}^{max} but never greater than one, *i.e.* $\hat{\tau}_{pk} = \min[\theta \tau_{pk}^{max}, 1]$, with $\theta \in (0, 1)$. The Armijo rule is then applied to give $\tau_{pk} = \beta^\Omega \hat{\tau}_{pk}$, with β chosen to be in the interval $[0.1, 0.5]$, dependent on the confidence on the initial step $\hat{\tau}_{pk}$.

Dual step-size rule:

The dual step-size strategy used by the algorithm follows that described in Akrotirianakis and Rustem ([2]). It is a modification of the strategy suggested by Yamashita in [80] and Yamashita and Yabe [81].

For a fixed value of the barrier parameter μ , we determine a stepsize τ_{dk}^i along the direction Δz_k^i for each dual variable $z_k^i, i = 1, 2, \dots, N$, such that the following box constraints are satisfied:

$$\tau_{dk}^i = \max_{1 \leq i \leq N} \left\{ \tau_i > 0 : LB_k^i \leq (\bar{x}_k^i + \tau_{pk} \Delta \bar{x}_k^i) (z_k^i + \tau \Delta z_k^i) \leq UB_k^i \right\} \quad (2.32)$$

with LB_k^i and UB_k^i defined, as in Step 2.4 of the algorithm, as:

$$LB_k^i = \min\{\frac{1}{2}\hat{m}_k\mu, (\bar{x}_{k+1}^i)(z_k^i)\}$$

and

$$UB_k^i = \min\{2\hat{M}_k\mu, (\bar{x}_{k+1}^i)(z_k^i)\}.$$

The parameters $\hat{m}_k$ and $\hat{M}_k$ are chosen to satisfy

$$0 < \hat{m}_k \leq \min\{1, \frac{(1-\eta)(1-\frac{\eta}{(M_0)^{\eta'}})\min_i\{(\bar{x}_k^i)(z_k^i)\}}{\mu}\}, (2.33)$$

and

$$\hat{M}_k \geq \max\{1, \frac{\max_i\{(\bar{x}_k^i)(z_k^i)\}}{\mu}\} > 0, (2.34)$$

for a fixed value of μ , with $\eta \in (0, 1)$ and M_0 a large positive number. These two parameters are fixed to constants which satisfy (2.33) and (2.34).

The common dual step τ_{dk} , is the minimum of all individual step-lengths τ_{dk}^i with the restriction of always being no greater than one. Numerically:

$$\tau_{dk} = \min\left\{1, \min_{1 \leq i \leq N} \{\tau_{dk}^i\}\right\}.$$

The step-lengths for the dual variables λ can be set to either τ_{dk} or 1. It is clear from the construction of the lower and upper bounds that

$$\frac{LB_k^i}{\bar{x}_k^i + \tau_{pk}\Delta\bar{x}_k^i} \leq z_k^i \leq \frac{UB_k^i}{\bar{x}_k^i + \tau_{pk}\Delta\bar{x}_k^i}$$

for all $i = 1, \dots, N$.

The dual step-size rule defines a neighbourhood of the central path beyond which the iterates cannot stray. It does so by setting a lower bound which prevents the products $\bar{x}_k^i z_k^i$ from becoming very much smaller than the current value of the barrier parameter, preventing them reaching the boundary of the feasible region prematurely. The dual step-size rule also prevents the variables $\bar{x}_k^i$ and z_k^i from becoming unbounded by setting an upper bound

on them that is proportional to the barrier parameter.

The step-size rule used is as described in ([2])), which differs from that that proposed in Yamashita and Yabe ([81]) in the definition of the lower bounds, $LB_k^i, i = 1, \dots, N$

2.1.9 Selection of barrier parameter:

The selection of the barrier parameter also follows the strategy outlined in Akrotirianakis and Rustem ([2]). It is restated here for convenience

$$\begin{aligned}
 \mu^{l+1} &= \min\{0.95\mu^l, 0.01(0.95)^k \|E(v_k, \bar{x}_k, \lambda_k)\|_2\} \\
 \text{If } \|E(v_k, \bar{x}_k, \lambda_k; \mu^l)\|_2 &\leq 0.1\eta\mu^l \text{ then} \\
 \text{If } \mu^l < 10^{-4} &\text{ then} \\
 \mu^{l+1} &= \min\{0.95\mu^l, 0.01(0.95)^{k+2\sigma} \|E(v_k, \bar{x}_k, \lambda_k)\|_2\} \\
 \text{Else } \mu^{l+1} &= \min\{0.95\mu^l, 0.01(0.95)^{k+\sigma} \|E(v_k, \bar{x}_k, \lambda_k)\|_2\},
 \end{aligned} \tag{2.35}$$

where $\sigma > 0$ is a parameter which accelerates the decrease of μ at appropriate points. The algorithm checks if the current point is close enough to the central path (*i.e.* if $\|E(v_k, \bar{x}_k, \lambda_k; \mu^l)\|_2 \leq 0.1\eta\mu^l$) and to the optimum solution (*i.e.* if $\mu^l < 10^{-4}$), in which case the barrier parameter is multiplied by $(0.85)^{2\sigma}$, reducing it at a faster rate. If the current point is close to the central path but not an optimum solution, the barrier parameter is multiplied by the larger factor $(0.85)^\sigma$.

Algorithm :**STEP 0: Initialisation:**

Given $\tilde{x}^0 \in \mathbb{R}^N, > 0$ and $\tilde{v}^0 \in \mathbb{R}^1$, choose $\tilde{\lambda}^0 \in \mathbb{R}^q$, and $\chi \in (0, \frac{1}{2}]$. Set $\tilde{v}^0 = \max_j f^j(\tilde{x}^0)$, penalty and

barrier parameters $c_0 > 0, \mu^0 > 0$ and parameters: $\theta, \epsilon_0, \eta, \rho \in (0, \frac{1}{2}), \delta > 0$.

Set $l = 0, k = 0$ and new-merit = **false**.

STEP 1: Test for convergence of outer iterations:

If $\|E(\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l; \mu^l)\|_2 / (1 + \|(\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l)\|_2) \leq \epsilon_0$, then STOP.

STEP 2: Start of inner iterations: (μ is fixed to μ^l throughout this step)

Set $(v_k, \bar{x}_k, z_k, \lambda_k) = (\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l)$

Step 2.1: Test for convergence of inner iterations:

If $(\|E(v_k, \bar{x}_k, \lambda_k; \mu^l)\|_2 \leq \eta\mu^l) \text{ and } \|G_k\|_2^2 \leq \epsilon_G)$ then

Set $(\tilde{v}^{l+1}, \tilde{x}^{l+1}, \tilde{z}^{l+1}, \tilde{\lambda}^{l+1}) = (v_k, \bar{x}_k, \lambda_k)$ and GOTO step 3

Step 2.2: Solve Newton system (2.14) to obtain $\Delta w_k = (\Delta v_k, \Delta \bar{x}_k, \Delta z_k, \Delta \lambda_k)$ **Step 2.3: Penalty parameter selection:**

Set $c_{k+1} = c_k$

If (.not.new-merit) then

$$\mathcal{M}_{num} = \Delta v_k - \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1} - c_k \langle G_k, G_k \rangle + (\chi + \frac{1}{2}) \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle$$

If $(\mathcal{M}_{num} > 0 \text{ and } \|G_k\|_2^2 > \epsilon_G)$ then

$$c_{k+1} = \max \left\{ \frac{\Delta v_k - \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1} + (\chi + \frac{1}{2}) \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle}{\langle G_k, G_k \rangle}, c_k + \delta \right\}.$$

Step 2.4: Step-length selection rules:

$$\text{Set } \tau_{pk}^{\max} = \min_{1 \leq i \leq N} \left\{ \frac{(-\bar{x}_k^i)}{(\Delta \bar{x}_k^i)} \middle| (\Delta \bar{x}_k^i) < 0 \right\}$$

$$\text{and } \hat{\tau}_{pk} = \min\{\theta \tau_{pk}^{\max}, 1\}.$$

If $(\mathcal{M}_{num} > 0 \text{ and } (0 < \|G_k\|_2^2 \leq \epsilon_G)) \text{ or new-merit}$ then

Let $\tau_{pk} = \beta^\Omega \hat{\tau}_{pk}$, where Ω is the smallest non-negative integer such that

$$\|E(w_k + \tau_k \Delta w_k)\|_2^2 - \|E(w_k)\|_2^2 \leq \rho \tau_k \langle \nabla E'(w_k) E(w_k), \Delta w_k \rangle$$

new-merit = true

Else

Let $\tau_{pk} = \beta^\Omega \hat{\tau}_{pk}$, where Ω is the smallest non-negative integer such that

$$F(v_{k+1}, \bar{x}_{k+1}; c_{k+1}, \mu) - F(v_k, \bar{x}_k; c_{k+1}, \mu) \leq \rho \tau_{pk} \phi(\Delta \bar{x}_k, c_{k+1}; \mu)$$

with $\bar{x}_{k+1} = x_k + \tau_{pk} \Delta \bar{x}_k$ and $v_{k+1} = \max_j f^j(x_{k+1})$.

Set $LB_k^i = \min\{\frac{1}{2} \hat{m} \mu, (\bar{x}_{k+1}^i)(z_k^i)\}$ and $UB_k^i = \min\{2 \hat{M} \mu, (\bar{x}_{k+1}^i)(z_k^i)\}$

for some $\hat{m}, \hat{M} > 0$.

For $i = 1, 2, \dots, N$ find:

$$\tau_{dk}^i = \max_{1 \leq i \leq N} \left\{ \tau_i > 0 : LB_k^i \leq (\bar{x}_k^i + \tau_{pk} \Delta \bar{x}_k^i)(z_k^i + \tau \Delta z_k^i) \leq UB_k^i \right\}$$

Set $\tau_{dk} = \min\left\{1, \min_{1 \leq i \leq N} \{\tau_{dk}^i\}\right\}$.

Set $\lambda_{k+1} = \lambda_k + \tau_{dk} \Delta \lambda_k$.

End if

Step 2.5: Set $k = k + 1$ and GOTO Step 2.1

STEP 3: Reduce barrier parameter as described in section §2.1.9.

new-merit = false

STEP 4: Set $l = l + 1$ and GOTO Step 1.

Algorithm 1.

Chapter 3

Convergence Analysis and Numerical Experiments

3.1 Convergence Results

The algorithm is shown to be globally convergent, in the sense that it guarantees progress towards a solution from any starting point. Through a series of proofs, we show that the search direction $\Delta\bar{x}_k$ is a descent direction for the merit function at the current point x_k , and that the merit function is monotonically decreasing for a fixed value of the barrier parameter μ .

Lemma 1 (descent property of $\Delta\bar{x}_k$ when $\|G_k\|_2^2 > \epsilon_G$.)

Let f and g be differentiable functions and let there exist a small $\epsilon_G > 0$ such that $\|G_k\|_2^2 > \epsilon_G$ and $\Delta\bar{x}_k$ be the search direction for variable $\bar{x}$ calculated by the solution of the Newton system (2.14) and c_{k+1} chosen as in Step 2.3 of the algorithm above. Then $\Delta\bar{x}_k$ is a descent direction for the merit function. Furthermore

$$\phi(\Delta\bar{x}_k, c_{k+1}, \mu) \leq -\chi\langle\Delta\bar{x}_k, Q_k\Delta\bar{x}_k\rangle \leq 0. \quad (3.1)$$

Proof Step 2.3 of the algorithm checks if

$$\Delta v_k - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} - c_k \langle G_k, G_k \rangle + (\chi + \frac{1}{2}) \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq 0. \quad (3.2)$$

If this is satisfied' then by setting $c_{k+1} = c_k$ and rearranging (3.2) we obtain

$$\phi(\Delta \bar{x}_k, c_{k+1}; \mu) \leq -\chi \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq 0 \quad (3.3)$$

On the other hand, if (3.2) is not satisfied by setting

$$c_{k+1} \geq \frac{\Delta v_k - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + (\chi + \frac{1}{2}) \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle}{\langle G_k, G_k \rangle} \quad (3.4)$$

and substituting into (2.30), it can be verified that (3.1) holds.

The above Lemma establishes the descent property of the search direction for $\|G_k\|_2^2 > \epsilon_G$. We now go on to demonstrate the descent property for the merit function when $\|G_k\|_2^2 = 0$.

Lemma 2 (Finiteness of penalty parameter and descent property of the search direction when $G_k = 0$.) *Let f and g be differentiable functions and let $(\Delta v_k, \Delta \bar{x}_k, \Delta \lambda_k)$ be the solution to the Newton system (2.14). If $G_k = 0$ for some k , then the descent property $\phi(\Delta \bar{x}_k, c_k, \mu^l) \leq -\frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq 0$ is satisfied for any choice of the penalty parameter, $c_k \in [0, \infty)$. Therefore the search direction $\Delta \bar{x}_k$ obtained from the solution of the Newton system of equations is still a descent direction after feasibility of the equality constraints has been achieved.*

Proof If $G_k = 0$, then from (2.19) and the definition of G_k :

$$\nabla f'_k \Delta x_k - \mathbf{1} \Delta v_k + \Delta s_k = 0, \quad (3.5)$$

$$\nabla g'_k \Delta x_k = 0.$$



Pre-multiplying (2.17) by $\Delta \bar{x}_k$ we get

$$\begin{aligned} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \langle \Delta \bar{x}_k, \nabla G_k \Delta \lambda_k \rangle - \langle \Delta \bar{x}_k, \Delta z_k \rangle = \\ -\frac{1}{2} \langle \Delta \bar{x}_k, \nabla G_k \lambda_k \rangle + \langle \Delta \bar{x}_k, z_k \rangle. \end{aligned} \quad (3.6)$$

Rearranging the equation above yields

$$\langle \Delta \bar{x}_k, \nabla G_k [\lambda_k + \Delta \lambda_k] \rangle = -\langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \langle \Delta \bar{x}_k, z_k + \Delta z_k \rangle. \quad (3.7)$$

Considering (2.30) for $G_k = 0$ yields:

$$\phi(\Delta \bar{x}_k, c_{k+1}, \mu) = \Delta v_k - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle. \quad (3.8)$$

Multiplying (2.19) by $(\lambda_k + \Delta \lambda_k)'$ results in

$$\Delta v_k = \langle \Delta \bar{x}_k, \nabla G_k [\lambda_k + \Delta \lambda_k] \rangle, \quad (3.9)$$

when $G_k = 0$.

Substituting for Δv_k in (3.8) with (3.9) and using equation (3.7) establishes that

$$\begin{aligned} \phi(\Delta \bar{x}_k, c_{k+1}, \mu) &= -\langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \langle \Delta \bar{x}_k, z_k + \Delta z_k \rangle \\ &\quad - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \\ &= -\frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \langle \Delta \bar{x}_k, z_k \rangle + \langle \Delta \bar{x}_k, \Delta z_k \rangle + \\ &\quad - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1}, \end{aligned} \quad (3.10)$$

at $G_k = 0$. Solving equation (2.18) for Δz_k gives

$$\Delta z_k = -z_k - \bar{X}_k^{-1} Z_k \Delta \bar{x}_k + \mu \bar{X}_k^{-1} \mathbf{1}. \quad (3.11)$$

Multiplying (3.11) by $\Delta \bar{x}_k'$ yields

$$\langle \Delta \bar{x}_k, \Delta z_k \rangle = -\Delta \bar{x}_k' \bar{X}_k^{-1} Z_k \Delta \bar{x}_k - \Delta \bar{x}_k' z_k + \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1}. \quad (3.12)$$

Substituting equations (3.12) into (3.10) produces

$$\begin{aligned} \phi(\Delta \bar{x}_k, c_{k+1}, \mu) &= -\frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle - \Delta \bar{x}_k' \bar{X}_k^{-1} Z_k \Delta \bar{x}_k \\ &\leq -\frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle - \frac{1}{2} \langle \Delta \bar{x}_k, \bar{X}_k^{-1} Z_k \Delta \bar{x}_k \rangle \\ &= -\frac{1}{2} \langle \Delta \bar{x}_k, (Q_k + \bar{X}_k^{-1} Z_k) \Delta \bar{x}_k \rangle \\ &\leq -\frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \\ &\leq 0. \end{aligned} \quad (3.13)$$

Therefore, the condition in Step 2.3 of the algorithm is always satisfied when $G_k = 0$ and c_{k+1} is set to c_k . It can be verified that the descent property of the search direction still holds when the value of the penalty parameter is unchanged (see Step 2.4 of the algorithm).

Lemma 3 (Descent property of the search direction for $0 < \|G_k\|_2^2 \leq \epsilon_G$) *The direction Δw_k calculated by solving the Newton System (2.14) with c_{k+1} chosen Algorithm 1 is a descent direction for Merit_2 when $0 < \|G_k\|_2^2 \leq \epsilon_G$*

Proof From (2.15),

$$\Delta w_k = -\nabla E_k^{-1} E_k$$

When $0 < \|G_k\|_2^2 \leq \epsilon_G$, Step 2.4 of the algorithm uses merit function $D_k = \langle E_k, E_k \rangle$. Then the descent property is established from

$$\nabla D_k' \Delta w_k = 2(\nabla E_k' E_k)' [-\nabla E_k^{-1} E_k] = -2\langle E_k, E_k \rangle \leq 0.$$

Corollary 1 *If $\|\Delta w_k\|=0$, the algorithm chooses $c_{k+1}=c_k$.*

Proof From (2.19) and the definition of G_k , when $\Delta w_k = 0$, $G_k = 0$, therefore this situation is covered by Lemma 2. (See Akrotirianakis and Rustem ([2])).

Lemma 4 *Let f and g be twice continuously differentiable functions. Then, for a fixed value μ^l of μ :*

- *There exists a constant $c_{k+1} \geq 0$, satisfying Step 2.3 of the algorithm.*
- *Assuming that the sequence $\bar{x}_k$ is bounded, c_k is increased finitely often. That is, there exists an integer $k_* \geq 0$, and $c_* \geq 0$, such that for all $k \geq k_*$, $c_k = c_* < \infty$.*

Proof Item 1 follows directly from the preceding Lemmas, since a finite value c_{k+1} is always generated, in Step 2.3. Item 2 can be shown by contradiction. Assume that $c_k \rightarrow \infty$ as $k \rightarrow \infty$. From the definition of c_{k+1} in Step 2.3, we can deduce that if $c_k \rightarrow \infty$ then $\|G_k\|_2^2 \rightarrow 0$. Hence, there exists an integer k_1 such that for all $k \geq k_1$ we have $\|G_k\|_2^2 \leq \epsilon_G$. In Step 2.4 of the algorithm, however, when $\|G_k\|_2^2 \leq \epsilon_G$, the algorithm stops increasing the penalty parameter and switches to the second merit function. Therefore, the maximum value that c_k can attain is $c_* = c_{k_1} = M_{num}/\epsilon_G$ where M_{num} and c_* are finite values. Hence, $c_* < \infty$, contrary to our assumption that $c_k \rightarrow \infty$ as $k \rightarrow \infty$. Hence, the penalty parameter does not increase indefinitely and there exists an integer $k_* \geq 0$ such that for all $k \geq k_*$ we have $c_k < \infty$.

3.1.1 Convergence of the inner iterations

Consider the objective function

$$\max_{\alpha \in \Xi_m^+} \mathfrak{J}(x, \alpha) = \max_{\alpha \in \Xi_m^+} \langle \alpha, f \rangle = \max_j (f^j),$$

and its linear approximation with respect to x at x_k

$$\max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(x, \alpha) = \max_{\alpha \in \Xi_m^+} \langle \alpha, f(x_k) + \nabla f'(x_k)(x - x_k) \rangle = \max_j (f_k^j + \langle \nabla f_k^j, \Delta x_k \rangle).$$

This objective function is used in the following Theorem.

Theorem 1 (Monotonic descent of Merit_1) *Let:*

- *Assumption A1 hold;*
- *the (approximate) Hessian Q_k be such that, for each iteration k ,*

$$m_1 \|y\|_2^2 \leq y' Q_k y \leq M_1 \|y\|_2^2; M_1 > m_1 > 0; 0 \neq y \in R^n \quad (3.14)$$

- *for each k , there exists a bounded point $\Delta w_k = (\Delta v_k, \Delta \bar{x}_k, \Delta z_k, \Delta \lambda_k)$ as a solution to the Newton system (2.14)*
- *there exists an iteration k_* , small $\epsilon_G > 0$, $\|G_k\|_2^2 \notin (0, \epsilon_G)$ and a scalar $c_* \geq 0$, $c_* = c_*(\epsilon_G)$, such that the condition*

$$\Delta v_k - \mu \Delta \bar{x}_k' \bar{X}_k^{-1} \mathbf{1} - c_k(\epsilon_G) \langle G_k, G_k \rangle + \frac{1}{2} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq 0,$$

is satisfied for all $k \geq k_*$ and $c_{k+1}(\epsilon_G) = c_k(\epsilon_G) = c_*(\epsilon_G)$.

Then the step-size computed in Step 2.4 is such that $\tau_{pk} \in [0, 1]$ and hence the sequence $\{v_k + \frac{1}{2}c_k \langle G_k, G_k \rangle - \mu \sum \ln(\bar{x}_k^i)\}$ is monotonically decreasing for $k \geq k_*$ and μ fixed.

Proof From Step 2.4 of the algorithm and for $\bar{x}_{k+1} = \bar{x}_k + \tau_{pk} \Delta \bar{x}_k$:

$$\begin{aligned} v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i \\ = \max_{\alpha \in \Xi_m^+} \langle \alpha, \bar{f}_k + \tau_{pk} \Delta \bar{x}_k \rangle + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i. \end{aligned} \quad (3.15)$$

Considering the second order Taylor's series expansion of the above with remainder for the case $\|G_k\|_2^2$, we obtain the equation:

$$\begin{aligned} & v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i \\ & \leq \max_{\alpha \in \Xi_m^+} \langle \alpha, f_k + \tau_{pk} \nabla f'_k \Delta x_k \rangle + \frac{1}{2} \tau_{pk}^2 \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \frac{1}{2} c_* \{ \langle G_k, G_k \rangle + 2\tau_{pk} \langle G_k, \nabla G'_k \Delta w_{pk} \rangle \} \\ & \quad - \mu \sum_i \ln \bar{x}_k^i - \tau_{pk} \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \xi_k \tau_{pk}^2 \|\Delta \bar{x}_k\|_2^2, \end{aligned} \quad (3.16)$$

where

$$\begin{aligned} \xi_k = \int_0^1 (1-t) & \left\| \max_{\alpha \in \Xi_m^+} \{ \sum_i \alpha^i \nabla^2 f^i(x, t) \} + c \langle \nabla^2 G(\bar{x}(t)), \nabla^2 G(\bar{x}(t)) \rangle \right. \\ & \left. + c \sum_i G^i(\bar{x}(t)) \nabla^2 G^i(\bar{x}(t)) - \mu \sum_i \frac{\Delta \bar{x}_k^i}{\bar{x}_k^i} - Q_k \right\| dt \end{aligned} \quad (3.17)$$

and

$$w_{pk} = (v_k, \bar{x}_k)' \text{ and } \Delta w_{pk} = (\Delta v_k, \Delta \bar{x}_k)'.$$

Using

$$\max_{\alpha \in \Xi_m^+} \langle f(x_k) + \tau_{pk} \nabla f'(x_k) \Delta x_k, \alpha \rangle = \max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(x_k + \Delta x_k, \alpha),$$

we can write

$$\begin{aligned} & v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i \\ & \leq \max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(x_k + \tau_{pk} \Delta x_k, \alpha) + \frac{1}{2} \tau_{pk}^2 \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \frac{1}{2} c_* \{ \langle G_k, G_k \rangle + 2\tau_{pk} \langle G_k, \nabla G'_k \Delta w_{pk} \rangle \} \\ & \quad - \mu \sum_i \ln \bar{x}_k^i - \tau_{pk} \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \xi_k \tau_{pk}^2 \|\Delta \bar{x}_k\|_2^2. \end{aligned} \quad (3.18)$$

By convexity of $\max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(x, \alpha)$:

$$\begin{aligned} \max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(\bar{x}_k + \tau_{pk} \Delta \bar{x}_k, \alpha) & \leq \max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(\bar{x}_k, \alpha) + \tau_{pk} [\max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(\bar{x}_k + \Delta \bar{x}_k, \alpha) - \max_{\alpha \in \Xi_m^+} \mathfrak{J}_k(\bar{x}_k, \alpha)] \\ & \leq v_k + \tau_{pk} [(v_k + \Delta v_k) - v_k]. \end{aligned}$$

Therefore

$$\begin{aligned}
& v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i \\
& \leq v_k + \tau_{pk} \Delta v_k + \frac{1}{2} \tau_{pk}^2 \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + \frac{1}{2} c_* \{ \langle G_k, G_k \rangle + 2 \tau_{pk} \langle G_k, \nabla G'_k \Delta w_{pk} \rangle \} \\
& \quad - \mu \sum_i \ln \bar{x}_k^i - \tau_{pk} \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \xi_k \tau_{pk}^2 \|\Delta \bar{x}_k\|_2^2
\end{aligned} \tag{3.19}$$

Rearranging this, and using (2.20) yields:

$$\begin{aligned}
& v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i - v_k - \frac{1}{2}c_* \langle G_k, G_k \rangle + \mu \sum_i \ln \bar{x}_k^i \\
& \leq \tau_{pk} \Delta v_k + \frac{1}{2} \tau_{pk}^2 \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle + c_* \tau_{pk} \langle G_k, \nabla G'_k \Delta w_{pk} \rangle - \mu \tau_{pk} \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \xi_k \tau_{pk}^2 \|\Delta \bar{x}_k\|_2^2 \\
& = \tau_{pk} [\Delta v_k + c_* \langle G_k, \nabla G'_k \Delta w_k \rangle - \mu \Delta \bar{x}'_k \bar{X}_k^{-1} \mathbf{1} + \frac{1}{2} \tau_{pk} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle] + \xi_k \tau_{pk}^2 \|\Delta \bar{x}_k\|_2^2 \\
& \leq \tau_{pk} [\phi(\Delta \bar{x}_k, c_*, \mu) + \xi_k \tau_{pk} \|\Delta \bar{x}_k\|_2^2].
\end{aligned} \tag{3.20}$$

Moreover, from (3.14) above,

$$\|\Delta \bar{x}_k\|_2^2 \leq \frac{1}{m_1} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle,$$

and from Lemmas 2 and 4 we have

$$\langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq -\frac{\phi(\Delta \bar{x}_k, c_k; \mu)}{\chi}$$

which combine to yield

$$\|\Delta \bar{x}_k\|_2^2 \leq \frac{1}{m_1} \langle \Delta \bar{x}_k, Q_k \Delta \bar{x}_k \rangle \leq -\frac{\phi(\Delta \bar{x}_k, c_k, \mu)}{\chi m_1}.$$

Substituting this into (3.20) above gives:

$$\begin{aligned}
& v_{k+1} + \frac{1}{2}c_* \langle G_{k+1}, G_{k+1} \rangle - \mu \sum_i \ln \bar{x}_{k+1}^i \\
& \quad - v_k - \frac{1}{2}c_* \langle G_k, G_k \rangle + \mu \sum_i \ln \bar{x}_k^i \\
& \leq \tau_{pk} \phi(\Delta \bar{x}_k, c_*, \mu) [1 - \tau_{pk} (\frac{\xi_k}{\chi m_1})]
\end{aligned} \tag{3.21}$$

The stepsize strategy in Step 2.4 of the algorithm determines τ_{pk} such that $\rho \leq 1 - \tau_{pk} (\frac{\xi_k}{\chi m_1}) \leq$

$\frac{1}{2}$. Since by Lemmas 2 and 4, the descent property $\phi(\Delta\bar{x}_k, c_{k+1}, \mu) \leq 0$ holds, there exists a τ_{pk} to ensure that (3.21) and hence (2.29) holds. Since $\phi(\Delta\bar{x}_k, c_{k+1}, \mu) \leq 0$, it follows that $v_k + \frac{1}{2}c_k\langle G_k, G_k \rangle - \mu \sum_i \ln \bar{x}_k^i$ is monotonically decreasing.

Lemma 5 *Let the assumptions of Theorem 1 be satisfied and let, for some k_0 and $k \geq k_0$, the set*

$$\begin{aligned} U = \{ \bar{x} \in R^N, v \in R^1 | v - \frac{1}{2}c_*\langle G(v, \bar{x}), G(v, \bar{x}) \rangle + \mu \sum_i \ln \bar{x}^i \\ \leq v_{k_0} - \frac{1}{2}c_*\langle G_{k_0}, G_{k_0} \rangle + \mu \sum_i \ln \bar{x}_{k_0}^i \} \end{aligned} \quad (3.22)$$

be compact. Then

$$\lim_{k \rightarrow \infty} \phi(\Delta\bar{x}_k, c_*, \mu^l) = 0. \quad (3.23)$$

Proof The scalar $\rho \in (0, \frac{1}{2})$ determined in the step-size strategy described in Step 2.4 of the algorithm is such that $\rho \leq 1 - \tau_{pk}(\frac{\xi_k}{\chi^{m_1}}) \leq \frac{1}{2}$, and by solving for τ_{pk} we get,

$$\frac{\frac{1}{2}}{(\frac{\xi_k}{\chi^{m_1}})} \leq \tau_{pk} \leq \frac{1-\rho}{(\frac{\xi_k}{\chi^{m_1}})}$$

Hence, the largest value that τ_{pk} can take whilst ensuring satisfaction of the Armijo rule in Step 2.4 is the choice $\tau_0 = \min\{1, \frac{1-\rho}{(\frac{\xi_k}{\chi^{m_1}})}\}$. Clearly, the choice of the stepsize $\tau_{pk} = \beta^\Omega \hat{\tau}_{pk}$ in Step 2.4 is in the range $\tau_{pk} \in [\beta\tau_0, \tau_0]$ and thereby also satisfies (2.29). As f and g are twice continuously differentiable and as U is a compact set, there is a scalar $\bar{M} < \infty$ such that $\xi_k \leq \bar{M}$. Thus, as $m_1 > 0$, we have $\tau_k \geq \bar{\tau} > 0$, for all k , for the step-size strategy defined by (2.29) and for some $\bar{\tau} = \min\{1, \frac{1-\rho}{(\frac{\bar{M}}{\chi^{m_1}})}\}$. Hence, the stepsize τ_{pk} is always bounded away from zero. The boundedness of $v_k - \frac{1}{2}c_*\langle G_k, G_k \rangle + \mu \sum_i \ln \bar{x}_k^i$ on U and

$$\phi(\Delta v_k, \Delta \bar{x}_k, c_*, \mu) < 0$$

imply

$$\begin{aligned} 0 &\leq \rho \sum_k \tau_{pk} |\phi(\Delta \bar{x}_k, c_*, \mu)| \\ &\leq \sum_k [v_k - \frac{1}{2} c_* \langle G_k, G_k \rangle + \mu \sum_i \ln \bar{x}_k^i] - [v_{k+1} - \frac{1}{2} c_* \langle G_{k+1}, G_{k+1} \rangle + \\ &\quad \mu \sum_i \ln \bar{x}_{k+1}^i] < \infty, \end{aligned}$$

which yields (3.23). Therefore, from Lemmas 2 and 3, $\{\phi(\Delta v_k, \Delta \bar{x}_k, c_*, \mu)\}$ is convergent and bounded below. From (2.26), this implies that $\{\Delta \bar{x}_k\}$ is also bounded.

Lemma 6 *Let the assumptions of Theorem 1 be satisfied. Then, inequality (3.1) and Lemmas , 2 and 5 imply*

$$\lim_{k \rightarrow \infty} \|\Delta \bar{x}_k\| = 0, \quad (3.24)$$

and

$$\lim_{k \rightarrow \infty} \{\Delta v_k\} = 0. \quad (3.25)$$

Proof The proof of (3.24) and (3.25) follows from (3.23) and (3.1).

Corollary 2 *The sequence $\{\bar{x}_k\}$ generated by the algorithm is bounded above and away from zero.*

Proof Assume to the contrary, that the sequence $\{\|\bar{x}_k\|\} \rightarrow 0$. Then

$$\{-\sum_i \ln \bar{x}_k^i\} \rightarrow \infty.$$

From Assumption A5 that the feasible region is bounded, the sequences $f(x_k)$, $\{\|G(v_k, \bar{x}_k)\|\}$ and $\{\bar{x}_k\}$ are also bounded. Hence the sequence

$$\{v_k + \frac{1}{2}c_* \langle G_k, G_k \rangle - \mu \sum_i \ln \bar{x}_k^i\} \rightarrow \infty.$$

This contradicts the monotonic decrease of this sequence established in Theorem 1. As the iterates are always feasible in an interior point algorithm, boundedness of the feasible region implies boundedness of the iterates.

Lemma 7 *For fixed μ , the lower bounds LB_k^i and the upper bounds $UB_k^i, i=1,2,\dots,N$ of the box constraints in the dual step-size rule are bounded away from zero and bounded from above respectively, if the corresponding $\bar{x}_k^i$ of the iterates $\bar{x}_k$ are also bounded above and away from zero.*

Proof From the dual step-size rule, the proof is established by Yamashita in [81].

This implies that $\{z_k\}$ is also a bounded sequence. We have thus established that $\{\bar{x}_k\}$ and $\{z_k\}$ are bounded sequences.

Lemma 8 *Let (3.14) be satisfied and $\{w_k\}$ be a sequence of vectors generated by the algorithm. Then the sequence of vectors, $\{\lambda_k + \Delta\lambda_k\}$, and $\{z_k + \Delta z_k\}$ are bounded.*

Proof : Making $\Delta\bar{x}_k$ the subject of equation (2.19) establishes G_k, v_k and ∇G_k , as bounded sequences. From (3.14) in Theorem 1 and Assumption A4, Q_k is bounded from below and away from zero. Rearranging (2.17) and multiplying by $\Delta\bar{x}$ gives:

$$-\langle \Delta\bar{x}_k, \nabla G_k(\lambda_k + \Delta\lambda_k) \rangle = \langle \Delta\bar{x}_k, Q_k \Delta\bar{x}_k \rangle + \langle \Delta\bar{x}_k, z_k + \Delta z_k \rangle \quad (3.26)$$

Rearranging the above equation gives

$$-\langle \Delta\bar{x}_k, Q_k \Delta\bar{x}_k \rangle = \langle \Delta\bar{x}_k, \nabla G_k(\lambda_k + \Delta\lambda_k) \rangle + \langle \Delta\bar{x}_k, z_k + \Delta z_k \rangle \quad (3.27)$$

From (3.14), we have

$$-m_1 \|\Delta \bar{x}_k\|_2^2 \geq -\Delta \bar{x}_k' Q_k \Delta \bar{x}_k \geq -M_1 \|\Delta \bar{x}_k\|_2^2 \quad (3.28)$$

Therefore

$$-m_1 \|\Delta \bar{x}_k\|_2^2 \geq \langle \Delta \bar{x}_k, \nabla G_k(\lambda_k + \Delta \lambda_k) \rangle + \langle \Delta \bar{x}_k, z_k + \Delta z_k \rangle \geq -M_1 \|\Delta \bar{x}_k\|_2^2 \quad (3.29)$$

From Lemmas 5, 6 and the boundedness of Q_k , we establish that $\{\lambda_k + \Delta \lambda_k\}$ and $\{z_k + \Delta z_k\}$ are also bounded sequences.

Lemma 9 *Let $\{w_k\}$ be a sequence of vectors generated by the algorithm. Then the sequence of vectors, $\{\Delta \bar{x}_k, \lambda_k + \Delta \lambda_k, z_k + \Delta z_k\}$ is bounded.*

Proof The result follows from Lemmas 6, 7 and 8 above.

3.1.2 Convergence of the Outer Iterations

Theorem 2 *Let the point satisfying the perturbed first order (KKT) conditions of (2.4) for a fixed value of μ be given by $(v_*(\mu), \bar{x}_*(\mu), z_*(\mu), \lambda_*(\mu))$ and the assumptions of Theorem 1 be satisfied. Then the algorithm either terminates at*

$$(v_*(\mu), \bar{x}_*(\mu), z_*(\mu), \lambda_*(\mu))$$

or it generates an infinite sequence $\{\bar{x}_k\}$ in which there exists a subsequence $\{\bar{x}_k\}$ with $k_* \leq k \in K \subset \{0, 1, \dots\}$ and every accumulation point of the infinite sequence $\{\bar{x}_k\}$ is $\bar{x}_*$.

Proof: By Lemma 2, there exists a subsequence $\{\bar{x}_k\}$ such that $\|\Delta \bar{x}_k\| \rightarrow 0$. Let there exist $(v_*(\mu), \bar{x}_*(\mu), z_*(\mu), \lambda_*(\mu))$, such that $\{v_k\} \rightarrow v_*(\mu)$, $\{\bar{x}_k\} \rightarrow \bar{x}_*(\mu)$, $\{z_k\} \rightarrow z_*(\mu)$, $k_* \leq k \in K$. The existence of such a point is ensured since, by assumption A2, corollary 2

and Lemmas 7 and 8, the sequence $(v_*(\mu), \bar{x}_*(\mu), z_*(\mu), \lambda_*(\mu))$ is bounded for μ fixed and by Theorem 1, the algorithm decreases $\varphi(\bar{x}_k) - \frac{1}{2}c_*\langle G_k, G_k \rangle + \mu \sum_i \ln \bar{x}_k^i$ sufficiently at each iteration, thereby ensuring $\bar{x}_k \in U$, where U is compact. It is shown that for sufficiently large k , unit step-sizes are attained by establishing that

$$\lim_{k \rightarrow \infty} \|z_k + \Delta z_k - \mu \bar{X}_{k+1}^{-1} \mathbf{1}\| = 0 \quad (3.30)$$

Adding $-\mu \bar{X}_{k+1}^{-1} \mathbf{1}$ to both sides of equation (2.18) and rearranging gives

$$\|z_k + \Delta z_k - \mu \bar{X}_{k+1}^{-1} \mathbf{1}\| \leq \|-\bar{X}_k^{-1} Z_k\| \|\Delta \bar{x}_k\| + \mu \|\bar{X}_k^{-1} - \bar{X}_{k+1}^{-1}\| \|\mathbf{1}\|. \quad (3.31)$$

Furthermore,

$$\|\bar{X}_k^{-1} - \bar{X}_{k+1}^{-1}\|_2^2 \leq N \max_{1 \leq i \leq N} \left\{ \left(\frac{1}{\bar{x}_k^i} - \frac{1}{\bar{x}_{k+1}^i} \right)^2 \right\} = N \max_{1 \leq i \leq N} \left\{ \frac{(\tau_{pk})^2 (\Delta x_k^i)^2}{(\bar{x}_k^i)^2 (\bar{x}_{k+1}^i)^2} \right\}$$

Since $\tau_{pk} \in (0, 1]$, $(\Delta \bar{x}_k^i)^2 \leq \|\Delta \bar{x}_k\|_2^2$ and the sequence $\{\bar{x}_k\}$ is bounded away from zero, from the above inequality and (3.24), we can derive

$$\lim_{k \rightarrow \infty} \|\bar{X}_k^{-1} - \bar{X}_{k+1}^{-1}\|_2^2 \leq m \lim_{k \rightarrow \infty} \max_{1 \leq i \leq m} \left\{ \frac{\|\Delta \bar{x}_k\|_2^2}{(\bar{x}_k^i)^2 (\bar{x}_{k+1}^i)^2} \right\} = 0. \quad (3.32)$$

Consequently, letting $k \rightarrow \infty$ in (3.31) and using (3.24) and (3.32), we can deduce that

$$\lim_{k \rightarrow \infty} \|z_k + \Delta z_k - \mu \bar{X}_{k+1}^{-1} \mathbf{1}\| = 0.$$

As a result, for sufficiently large k ,

$$z_{k+1} = z_k + \Delta z_k. \quad (3.33)$$

Moreover, using $\Delta z_k = -z_k + \mu \bar{X}_k^{-1} \mathbf{1} - \bar{X}_k^{-1} Z_k \Delta \bar{x}_k$, and for k sufficiently large, the complementarity condition becomes

$$\bar{X}_{k+1} z_{k+1} = \bar{X}_{k+1} (z_k + \Delta z_k) = \bar{X}_{k+1} \bar{X}_k^{-1} (\mu \mathbf{1} - Z_k \Delta \bar{x}_k) \quad (3.34)$$

From (3.24) and the fact that the elements of the diagonal matrix $\bar{X}_{k+1}\bar{X}_k^{-1}$ can be written as $\frac{(\bar{x}_{k+1}^i)}{(\bar{x}_k^i)} = 1 + \tau_{pk}\frac{\Delta\bar{x}_k^i}{(\bar{x}_k^i)}$, for all $i = 1, 2, \dots, N$, it is clear that

$$\lim_{k \rightarrow \infty} \bar{X}_{k+1}\bar{X}_k^{-1} = I_N \quad (3.35)$$

where I_N is the $N \times N$ identity matrix. Therefore, letting $k \rightarrow \infty$ in (3.34) and using (3.24) and (3.35) yields

$$\lim_{k \rightarrow \infty} \bar{X}_{k+1}z_{k+1} = \bar{X}_*(\mu)z_*(\mu) = \mu\mathbf{1}. \quad (3.36)$$

Also, for $k \rightarrow \infty$, the last equation of the Newton system (2.14), (3.24) and (3.25) yield:

$$\lim_{k \rightarrow \infty} \nabla G'_k \Delta \bar{x}_k = G(\bar{x}_*(\mu)) = 0. \quad (3.37)$$

The second equation of the Newton system (2.14) can be written as

$$\nabla G_k \lambda_{k+1} - z_{k+1} = -Q_k \Delta \bar{x}_k. \quad (3.38)$$

Letting $k \rightarrow \infty$ and using (3.24), we get

$$\lim_{k \rightarrow \infty} \|\nabla G_k \lambda_{k+1} - z_{k+1}\| = 0 \quad (3.39)$$

or equivalently

$$\nabla G(\bar{x}_*(\mu)) \lambda_*(\mu) - z_*(\mu) = 0 \quad (3.40)$$

Also, from (3.33) and considering equation (2.16),

$$1 = \langle \lambda_{k+1}, \mathbf{1} \rangle \quad (3.41)$$

and therefore letting $k \rightarrow \infty$, we get

$$1 = \langle \lambda_*(\mu), \mathbf{1} \rangle \quad (3.42)$$

From (3.42), (3.40), (3.38) and (3.37), we conclude that $(v_*(\mu), \bar{x}_*(\mu), z_*(\mu), \lambda_*(\mu))$ is a solution of the perturbed KKT conditions given in (2.14). As observed in Akrotirianakis and Rustem ([2]), an immediate consequence of Theorem 2 is that, for any convergent subsequence produced by the algorithm, for $\mu = \mu^l$, there is an iteration $k \geq \tilde{k}$, such that

$$\|E(v_k, \bar{x}_k, \lambda_k, z_k; c_*, \mu)\|_2 \leq \eta \mu^l \quad (3.43)$$

for all $k \geq \tilde{k}$ where $\eta \geq 0$. Record the value of the current iterate as

$$(\tilde{v}^l, \tilde{\bar{x}}^l, \tilde{z}^l, \tilde{\lambda}^l) = (v_k, \bar{x}_k, z_k, \lambda_k),$$

and set $\mu^{l+1} < \mu^l$. In this way we generate a sequence of approximate central points $\{\tilde{v}^l, \tilde{\bar{x}}^l, \tilde{z}^l, \tilde{\lambda}^l\}$.

In the manner presented in Akrotirianakis and Rustem ([2]), in the rest of this section it will be shown that the sequence of approximate central points converge to a KKT point $(v_*, \bar{x}_*, z_*, \lambda_*)$ of the original constrained minimax problem (2.7).

For a given and sufficiently small $\varepsilon \geq 0$, considering the set of all approximate central points, generated by Algorithm 1,

$$\begin{aligned} S(\varepsilon) &= \{(\tilde{v}^l, \tilde{\bar{x}}^l, \tilde{z}^l, \tilde{\lambda}^l) | \varepsilon \leq \|E(\tilde{v}^l, \tilde{\bar{x}}^l, \tilde{z}^l, \tilde{\lambda}^l; c_*, \mu^l)\|_2 \\ &\leq \|E(\tilde{v}^0, \tilde{\bar{x}}^0, \tilde{z}^0, \tilde{\lambda}^0; c_*, \mu^0)\|_2, \forall \mu^l < \mu^0\}. \end{aligned} \quad (3.44)$$

If $\varepsilon > 0$, the step-size rules described in previous sections guarantee that $\tilde{x}^l$ and $\tilde{z}^l \in S(\varepsilon)$ are bounded away from zero for $l \geq 0$. As a result $(\tilde{x}^l)'(\tilde{z}^l)$ are also bounded away from zero in $S(\varepsilon)$.

The following lemma shows that the sequence $\{\tilde{\lambda}^l\}$ is bounded if the sequence $\{\tilde{z}^l\}$ is also bounded.

Lemma 10 *Assumption A2 holds, and the iterates $\tilde{x}^l$ are in a compact set for $l \geq 0$, then*

there exists a constant $M_1 > 0$ such that

$$\|\tilde{\lambda}'\| \leq M_1(1 + \|\tilde{z}'\|).$$

Proof We define $r' = \tilde{\nabla}G(\tilde{x}')\lambda - \tilde{z}'$, and solve for $\nabla G(\tilde{x}')\tilde{\lambda}'$ to obtain

$$\nabla G(\tilde{x}')\tilde{\lambda}' = \tilde{z}' + r'.$$

From the assumptions made, the above equation can be written as:

$$\tilde{\lambda}' = [\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')(\tilde{z}') + [\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')(\tilde{r}') \quad (3.45)$$

Taking norms of both sides of the above equation yields:

$$\begin{aligned} \|\tilde{\lambda}'\| &\leq \|[\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')\| \|\tilde{z}'\| \\ &\quad + \|[\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')\| \|\tilde{r}'\| \\ &\leq M_1(1 + \|\tilde{z}'\|) \end{aligned} \quad (3.46)$$

where the constant M_1 is defined as

$$M_1 \geq \max\{\|[\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')\|, \|\nabla G(\tilde{x}')\nabla G(\tilde{x}')']^{-1}\nabla G(\tilde{x}')\| \|\tilde{r}'\|\}.$$

and is finite, according to our assumptions.

Theorem 3 *Let $\{\tilde{\mu}^l\}$ be a positive monotonically decreasing sequence of barrier parameters with $\{\tilde{\mu}^l\} \rightarrow 0$, and let $\{\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l\}$ be a sequence of approximate central points satisfying (3.43) for $\mu = \tilde{\mu}^l, l \geq 0$. Then, the sequence $\{\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l\}$ is bounded and all of its limit points $(\tilde{v}^*, \tilde{x}^*, \tilde{z}^*, \tilde{\lambda}^*)$ satisfy the KKT conditions of problem (2.8).*

Proof From Lemma (10), the sequence $\{\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l\}$ is bounded and remains in the compact set $S(\varepsilon)$. Thus it has a limit point in $S_2 \in \varepsilon$, denoted by $(\tilde{v}^*, \tilde{x}^*, \tilde{z}^*, \tilde{\lambda}^*)$. From (3.43) and the fact that $\{\tilde{\mu}^l\} \rightarrow 0$, we easily obtain that

$$\lim_{l \rightarrow \infty} \|E(\tilde{v}^l, \tilde{x}^l, \tilde{z}^l, \tilde{\lambda}^l; c_*, \mu)\|_2 = 0$$

and $(\tilde{v}^*, \tilde{x}^*, \tilde{z}^*, \tilde{\lambda}^*)$ is a KKT point of the original constrained problem (2.8).

3.2 Numerical Results

The algorithm presented in preceding section has been implemented in standard C, and interfaced with AMPL, the powerful mathematical programming language, which provides it with first and second derivative information using automatic differentiation, and allows the optimisation problems to be written in a simple algebraic notation [23]. The code was tested on a 2.4 GHz workstation with 512 KB of RAM, and 2.4Gb of memory. The test set consists of thirty-three minimax problems drawn from the CUTE set, available in AMPL format from the web-page detailed at [8]. The problem characteristics are summarised in Table 1 below, which uses the following abbreviations:

- GECO - General Convex problem,
- QCQO - Quadratically constrained problem,
- LO - Linear optimisation problem,
- NC - Nonconvex optimisation problem.
- No. of vars - Number of variables

- No. of constr - Number of constraints
- Var. Types - variable types
- Cons. types - constraint types

The algorithm parameters were set as follows. In Step 1, the accuracy of the stopping criteria is $\epsilon_0 = 10^{-6}$. In Step 2.3, ϵ_G was set to 10^{-4} and $\delta = 10$. In Step 2.4, we set $\theta = 0.9995, \beta = 0.5, \rho = 10^{-4}, m_1 = 1$ and $M_1 = 10$. In the barrier selection rule we set $\sigma = 6$. The significance of the choice of value of ϵ_G is discussed further below. All other parameters were chosen to be consistent with similar algorithms e.g LOQO.

Furthermore, the algorithm requires the Hessian of the Lagrangian to be positive definite at every iteration in order to achieve convergence. Indefinite Hessians at some iteration may cause the descent condition not to be satisfied and the algorithm may diverge. The treatment of indefinite systems is therefore critical. The algorithm deals with indefinite systems by perturbing the indefinite Hessian of the Lagrangian using the modified Cholesky factorisation described in [29] and [58]. Given a symmetric matrix A , modified Cholesky factorisation produces a perturbation E such that $A + E$ is positive definite, where E is a non-negative diagonal matrix that is zero if A is positive definite. The algorithm, taken from ([58]) is described in ([1])

The algorithm demonstrates fast local convergence, in terms of number of iterations, mainly due to the selection of the step-sizes and reduction of the barrier parameter. The numerical results are summarised in Table 2, where we use the following abbreviations:

- Itns : The total number of iterations required to find the optimal solution of the problem.
- c_0 : Initial value of the penalty parameter.
- c_* : Final value of the penalty parameter.
- k_* : The iteration after which the penalty parameter was unchanged.

The performance of the minimax algorithm in terms of number of iterations is provided, along with a comparison with the results obtained using LOQO and MINOS on the NEOS

server 4.0. LOQO is an infeasible primal dual interior point algorithm which uses a hybrid filter and merit function approach. MINOS uses a linear simplex method for linear problems, a reduced gradient method with quasi-Newton approximations for nonlinear objective functions, and sparse projected Lagrangian method for nonlinear constraints. The codes used for LOQO and MINOS were also interfaced with AMPL to use the exact second derivatives provided. We also demonstrate the attainment of the threshold values of the penalty term, c , beyond which the algorithm does not need to increase the penalty further. In addition, we compare the performance of algorithm using the single augmented Lagrangian merit function with the performance when both merit functions are used. Using the l_2 norm of the perturbed conditions as a single merit function may lead the algorithm to converge to a saddle point or a maximum when a minimum is sought, as illustrated by the examples in Appendix 1. The choice of value of ϵ_G is therefore significant to the performance of the algorithm - if ϵ_G is too large, the algorithm may switch too soon or too frequently to the secondary merit function with an increased chance of convergence to a saddle point or even a minimum when a maximum point is sought. The value 10^{-4} for the feasibility tolerance parameter ϵ_G was found to be large enough to trigger the switch to the secondary merit function in some test problems without causing it to converge to an incorrect solution point. It was therefore set to this value for our test problems in order to enable the comparison between the single and switching merit functions provided in Table 3.2. The starting points of the test problems are detailed in Appendix 2. To demonstrate the switch between the two merit functions, the progress of the algorithm for `chaconn1` and `mifflin1` is shown in Tables 3.3 and 3.4.

Table 3.1: The Test Set

Problem	No. of vars	No. of constr	Var. types	Problem type	Cons. types
cb2	2	3	all nlin	all nlin	GECO
cb3	2	3	all nlin	all nlin	GECO

Continued on Next Page...

Table 3.1 – Continued

Problem	No. of vars	No. of constr	Var. types	Problem type	Cons. types
chaconn1	2	3	all nlin	all nlin	GECO
chaconn2	2	3	all nlin	all nlin	GECO
congigmz	2	5	all nlin	2 nlin	QCQO
coshfun	60	20	all nlin	20 nlin	NC
demymalo	2	3	all nlin	1 nlin	QCQO
gigomecz1	2	3	all nlin	1 nlin	QCQO
goffin	50	50	all lin	all lin	LO
haifam	84	150	all nlin	all nlin	QCQO
haifas	6	9	all nlin	all nlin	QCQO
kiwcresc	2	2	all nlin	all nlin	QCQO
madsen	2	6	all nlin	all nlin	GECO
madsschj	80	158	all nlin	all nlin	QCQO
makela1	2	2	all nlin	1 nlin	QCQO
makela2	2	3	all nlin	all nlin	QCQO
makela3	20	20	all nlin	all nlin	QCQO
makela4	20	40	all lin	all lin	LO
mifflin1	2	2	all nlin	1 nlin	QCQO
mifflin2	2	2	all nlin	all nlin	QCQO
minmaxrb	2	4	1 nlin	2 nlin	QCQO
oet1	2	1002	all lin	all lin	LO
oet2	2	1002	all nlin	1000 nlin	NC
polak1	2	2	all nlin	all nlin	GECO
polak2	10	2	all nlin	2 nlin	GECO
polak3	11	10	all nlin	all nlin	QCQO
polak4	2	3	all nlin	all nlin	QCQO

Continued on Next Page...

Table 3.1 – Continued

Problem	No. of vars	No. of constr	Var. types	Problem type	Cons. types
polak5	2	2	all nlin	all nlin	GECO
polak6	4	4	all nlin	all nlin	GECO
pt	2	500	all lin	all lin	LO
rosenmmx	4	4	all nlin	all nlin	QCQO
spiral	2	2	all nlin	all nlin	GECO
womflet	2	3	all nlin	all nlin	GECO

In table 3.1, *lin* or *nlin* refers to the number of variables occurring linearly or nonlinearly in the objective function and/or constraints of the problem.

Table 3.2 Numerical results for single and double merit function minimax algorithms in comparison with LOQO.]

	MINOS		LOQO		Minimax algorithm					
					Single Merit Function			Switching Merit Function		
	Itns		Itns	c_0	Itns	c_*	k_*	Itns	c_*	k_*
cb2	11		8	0.1	8	6.39e11	8	8	1.71e8	6
cb3	2		9	0.1	9	4.53e11	9	9	6.94e8	7
chaconn1	24		9	0.1	9	3.08e12	9	9	3.7e8	7
chaconn2	2		9	0.1	9	4.53e11	9	9	6.94e8	7
congigmz	16		33	0.1	27	1.30e12	25	24	3.55e8	22
coshfun	450		22	0.0	48	1.81e11	27	48	2.693e6	21
demymalo	30		16	0.1	15	1.14e12	15	15	9.91e7	13
gigomecz1	40		16	0.1	15	2.03e10	15	15	1.75e8	14

Continued on Next Page...

Table 3.2 – Continued

	MINOS	LOQO	Minimax algorithm						
	Itns	Itns	c_0	Single Merit Function			Switching Merit Function		
				Itns	c_*	k_*	Itns	c_*	k_*
goffin	1	11	0.0	3	4.88e8	3	3	4.88e8	3
haifam	F	max	0.1	23	3.94e7	23	23	3.94e7	23
haifas	59	12	0.1	10	1.03e12	10	10	4.55e8	8
kiwcresc	17	14	10	11	9.45e10	11	11	8.51e8	10
madsen	27	11	0	23	7.42e8	19	19	1.91e6	18
madsschj	1083	22	0.0	29	3.81e5	29	29	3.81e5	29
makela1	26	14	1.0	13	1.69e12	13	13	1.19e8	11
makela2	21	12	0.1	11	2.63e9	11	11	7.5e7	10
makela3	22	11	0.1	10	1.65e2	8	10	1.65e2	8
makela4	1	11	0.1	4	0.1	1	4	0.1	1
mifflin1	24	9	0.1	9	2.81e12	9	9	2.52e8	7
mifflin2	11	10	1	10	4.34e10	10	12	6.2e7	8
minmaxrb	24	10	1	15	1e12	15	17	2.94e8	15
oet1	159	14	0.0	17	730.57	1	17	730.57	1
oet2	F	162	0.1	159	2.12e10	155	160	2.12e10	155
polak1	F	11	0.1	7	2.04e11	7	7	5.3e7	5
polak2	136	17	0.1	9	2.44e8	9	9	2.44e8	9
polak3	F	22	1	20	2.55e12	20	20	1	1
polak4	7	12	1	9	1.44e8	8	9	1.44e8	8
polak5	16	38	0.1	14	6.04e8	9	14	6.04e8	9
polak6	232	60	1	19	3.09e8	19	22	3.09e8	19
pt	1	14	0.0	31	1.06e12	31	32	3.77e8	29
rosenmmx	87	18	0.1	16	3.27e8	16	16	3.27e8	16
spiral	62	24	1	23	1.41e12	22	23	8.32e8	17

Continued on Next Page...

Table 3.2 – Continued

	MINOS	LOQO	c_0	Minimax algorithm					
				Single Merit Function			Switching Merit Function		
				Itns	c_*	k_*	Itns	c_*	k_*
womflet	35	88	0.1	15	9e11	15	15	8.88e8	13

In the reported numerical results for the minimax algorithm, in Table 3.2, a small choice for c_0 , the initial value of the penalty parameter, was often adequate. Convergence was achieved in all cases, although for some cases the final penalty parameter value was required to become large to achieve this in the case of the single merit function algorithm. This problem is tackled with the use of the secondary merit function algorithm, in which a switch to a second merit function is triggered once the two-norm of the equality constraints falls below the feasibility tolerance threshold. For most of the problems in Table 3.2, the minimax algorithms do not perform significantly worse than LOQO or MINOS. The exceptions to this are the linear problems (goffin, makela4, pt) for MINOS and problems coshfun and madsen for LOQO. MINOS and LOQO both have special features for solving linear problems. For the problems in the test set, the minimax algorithms most frequently provide solutions in the smallest number of iterations. For the test problems for which this is the case, the iteration counts for the minimax algorithms are in bold type in Table 2.

Itn Count	Merit ftn used	penalty parameter, c	Obj ftn	Variables
1	Aug. Lag.	10.1	3.94732	(1.00258, 0.28172)
2	Aug. Lag.	10.1	2.02875	(1.17320, 0.84020)
3	Aug. Lag.	10.1	2.00637	(1.12765, 0.88404)
4	Aug. Lag.	10.1	1.95879	(1.13755, 0.89775)
5	Aug. Lag.	2.6e4	1.95294	(1.13914, 0.89915)
6	Aug. Lag.	3.3e6	1.95230	(1.13907, 0.89950)
7	Aug. Lag.	3.7e8	1.95223	(1.13904, 0.89955)
8	l_2 norm	3.7e8	1.95222	(1.13904, 0.89956)
9	l_2 norm	3.7e8	1.95222	(1.13904, 0.89956)

Table 3.3: Progress of the algorithm for chaconn1.

Itn Count	Merit ftn used	penalty parameter, c	Obj ftn	Variables
1	Aug. Lag.	45.08	-0.04027	(0.99856, 0.98040)
2	Aug. Lag.	45.08	-0.654271	(1.13630, 0.43687)
3	Aug. Lag.	45.08	-0.87024	(1.11624, 0.00233)
4	Aug. Lag.	45.08	-0.982121	(1.0176, 0.00003)
5	Aug. Lag.	1.9e4	-0.99805	(1.00195, 0)
6	Aug. Lag.	2.1e6	-0.99980	(1.00020, 0)
7	Aug. Lag.	2.5e8	-0.99998	(1.00002, 0)
8	l_2 norm	2.5e8	-0.99999	(1, 0)
9	l_2 norm	2.5e8	-0.99999	(1, 0)

Table 3.4: Progress of the algorithm for mifflin1.

Part III

ALM

ALM Model Notation and Convention

- T planning horizon.
- $t = 1, \dots, T + 1$ time periods.
- $i = 1, \dots, I$ asset type.
- $k = 1, \dots, K$ risk-free instrument.
- Ω^t all distinct paths at time t .
- ω_t vector of stochastic data observed at time t .
- $\omega^t \equiv \omega_0, \dots, \omega_t$, data path history up to period t .
- $e = (\omega, t)$ index denoting an event (node of the scenario tree), which can be identified by an ordered pair of scenario and time period.
- $\mathcal{N}$ set of all nodes in the scenario tree.
- $\mathcal{N}_t$ set of nodes in the scenario tree representing possible events at time t .
- $\mathcal{N}_I \equiv \mathcal{N} - (\mathcal{N}_0 \cup \mathcal{N}_T)$ set of all interior nodes of the scenario.
- $a(e)$ ancestor of an event $e \in \mathcal{N}$ parent in the scenario tree.
- $D(e)$ set of descendants of an event $e \in \mathcal{N}$ parent in the scenario tree.
- p_e branching probability of node n at time t , i.e $Prob[(e)|a(e)]$.
- P_e probability of being at node n at time t : if $e = (n, t)$, then $P_e = \prod_{i=1}^T p_{(a(n,i))}$.
- r_{ie} return on asset i at node e .
- cb_{ie} transaction cost of buying asset i at node e .
- cs_{ie} transaction cost of selling asset i at node e .
- $\mathbf{cb}_e$ vector of transaction costs of buying assets at node e .

- cs_e vector of transaction costs of selling assets at node e .
- R_e^P return on underlying index/portfolio at node e .
- $s, t = 1, \dots, T + 1$ time periods.
- $\mathbf{r}_t(\omega^t)$ stochastic vector of return values for the n assets for $t = 1, \dots, T$.
- $\Lambda \in R^{n \times n}$ covariance matrix associated with the return values.
- $\Gamma \in R^{n \times n^2}$ transformed skewness tensor associated with the return values.
- $\mathbf{r}_e$ stochastic realisation of $\mathbf{r}_t$ at node e .
- $\hat{\mathbf{r}}_e$ expectation of $\mathbf{r}_t(\omega^t)$ at node e , conditional on ω^{t-1} , $\hat{\mathbf{r}}_e = E(\mathbf{r}(\omega^t | \omega^{t-1}))$.
- $\mathbf{u} \otimes \mathbf{v} \equiv (u_1 v_1, u_2 v_2, \dots, u_n v_n)'$.
- $\mathbf{u} \otimes \mathbf{v}' \equiv (u_1 \mathbf{v}', u_2 \mathbf{v}', \dots, u_n \mathbf{v}')' = \mathbf{u} \mathbf{v}'$.
- $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}' \equiv \mathbf{u} \otimes \mathbf{v} \mathbf{w}' = (u_1 \mathbf{v} \mathbf{w}', u_2 \mathbf{v} \mathbf{w}', \dots, u_n \mathbf{v} \mathbf{w}')'$.

Variables

- B_{ie} amount purchased of asset i at node e of the scenario tree.
- S_{ie} amount sold of asset i at node e of the scenario tree.
- H_{ie} amount held of asset i at node e of the scenario tree.
- $\mathbf{B}_e$ vector indicating amount purchased of assets at node e .
- $\mathbf{S}_e$ vector indicating amount sold of assets at node e .
- $\mathbf{H}_e$ vector indicating asset holdings.
- mb_e money borrowed at node e .
- ml_e money lent at node e .
- A_e asset value at node e

- L_e liability value at node e
- W_e surplus wealth at node e .
- y_e payments made due to lapse or death at node e .
- W^{ini} Initial wealth at $t = 0$.
- H_i^{ini} Initial holding in asset i at $t = 0$.
- m^{ini} Initial endowment at $t = 0$.

Chapter 4

The Asset-Liability Management Problem

4.1 The Asset Liability Management Problem Statement

This chapter introduces a stochastic optimisation model for an asset liability management problem. The ALM problem considered is for an equity-linked capital guaranteed asset. This type of product has gained in popularity since the stock market crash of 1987 and is most frequently issued by banks. Insurance companies sometimes also include financial guarantees and other options to their products to make them more attractive to investors. These products offer the security of capital protection as well as fixed rate of return and a part of any potential gains in an underlying asset over a fixed period.

We develop a scenario based ALM model for a guaranteed investment product offered by an insurance company, with bonus provisions. The product is an equity-index linked investment which allows the holder to receive part of the returns generated by an equity index or portfolio of assets over a specified period. Such products are often referred to as guaranteed investment contracts. The product offers a guaranteed minimum rate of return $\bar{g}$ per period, along with a bonus if the return of the portfolio exceeds the guarantee. The bonus is determined according to a participation rate α , which indicates the percentage of the portfolio return paid to policy-holders. At maturity, the guaranteed return and added

bonuses along with the original capital invested are returned to the holder of the product. The issuer of the product needs to ensure the asset allocation is capable of generating a surplus wealth at the end of the planning horizon to cover costs.

The model is based on the ALM model described in Consiglio, Cocco and Zenios [17], but differs in a number of ways. Firstly, we do not model the shortfall account. Secondly, the objective function in the model in [17] uses the logarithmic utility function to incorporate the investors risk attitude, whereas we focus on optimising specific moment measures of the portfolio terminal wealth distribution. As in [17], the portfolio optimisation models developed include linear transaction costs.

The premium collected forms the initial endowment m^{ini} which is allocated across the assets. Liabilities grow at a rate which is at least equal to the guarantee. Excess returns over $\bar{g}$ are returned to the policyholders according to the participation rate α . The dynamics of the liability account are given by

$$L_e = (1 - \kappa_e)L_{a(e)}(1 + \max[\alpha R_e^P, \bar{g}]),$$

where κ_e is the lapse probability of the policyholder as calculated by actuaries. The discontinuity caused by the $\max$ operator in the model can be circumvented by introducing variables π_e^+ and π_e^- to measure the portfolio excess return over and shortfall below the guarantee respectively. These variables satisfy

$$\begin{aligned} \alpha R_e^P - \bar{g} &= \pi_e^+ - \pi_e^- \\ \pi_e^+ &\geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0. \end{aligned} \tag{4.1}$$

During each period, the insurance company makes payments due to death or lapse given by

$$y_e = \kappa_e L_{a(e)}(1 + \bar{g} + \pi_e^+) \tag{4.2}$$

Constraints

In addition to those defined above, we impose the following constraints on the model:

Cash balance constraints: These constraints ensure zero net cash flow at every node.

$$\sum_{i=1}^I S_{ie}(1 - cs_{ie}) - \sum_{i=1}^I B_{ie}(1 + cb_{ie}) - y_e + mb_e - ml_e - (1 + r_{be}^f)mb_{a(e)} + (1 + r_{le}^f)ml_{a(e)} = 0, e \in N_I. \quad (4.3)$$

$$\sum_{i=1}^I S_{ie}(1 - cs_{ie}) - \sum_{i=1}^I B_{ie}(1 + cb_{ie}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in N_0. \quad (4.4)$$

Asset inventory constraint:

The asset inventory constraints give the amounts invested in each asset class at each node.

$$H_{ie} = (1 + r_{ie})H_{ia(e)} + B_{ie} - S_{ie}, e \in N_I. \quad (4.5)$$

$$H_{ie} = (1 + r_{ie})H_{ia(e)}, e \in N_T. \quad (4.6)$$

$$H_{ie} = H_i^{ini} + B_{ie} - S_{ie}, e \in N_0. \quad (4.7)$$

Asset weight restrictions: These limit the amounts invested in an asset class to be within a certain range, which may be a proportion of the entire fund.

$$H_{ie}^L \leq H_{ie} \leq H_{ie}^U, \quad (4.8)$$

$$B_{ie}^L \leq B_{ie} \leq B_{ie}^U,$$

$$S_{ie}^L \leq S_{ie} \leq S_{ie}^U,$$

$$e \in N, i = 1, \dots, I.$$

End of horizon constraints:

$$mb_e = 0, ml_e = 0, B_{ie} = 0, S_{ie} = 0, e \in \mathcal{N}_T. \quad (4.9)$$

Liability dynamics: These give the value of the liability at each node.

$$L_e = L_{a(e)}(1 + \bar{g} + \pi_e^+) \quad (4.10)$$

For convenience, we will adopt the notational convention used by de Athayde and Flores in ([18]) for the moment tensors. The second moment tensor is the $n \times n$ covariance matrix Λ . The third moments tensor is a $n \times n \times n$ cube. They transform this tensor into an $n \times n^2$ matrix by slicing each $n \times n$ layer and pasting them side by side in the same order. The resulting matrix form enables elegant formulas and differential calculus. Thus, if a general co-skewness is σ_{ijk} , for $n = 2$, the resulting matrix is

$$\begin{bmatrix} \sigma_{111} & \sigma_{112} & \sigma_{211} & \sigma_{212} \\ \sigma_{121} & \sigma_{122} & \sigma_{221} & \sigma_{222} \end{bmatrix}, \quad (4.11)$$

a 2×4 matrix with only four distinct elements.

Each stochastic realisation of r_t at node e can be expressed

$$\mathbf{r}_e = E(\mathbf{r}_t(\omega^t)) + \epsilon = \hat{\mathbf{r}}_e + \epsilon$$

where $\epsilon \sim Sk(0, \Lambda, \Gamma)$, such that the vector of return values at time t has a multivariate skewed distribution, with specified covariance matrix, Λ , and specified skewness tensor, Γ .

4.2 Scenario Generation

Scenario generation is an important part of the decision-making process for multi-period financial planning. Scenario-based stochastic models require an adequate representation of the underlying random data. This representation should be consistent with the assumptions regarding the probability distribution of the random data. A representation that is commonly

employed for stochastic problems is the scenario or event tree. The scenario tree is based on the assumption that there is a one-to-one correspondence between the sections and the nodes of the tree at stage t i.e. for any node at level t , each new observation of the random parameter must have only one immediate predecessor, but may have multiple successors. This is illustrated in the scenario tree in Figure 4.2 below. The stochastic programming approach determines the optimal decision for each node given the information available at that data point. The construction of a representative scenario tree is usually a fundamental problem in applications. For a fuller discussion on the requirements and a survey of scenario generation techniques, see Dupařova *et al* [19], or Kouwenberg and Zenios [47] for a fuller exposition on scenario generation for ALM and the advantages and disadvantages of the various methods. For a case study on a successful application of a global scenario generation system for ALM, see Mulvey and Thorlacius' paper on the Towers Perrin Scenario Generation system [55].

A class of methods for constructing trees are simulation techniques, discussed in Mulvey and Thorlacius [55] and Zenios [20], lattice models which are discussed for fixed-income portfolio models based on a underlying fine-grained interest-rate lattice (Jobst and Zenios, [38]). Kaut and Wallace [42] provide an overview of scenario generation methods and evaluate the suitability and quality of these methods on a given stochastic model.

In this report, we use the sequential optimisation method for moment matching scenario generation described in Gulpinar, Rustem and Settergren([34]). Let $\mathbf{x}$ denote a price vector, where $\mathbf{p}$ denotes the probability vector and $f_i(\mathbf{x}, \mathbf{p})$ is a mathematical expression of statistical property i , and $S V_i$ be the value of statistical property i . The scenario generator then solves a nonlinear optimisation problem of the form

$$\begin{aligned}
& \min_{\mathbf{x}, \mathbf{p}} \sum_i w_i (f_i(\mathbf{x}, \mathbf{p}) - S V_i)^2 \\
& \text{subject to} \\
& \sum_i p_i = 1, \\
& p_i \geq 0,
\end{aligned}$$

at each node of the scenario tree to generate probabilities and returns which match the first four moments of the historical data used as input for the scenario generator.

Once the tree has been generated, each scenario depicts a single realization of the stochastic equations over the defined planning horizon.

4.3 Data and Summary Statistics

The ALM portfolios are constructed from stock market indices of five regions: the US, Pacific (excluding Japan), Japan, Europe (excluding the UK) and the UK. These indices are proxied by the MSCI annual price index for these regions. We take the view of a US investor, thus the index to which the portfolio return is linked is the MSCI US index and investments are in US dollars. The summary statistics for the annual returns of the indices for the period 1971-2003, and the three sub-periods 1971-1981, 1982-1992 and 1993-2003 used in the rolling window estimation are provided in tables 3.1 to 4.3 below.

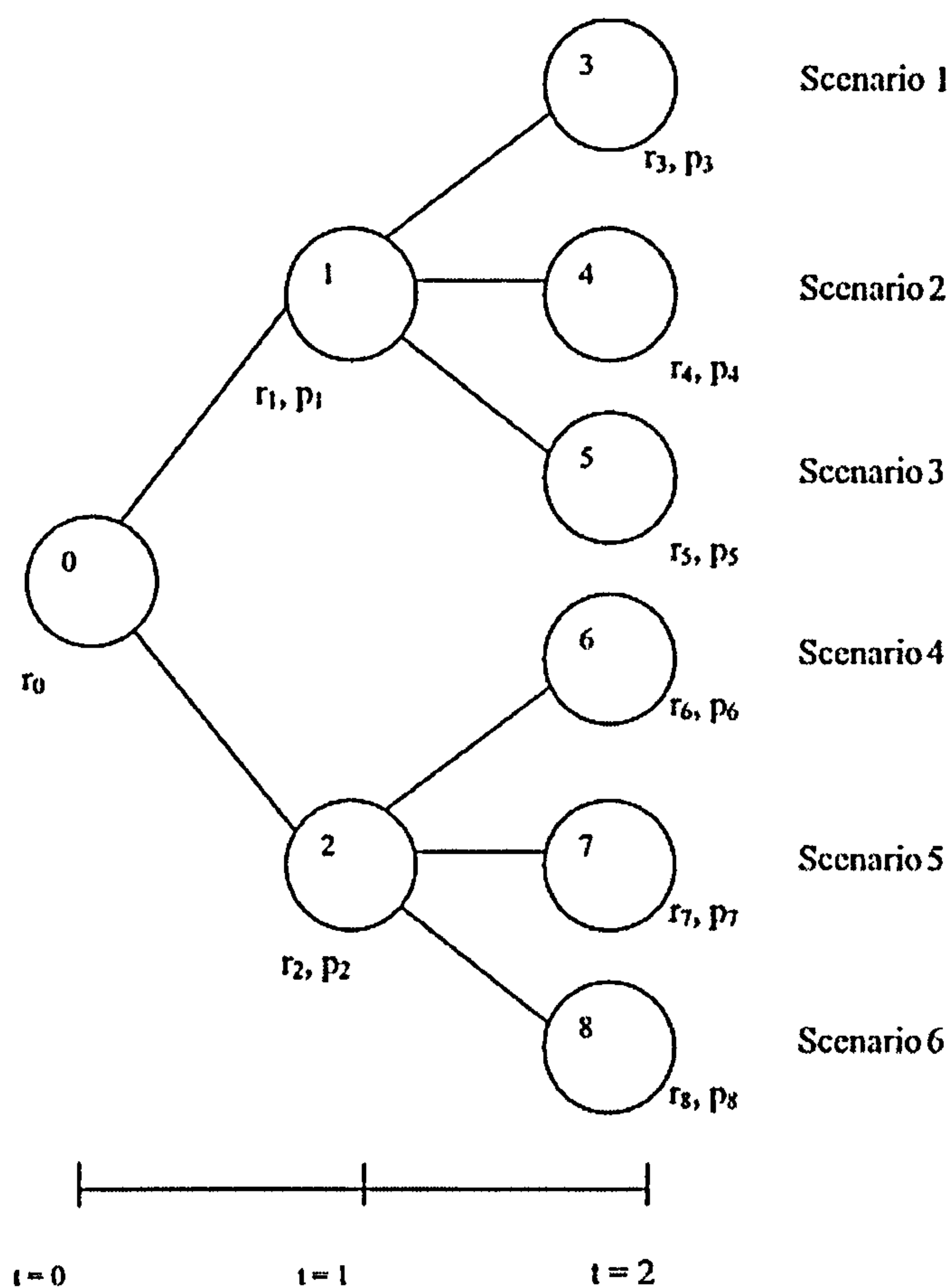


Figure 4.1: A two-period scenario tree, with 9 nodes and 6 scenarios .

4.3.1 Empirical Analysis: Evidence of Non-Normality and Skewness

The assumption of normal asset returns was assessed for the five regional MSCI indices using the Shapiro-Wilks normality test procedure in the *Analyze-it* package that is available as an excel-add on. The Shapiro-Wilks test is generally accepted to be better to small sample sizes (< 50). The Shapiro-Wilks Test, which was proposed in [71], calculates a test statistic, small values of which indicate a departure from normality. The Shapiro-Wilks' test is the preferred test of normality because of its good power properties as compared to a wide range of alternative tests [72]. The Shapiro-Wilks test rejects the assumption of normality for the US and UK indices for the 1971-2003 period and the UK and Europe ex UK for the

1993-2003 period at the 5% significance level (i.e. $p \leq 0.05$). The results of the normality tests are provided below.

Summary Statistics (1971-2003)					
	US	Pacific ex Japan	Japan	Europe ex UK	UK
Mean	0.053	0.018	0.048	0.054	0.033
Std. Dev.	0.171	0.28	0.274	0.182	0.287
Skew. Coeff.	-1.139	-0.779	-0.391	0.015	-2.391
Kurt. Coeff.	1.027	0.568	-0.502	-0.332	9.353

Table 4.1: Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1971 to 31 December 2003 (33 observations).

Summary Statistics (1971-1981)					
	US	Pacific ex Japan	Japan	Europe ex UK	UK
Mean	-0.005	0.026	0.122	0.022	-0.036
Std. Dev.	0.206	0.361	0.252	0.127	0.468
Skew. Coeff.	-1.019	-1.158	-0.217	-0.254	-1.57
Kurt. Coeff.	0.588	1.137	-0.451	-1.078	3.158

Table 4.2: Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1971 to 31 December 1981 (11 observations).

Summary Statistics (1982-1992)					
	US	Pacific ex Japan	Japan	Europe ex UK	UK
Mean	0.103	0.044	0.08	0.093	0.103
Std. Dev.	0.094	0.223	0.298	0.207	0.113
Skew. Coeff.	-0.318	-1.165	-1.06	0.556	0.417
Kurt. Coeff.	-1.004	1.328	1.311	-0.543	-0.062

Table 4.3: Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1982 to 31 December 1992 (11 observations).

Summary Statistics (1993-2003)					
	US	Pacific ex Japan	Japan	Europe ex UK	UK
Mean	0.065	0.001	-0.038	0.068	0.044
Std. Dev.	0.196	0.272	0.275	0.205	0.165
Skew. Coeff.	-0.877	-0.048	-0.076	-1.033	-0.718
Kurt. Coeff.	-0.543	-0.540	-1.322	-0.243	-1.37

Table 4.4: Year-end US dollar denominated returns on the equity indices of five regions from MSCI for the sample period December 31 1993 to 31 December 2003 (11 observations).

Normality Test (1971-2003)		
Region	Shapiro-Wilks Coeff.	p-value
US	0.9007	0.0048
Pacific ex. Japan	0.9510	0.1315
Japan	0.9704	0.4721
Europe ex. UK	0.9795	0.7557
UK	0.8035	<0.0001

Table 4.5: Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1971-2003

Normality Test (1971-1981)		
Region	Shapiro-Wilks Coeff.	p-value
US	0.9177	0.3001
Pacific ex. Japan	0.8858	0.1232
Japan	0.9423	0.5473
Europe ex. UK	0.9421	0.5450
UK	0.8734	0.0857

Table 4.6: Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1971-1981

Normality Test (1982-1992)		
Region	Shapiro-Wilks Coeff.	p-value
US	0.9235	0.3491
Pacific ex. Japan	0.9091	0.2380
Japan	0.9257	0.3692
Europe ex. UK	0.9385	0.5032
UK	0.9781	0.9544

Table 4.7: Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1982-1992

Normality Test (1993-2003)		
Region	Shapiro-Wilks Coeff.	p-value
US	0.8646	0.0660
Pacific ex. Japan	0.9601	0.7732
Japan	0.9488	0.6283
Europe ex. UK	0.8516	0.0448
UK	0.8311	0.0241

Table 4.8: Results of the Shapiro-Wilks test for the regional MSCI US dollar equity index returns for 1993-2003

Chapter 5

Robust Mean-Variance Asset Liability Management

5.1 Introduction

In this section, we describe an extension of the Markowitz model to the multi-period ALM problem of the previous section. We then present the robust variant of this MV model, in the presence of data uncertainty represented by rival scenario trees. A method of generating rival scenario trees is applied. The mean-variance and robust mean variance models are implemented in AMPL and back-tested. The results of the back-testing are also provided.

5.2 The Mean-Variance ALM Model

The total asset holding in each period is given by

$$A_t = \sum_i r_{it} H_{i(t-1)} = \mathbf{r}'_t \mathbf{H}_{t-1}$$

so the surplus wealth at time t after payment of the liability is

$$W_t = A_t - L_t.$$

The expected surplus wealth at time t after payment of the liability is then

$$\overline{W}_t = E[A_t - L_t] = E[W_t].$$

The variance of the terminal wealth is expressed as :

$$\begin{aligned} Var(W_T) &= E[W_T - \overline{W}_T]^2 \\ &= E[(A_T - L_T)^2] - [E(A_T - L_T)]^2 \\ &= E[(\mathbf{r}'_T \mathbf{H}_{T-1} - L_T^2)] - (E[(\mathbf{r}'_T \mathbf{H}_{T-1} - L_T)])^2 \\ &= E[(\mathbf{H}'_{T-1}(\mathbf{r}_T \mathbf{r}'_T) \mathbf{H}_{T-1} - 2\mathbf{r}'_T \mathbf{H}_{T-1} L_T + L_T^2)] - (E[(\mathbf{r}'_T \mathbf{H}_{T-1} - L_T)])^2. \end{aligned}$$

Given the possible events $e \in \mathcal{N}_t$, the expected wealth at time t is

$$\overline{W}_t = E \sum_{e \in \mathcal{N}_t} P_e W_e = E \sum_{e \in \mathcal{N}_t} P_e (A_e - L_e) = E \sum_{e \in \mathcal{N}_t} P_e (\mathbf{r}'_e \mathbf{H}_{a(e)} - L_e) = \sum_{e \in \mathcal{N}_t} P_e (\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e).$$

We can therefore represent the variance of the terminal wealth as:

$$\begin{aligned} Var(W_T) &= E[W_T - \overline{W}_T]^2 \\ &= E \sum_{e \in \mathcal{N}_T} P_e [W_e - \overline{W}_e]^2 = E \sum_{e \in \mathcal{N}_T} P_e [W_e^2] - [\overline{W}_e]^2 \\ &= \sum_{e \in \mathcal{N}_T} P_e [(\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e)]^2 - \left[\sum_{e \in \mathcal{N}_T} P_e (\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e) \right]^2, e \in \mathcal{N}_T \\ &= \sum_{e \in \mathcal{N}_T} P_e [\mathbf{H}'_{a(e)} (\hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e + \Lambda) \mathbf{H}_{a(e)} - 2\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} L_e + L_e^2] - \left[\sum_e P_e (\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e) \right]^2. \end{aligned} \tag{5.1}$$

The components of $\hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e + \Lambda$ in (5.1) reflect the risk arising from two types of uncertainty. The outer product term, $\hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e$, is the uncertainty due to the discrete distribution of realisations of ω^t . The continuous uncertainty due to the variability of the returns at a given realisation of ω^t is represented by the Λ term in (5.1), (see also [24] and [35]).

Given this, the ALM extension of the Markowitz model is:

NLP1

$$\begin{aligned}
& \min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \sum_{e \in \mathcal{N}_T} P_e [W_e - \overline{W}_e]^2 \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e - (1 + r_{be}^f)mb_{a(e)} + \\
& \quad (1 + r_{le}^f)ml_{a(e)} = 0, e \in \mathcal{N}_I, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, e \in \mathcal{N} \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{a(e)}(1 + \bar{g} + \pi_e^+), e \in \mathcal{N}_T, \\
& \alpha R_e^P - \bar{g} = \pi_e^+ - \pi_e^-, \\
& \pi_e^+ \geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_e W_e = \mathcal{W}, \\
& W_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W}_e = \sum_{e \in \mathcal{N}_T} P_e W_e.
\end{aligned}$$

5.3 Generation of Rival Scenario Trees

A number of different approaches have been used in the specification of parameter uncertainty in robust optimisation and estimation problems. In this analysis, the uncertainty is specified by rival scenario trees obtained from the uncertainty sets of the return distribu-

tion parameters. These uncertainty sets are generated by using historical data from different sub-periods as input for the parameter estimation in the scenario generator. The scenario generator uses the moment matching method of [34] described in the previous section to obtain the asset returns, scenario probabilities as well as other data for the optimisation model. The sub-periods are chosen from a rolling window of the observed data. This is known as a rolling window estimation. These uncertainty sets are then input to the scenario generator to create a number of alternative scenario trees for the optimisation problem described above. While keeping the data length fixed, we use rolling periods of historical data as input to the scenario generator. The method involves running simulations based on rolling time horizon length periods of historical data in chronological order. The iterations of the rolling scenario generation model a series of sequentially time-shifted periods that start after the previous period until all the historical market data has been traversed to run the scenario generation. The uncertainty structure generation approach can be outlined as follows:

- Step 1: Given an optimisation problem with a time horizon T , and a historical data set of length τ divide the historical data set, H_0 into τ/T data sets in chronological order.
- Step 2: Run τ/T simulations to create scenario trees using each of the data sets created in Step 1 as input.
- Step 3: Use the scenario tree output data to perform the robust portfolio optimisation described in the previous section.

This process is similar to the method used by Tütüncü and Koenig [76] to generate uncertainty sets for short and medium-term horizons. We choose the rolling window size to be a third of the data, thereby generating uncertainty sets from three non-overlapping periods.

5.4 The Robust Mean-Variance ALM Model

Now, we assume we have a number of rival scenario trees for the stochastic variable, and we are uncertain which is best. Each tree is defined by the pair $(\mathbf{r}_{je}, p_{je})$, where $e \in \mathcal{N}$,

$j = 1, \dots, J$. We obtain rival means and variances based on the rivalling scenario tree data. Thus, we extend the mean-variance model to robust minimax optimisation over the set of rival means and variances describing the probabilistic distribution of the stochastic variables of the asset liability management models. Given the return vectors $\mathbf{r}_{je} \in R^n$, and the node probabilities p_{je} , $j = 1, \dots, J$, the robust extension of the mean-variance ALM model can thus be written as

RobNLP1

$$\min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \mathcal{V}$$

subject to

$$\mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_{je} + mb_{je} - ml_{je} + m^{ini} = 0, e \in \mathcal{N}_0,$$

$$\mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0,$$

$$\mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, j = 1, \dots, J,$$

$$\mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e - (1 + r^f_{jbe})mb_{a(e)} + \\ (1 + r^f_{jle})ml_{a(e)} = 0, e \in \mathcal{N}_I, j = 1, \dots, J,$$

$$\mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, j = 1, \dots, J,$$

$$\mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, e \in \mathcal{N},$$

$$\mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, e \in \mathcal{N},$$

$$\mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, e \in \mathcal{N},$$

$$mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_{je} = 0, e \in \mathcal{N}_T,$$

$$L_{je} = (1 - \kappa_e)L_{ja(e)}(1 + \bar{g} + \pi^+_{je}), j = 1, \dots, J,$$

$$\alpha R_e^P - \bar{g} = \pi^+_{je} - \pi^-_{je}, j = 1, \dots, J,$$

$$\pi^+_{je} \geq 0, \pi^-_{je} \geq 0, \pi^+_{je}\pi^-_{je} = 0, e \in \mathcal{N},$$

$$\sum_{e \in \mathcal{N}_T} P_{je} W_{je} = \mathcal{W}, j = 1, \dots, J,$$

$$\sum_{e \in \mathcal{N}_T} P_{je} [W_{je} - \overline{W_{je}}]^2 \leq \mathcal{V}, j = 1, \dots, J,$$

$$W_{je} = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} - L_{je}, e \in \mathcal{N}_T,$$

$$\overline{W_{jT}} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1, \dots, J.$$

5.5 Numerical Experiments

Given a set of three rivaling scenario trees (i.e. $J = 3$) with return data for five assets and setting the expected terminal wealth, $\mathcal{W}$, to a given value, we do the following:

- solve problem NLP1 using each tree as input,
- evaluate the results of holding the optimal portfolio for tree j when the worst-case tree is valid,
- apply the minimax approach to problem RobNLP1 for robust optimisation over the set of trees.

The results are presented in the figures below. Figure 5.1 compares the minimax efficient frontier with the efficient frontiers computed using each of the rival scenario trees. Figure 5.2 illustrates the minimax frontier, which represents the worst-expected return for our set of rival scenario trees, for a given level of risk.

Whichever of the scenario trees is valid, the return obtained is at least as large as the minimax return. The scenario tree which produces the worst levels of return varies at different levels of risk. At the lowest levels of risk on the minimax frontier, Scenario Tree 1 produces the worst expected return. As the level of risk increases, Scenario Tree 2 becomes the worst-case. At the highest levels of risk, the worst expected return is yielded by Scenario Tree 3. The minimax strategy gives the portfolio holdings which maximise the worst expected return for a given level of risk (specified by the variance). Figure 5.2 also shows the performance of the minimax optimal portfolio holdings for each of the three scenario trees i.e. if the minimax optimal portfolio is held when Scenario Tree j is valid for $j = 1, 2, 3$.

Figure 5.3 compares the minimax frontier to the mean-risk profiles obtained if the optimal portfolio holdings for Scenario Trees 1, 2 and 3 and the worst-case tree is valid.

Figures 5.2 and 5.3 illustrate the non-inferiority of minimax i.e. the performance of the portfolio is guaranteed to improve should any of the scenario trees other than the worst-case scenario tree be actualised instead. However, should the worst-case be realised when the portfolio has been optimised in view of Scenario Tree 1, 2 or 3, 5.3 illustrates the resulting inferior portfolios.

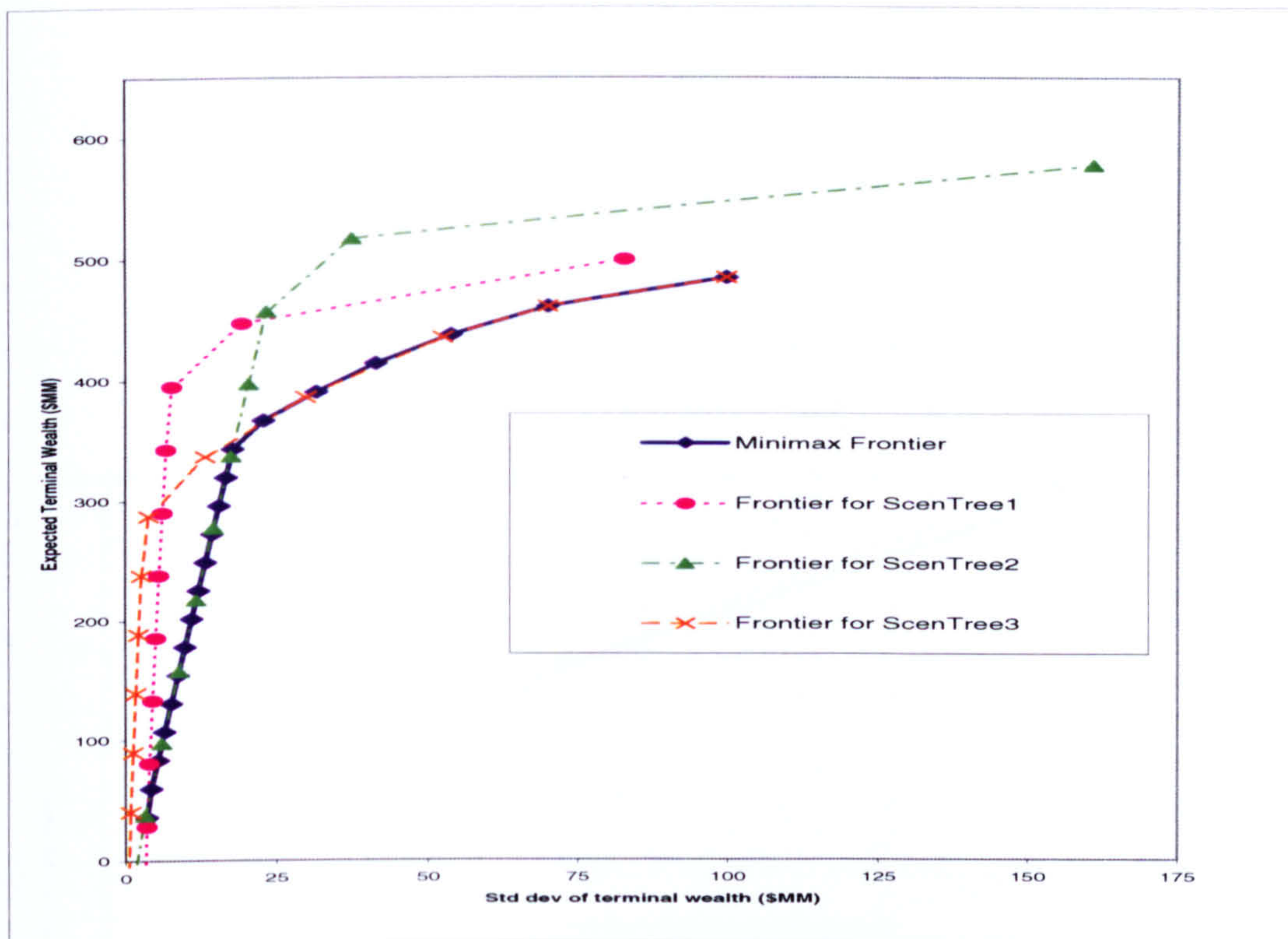


Figure 5.1: The minimax efficient frontier compared to the efficient frontiers computed using each of the three scenario trees.

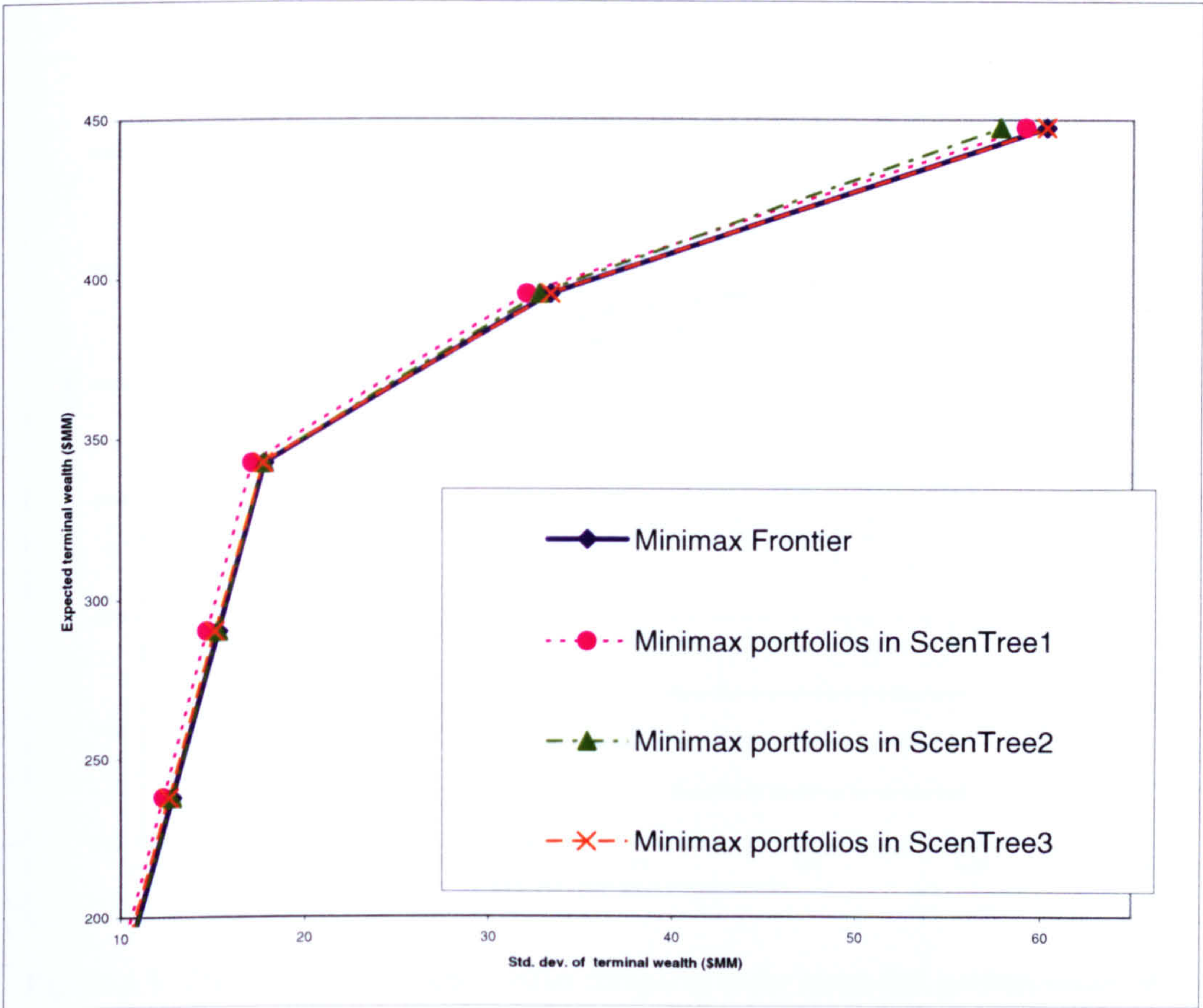


Figure 5.2: Noninferiority of minimax: A clipping of the minimax frontier and the mean-risk profile of the minimax optimal portfolio holdings evaluated using the three scenario trees. Performance improves if the worst-case scenario is not realized.

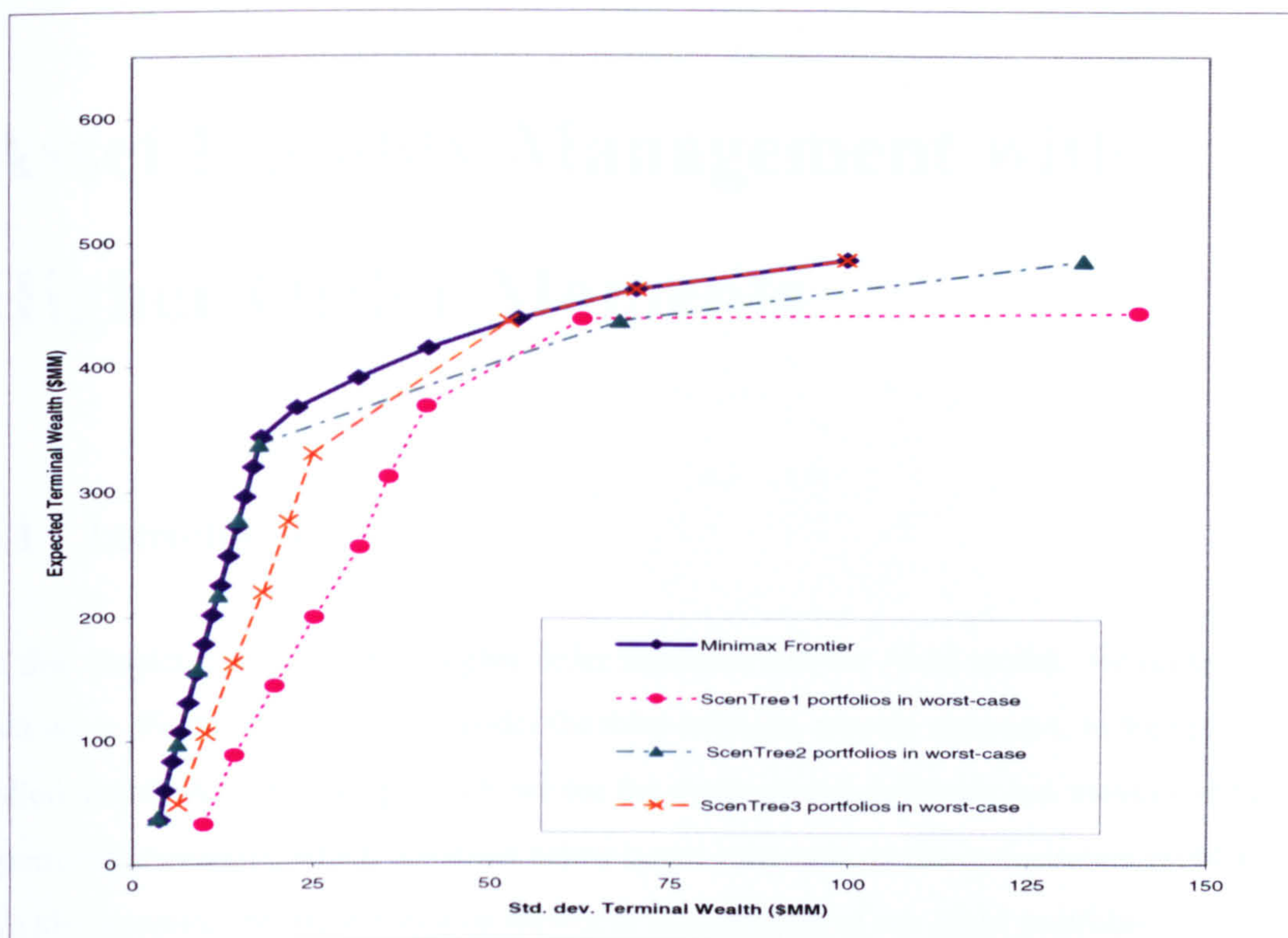


Figure 5.3: The minimax efficient frontier compared to the mean-risk profiles obtained from portfolio holdings selected on the efficient frontiers for scenario trees 1, 2 and 3 when the worst-case tree is valid. Performance deteriorates for all other Scenario Trees if the worst-case scenario tree is valid

Chapter 6

Asset Liability Management with Higher Order Moments

6.1 Introduction

In this chapter, we incorporate higher order moments into the ALM model. We do this in two ways. Firstly, we directly consider the third-moment, namely skewness, in the optimisation model. As a second approach, we use the more general downside risk measure in the optimisation models, which penalises below target variations, in the optimisation problem. We also examine the mean-variance-skewness relationships of our ALM portfolio.

6.2 The Mean-Variance-Skewness (MVS) ALM Model

The third moment measure, the skewness, of the terminal wealth can be defined in a similar manner to the variance as:

$$\begin{aligned} Skew(W_T) &= E[W_T - \overline{W_T}]^3 \\ &= E[W_T^3] - 3E[W_T^2]\overline{W_T} + 3E[W_T]\overline{W_T}^2 - \overline{W_T}^3 \\ &= E[W_T^3] - 3E[W_T^2]\overline{W_T} - 2\overline{W_T}^3. \end{aligned}$$

where

$$E[W_T^3] = E[\mathbf{H}'_{(T-1)}(\mathbf{r}'_T \otimes \mathbf{r}_T \mathbf{r}'_T) \mathbf{H}_{(T-1)} \otimes \mathbf{H}'_{(T-1)} - 3\mathbf{H}'_{(T-1)}(\mathbf{r}_T \mathbf{r}'_T) \mathbf{H}_{(T-1)} L_T + 3\mathbf{r}'_T \mathbf{H}_{(T-1)} L_T^2 + L_T^3]$$

Given the possible events $e \in \mathcal{N}_t$, the skewness of the terminal wealth can be expressed as

$$\begin{aligned} \text{Skew}(W_T) &= E[W_T - \overline{W_T}]^3 \\ &= E \sum_{e \in \mathcal{N}_T} P_e [W_e - \overline{W_T}]^3 \\ &= E \sum_{e \in \mathcal{N}_T} P_e [W_e^3 - 3W_e^2 \overline{W_T} + 3W_e \overline{W_T}^2 - \overline{W_T}^3] \\ &= E \left[\sum_{e \in \mathcal{N}_T} P_e [W_e^3] - 3 \sum_{e \in \mathcal{N}_T} P_e [W_e^2] \overline{W_T} + 3 \sum_{e \in \mathcal{N}_T} P_e W_e \overline{W_T}^2 - \overline{W_T}^3 \right] \\ &= E \left[\sum_{e \in \mathcal{N}_T} P_e [W_e^3] - 3 \sum_{e \in \mathcal{N}_T} P_e [W_e^2] \overline{W_T} + 2\overline{W_T}^3 \right]. \end{aligned}$$

where

$$\begin{aligned} E \sum_e P_e [W_e^3] &= \sum_e P_e [\mathbf{H}'_{a(e)}((\mathbf{r}_e \otimes \mathbf{r}_e \mathbf{r}'_e) + \Gamma) \mathbf{H}_{a(e)} \otimes \mathbf{H}'_{a(e)} - 3\mathbf{H}'_{a(e)}(\mathbf{r}_e \mathbf{r}'_e + \Lambda) \mathbf{H}_{a(e)} L_e \\ &\quad + 3\mathbf{r}'_e \mathbf{H}_{a(e)} L_e^2 + L_e^3] \end{aligned}$$

As before, Λ reflects the variability of the returns at a given realisation of ω' . The skewness of the returns at a given realisation of ω' is reflected by the Γ term.

The skewness can thus be expressed as:

$$\begin{aligned} \text{Skew}(W_T) &= E[W_T - \overline{W_T}]^3 \\ &= \sum_{e \in \mathcal{N}_t} P_e [\mathbf{H}'_{a(e)}((\hat{\mathbf{r}}'_e \otimes \hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e) + \Gamma) \mathbf{H}_{a(e)} \otimes \mathbf{H}'_{a(e)} - 3\mathbf{H}'_{a(e)}(\hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e + \Lambda) \mathbf{H}_{a(e)} L_e + 3\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} L_e^2 + L_e^3] \\ &\quad - 3 \sum_{e \in \mathcal{N}_t} P_e [\mathbf{H}'_{a(e)}(\hat{\mathbf{r}}_e \hat{\mathbf{r}}'_e + \Lambda) \mathbf{H}_{a(e)} - 2\hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} L_e + L_e^2] \overline{W_e} + 2\overline{W_e}^3. \end{aligned}$$

Hence the multistage asset liability management problem can be expressed as the maximization

sation of the skewness of the terminal wealth subject to (4.3)-(4.10):

ALMHOMSKEW:

$$\begin{aligned}
& \max_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \sum_{e \in \mathcal{N}_T} P_e [W_e - \overline{W}_T]^3 \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e - (1 + r_{be}^f)mb_{a(e)} + \\
& (1 + r_{le}^f)ml_{a(e)} = 0, e \in \mathcal{N}_I, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, e \in \mathcal{N} \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{a(e)}(1 + \bar{g} + \pi_e^+), e \in \mathcal{N}_T, \\
& \alpha R_e^P - \bar{g} = \pi_e^+ - \pi_e^-, \\
& \pi_e^+ \geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_e W_e \geq \mathcal{W}, \\
& \sum_{e \in \mathcal{N}_T} P_e [W_e - \overline{W}_T]^2 = \mathcal{V}, \\
& W_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W}_T = \sum_{e \in \mathcal{N}_T} P_e W_e.
\end{aligned}$$

6.3 The Mean-Downside Risk (MDR) ALM Model

Downside risk models seek to minimise the variation of the terminal wealth below a specified target. For a target wealth θ , we can write a downside risk formulation of the ALM problem as:

$$\begin{aligned}
& \min_{\{H, B, S\}} \sum_{e \in N_T} P_e [\max(0, \theta - W_e)]^u \\
& \text{subject to} \\
& S'_e(1 - c_{se}) - B'_e(1 + c_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in N_0, \\
& H_e = H^{ini} + B_e - S_e, e \in N_0, \\
& H_e = \hat{r}'_e H_{a(e)} + B_e - S_e, e \in N_I, \\
& S'_e(1 - c_{se}) - B'_e(1 + c_{be}) - y_e + mb_e - ml_e - (1 + r_{be}^f)mb_{a(e)} + \\
& (1 + r_{le}^f)ml_{a(e)} = 0, e \in N_I, \\
& H_e = \hat{r}'_e H_{a(e)}, e \in N_T, \\
& H_e^L \leq H_e \leq H_e^U, \\
& B_e^L \leq B_e \leq B_e^U, \\
& S_e^L \leq S_e \leq S_e^U, e \in N \\
& mb_e = 0, ml_e = 0, B_e = 0, S_e = 0, e \in N_T, \\
& L_e = (1 - \kappa_e)L_{a(e)}(1 + \bar{g} + \pi_e^+), e \in N_T, \\
& \alpha R_e^P - \bar{g} = \pi_e^+ - \pi_e^-, \\
& \pi_e^+ \geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0, e \in N, \\
& \sum_{e \in N_T} P_e W_e \geq \mathcal{W}, \\
& W_e = \hat{r}'_e H_{a(e)} - L_e, e \in N_T, \\
& \overline{W_T} = \sum_{e \in N_T} P_e W_e.
\end{aligned}$$

The power, u , determines the order of the lower-partial moment function in the objective. If $\theta = \sum_{e \in N_T} P_e W_e$, and $u = 2$, then the risk measure is the Lower Semi-Variance (LSV). For a downside risk formulation of the ALM problem in which we seek to minimise the lower semi-variance with respect to the terminal liability, we can express the problem as:

$$\begin{aligned}
& \min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \sum_{e \in \mathcal{N}_T} P_e [\max(0, \overline{W}_T - W_e)]^2 \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e - (1 + r_{be}^f)mb_{a(e)} + \\
& (1 + r_{le}^f)ml_{a(e)} = 0, e \in \mathcal{N}_I, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, e \in \mathcal{N} \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{a(e)}(1 + \bar{g} + \pi_e^+), e \in \mathcal{N}_T, \\
& \alpha R_e^P - \bar{g} = \pi_e^+ - \pi_e^-, \\
& \pi_e^+ \geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_e W_e \geq \mathcal{W}, \\
& W_e = \hat{\mathbf{r}}'_e \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W}_T = \sum_{e \in \mathcal{N}_T} P_e W_e.
\end{aligned}$$

The function $\max(0, \overline{W}_T - W_e)$ is discontinuous, and cannot be solved by standard non-linear solvers. For computational ease, the downside risk objective function can be reformulated by the introduction of artificial variables v_{je}^-, v_{je}^+ such that

$$\overline{W}_T - W_e = v_e^- - v_e^+, v_e^- \geq 0, v_e^+ \geq 0.$$

Thus, the downside risk ALM model becomes:

ALMHOMDSR

$$\begin{aligned}
& \min_{\{\mathbf{H}, \mathbf{B}, \mathbf{S}\}} \sum_{e \in \mathcal{N}_T} P_e (y_e)^2 \\
& \text{subject to} \\
& S'_e(1 - c_{se}) - B'_e(1 + c_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& H_e = H^{ini} + B_e - S_e, e \in \mathcal{N}_0, \\
& H_e = \hat{r}'_e H_{a(e)} + B_e - S_e, e \in \mathcal{N}_I, \\
& S'_e(1 - c_{se}) - B'_e(1 + c_{be}) - y_e + mb_e - ml_e - (1 + r^f_{be})mb_{a(e)} + \\
& (1 + r^f_{le})ml_{a(e)} = 0, e \in \mathcal{N}_I, \\
& H_e = \hat{r}'_e H_{a(e)}, e \in \mathcal{N}_T, \\
& H_e^L \leq H_e \leq H_e^U, \\
& B_e^L \leq B_e \leq B_e^U, \\
& S_e^L \leq S_e \leq S_e^U, e \in \mathcal{N} \\
& mb_e = 0, ml_e = 0, B_e = 0, S_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{a(e)}(1 + \bar{g} + \pi_e^+), e \in \mathcal{N}_T, \\
& \alpha R_e^P - \bar{g} = \pi_e^+ - \pi_e^-, \\
& \pi_e^+ \geq 0, \pi_e^- \geq 0, \pi_e^+ \pi_e^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_e W_e \geq \mathcal{W}, \\
& W_e = \hat{r}'_e H_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W_T} = \sum_{e \in \mathcal{N}_T} P_e W_e. \\
& \sum_{e \in \mathcal{N}_T} P_e W_e \geq \mathcal{W}, \\
& \overline{W_T} - W_e = v_e^- - v_e^+, \\
& v_e^- \geq 0, v_e^+ \geq 0.
\end{aligned}$$

6.4 A Global Optimisation Heuristic

The mean-variance-skewness model leads to a non-convex optimisation problem. For this reason, the results generated from its solution may not be globally optimal unless a global optimiser is used. In order to find at least a good local optimum, we simulate a rigorous

global search by using multiple starts of the local optimiser at interval points defined over the search space. For a vector of decision variables denoted by X , let X^{term} be the value of the decision variables at the end of a single run of the local optimiser. The simple heuristic algorithm is described below for a minimisation problem:

Step 1. Set $k = 0$, $Starts = 1000$, $X_k^0 = InitialValue$, $GlobalOptimum := M$. Set X_{Max} and X_{Min} . Let $Stepsize = (X_{Max} - X_{Min})/Starts$.

Step 2. Do while $k < Starts$

Call local optimisation algorithm, e.g. Algorithm 1. Solve optimisation problem. Set $k = k + 1$.

Step 3. If optimum found and $Objective \leq GlobalOptimum$,

let $GlobalOptimum = Objective$. GOTO Step 4.

Else

set $X_k^0 = X_{k-1}^{term}$ and GOTO Step 2 above.

Step 4. Set $X_k^0 = X_{k-1}^0 + Stepsize * k$. GOTO Step 2.

Algorithm 2.

If the solver fails, we warm-start it using the decision vector at the final iteration of the previous run as the initial point. If the problem is solved to optimality, we check it against the current best value of the objective function. If it is better, this becomes the new global optimum. The initial value, M , for the global optimum may be chosen to be the first value of the objective function obtained from a single run of the local optimiser on the problem. X_{Max} and X_{Min} are chosen from the bound constraints on the decision variables.

6.5 Numerical Experiments

6.5.1 The Mean-Variance-Skewness Relationship of the ALM Problem

The ALM Models described above were implemented in AMPL. The relationship between the mean, variance and skewness of the portfolio terminal wealth was then plotted by:

- Minimising the variance of the terminal wealth for a given level of expected terminal wealth, (to plot the Markowitz mean-variance frontier).
- Maximising the skewness of the terminal wealth for a given level of expected terminal wealth, (to plot the mean-skewness tradeoff).
- Maximising the skewness of the terminal wealth for a given level of variance, (to plot the variance-skewness tradeoff).

For the plots involving skewness, the global optimisation heuristic described in Section 6.4 was used to tackle the resulting non-convexity in the optimisation problems.

The familiar upward sloping convex curve in Figure 6.1 shows the positive trade-off between the variance (or standard deviation) and the expected terminal wealth i.e. for any level of skewness, the higher the expected terminal wealth the higher the variance/standard deviation of the portfolio.

The curve in Figure 6.2 shows the negative trade-off between the expected terminal wealth and the skewness. That is, low skewness portfolios have a higher expected terminal wealth in the positive skewness range. In the negative skewness range, as the negative skewness decreases the expected terminal wealth increases.

Figure 6.3 shows the positive trade-off between the skewness and the variance, similar to the mean-variance trade-off. Achieving greater positive skewness requires taking on greater risk in terms of variance or standard deviation. In the negative skewness range, higher levels of variance or standard deviation are associated with higher negative skewness.

Varying the guaranteed rate of the ALM from $g = 0\%$ to 12% has little effect on the mean-variance and skewness-variance frontiers. However, it introduces non-smoothness to the mean-skewness frontier of the ALM portfolio, as can be seen from Figure 6.4 below.

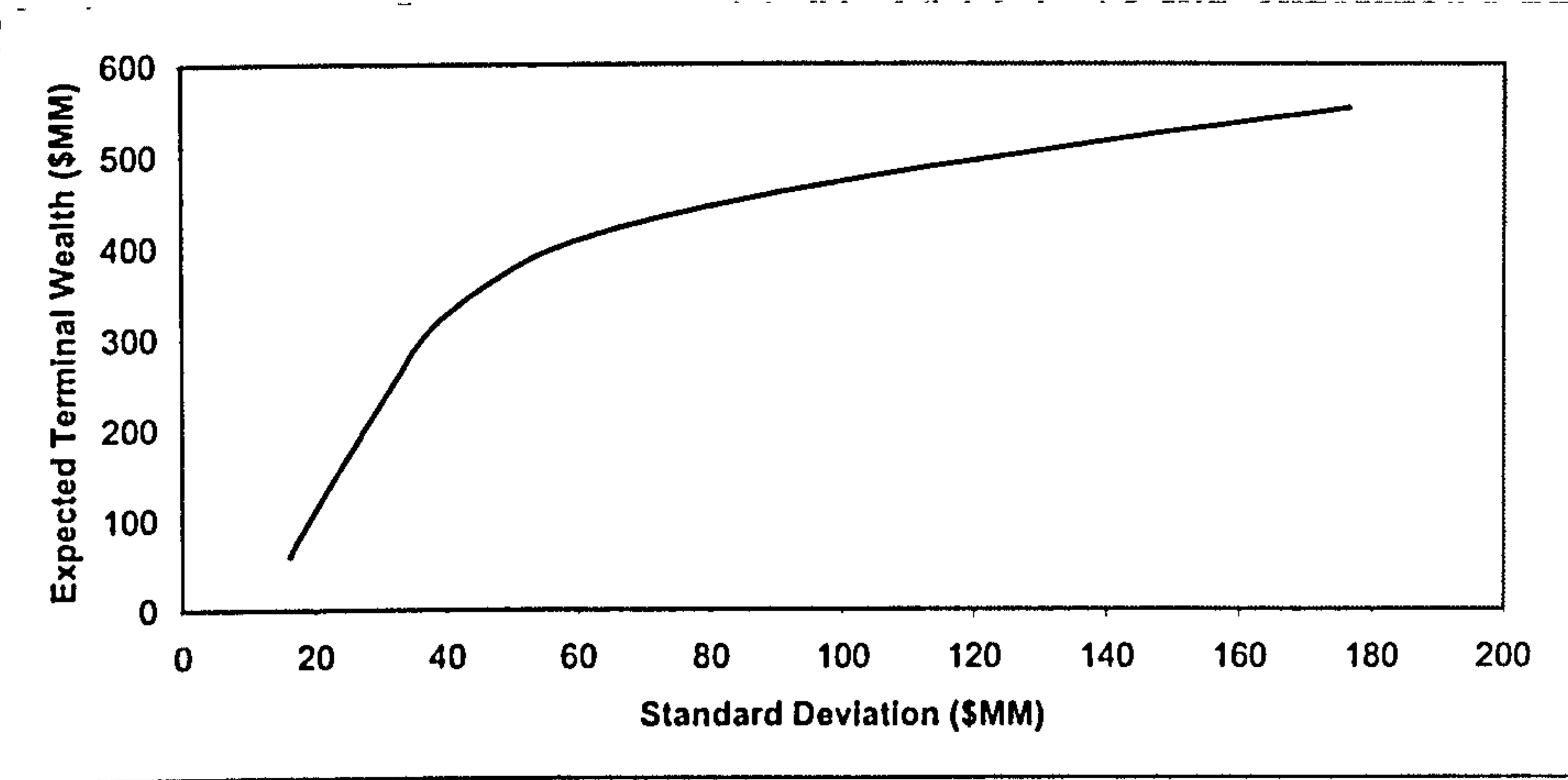


Figure 6.1: Mean-Variance Efficient frontier for the ALM problem. Guaranteed rate $g = 0\%$.

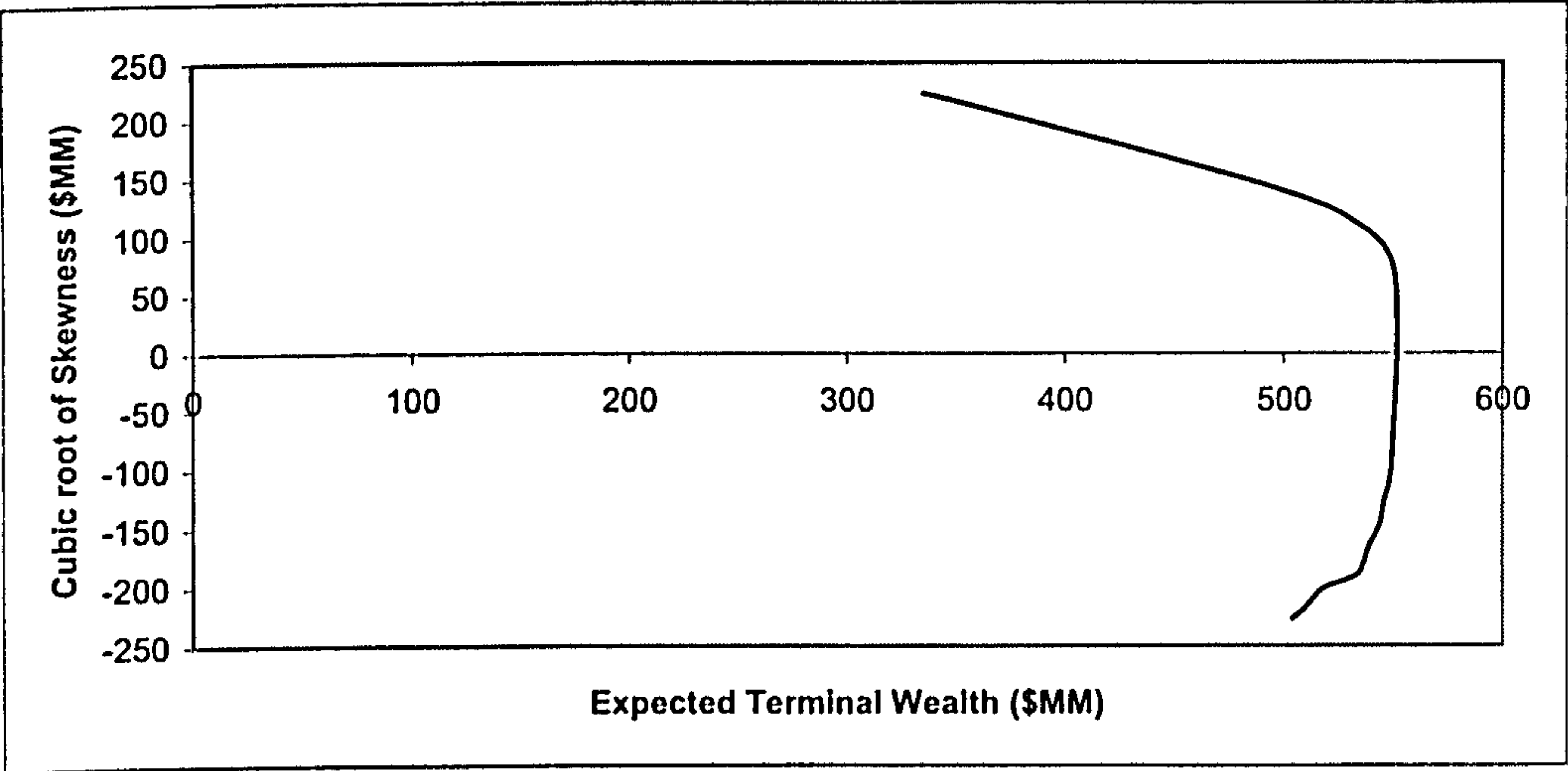


Figure 6.2: Mean-Skewness "Efficient" frontier for the ALM problem. Guaranteed rate $g = 0\%$.

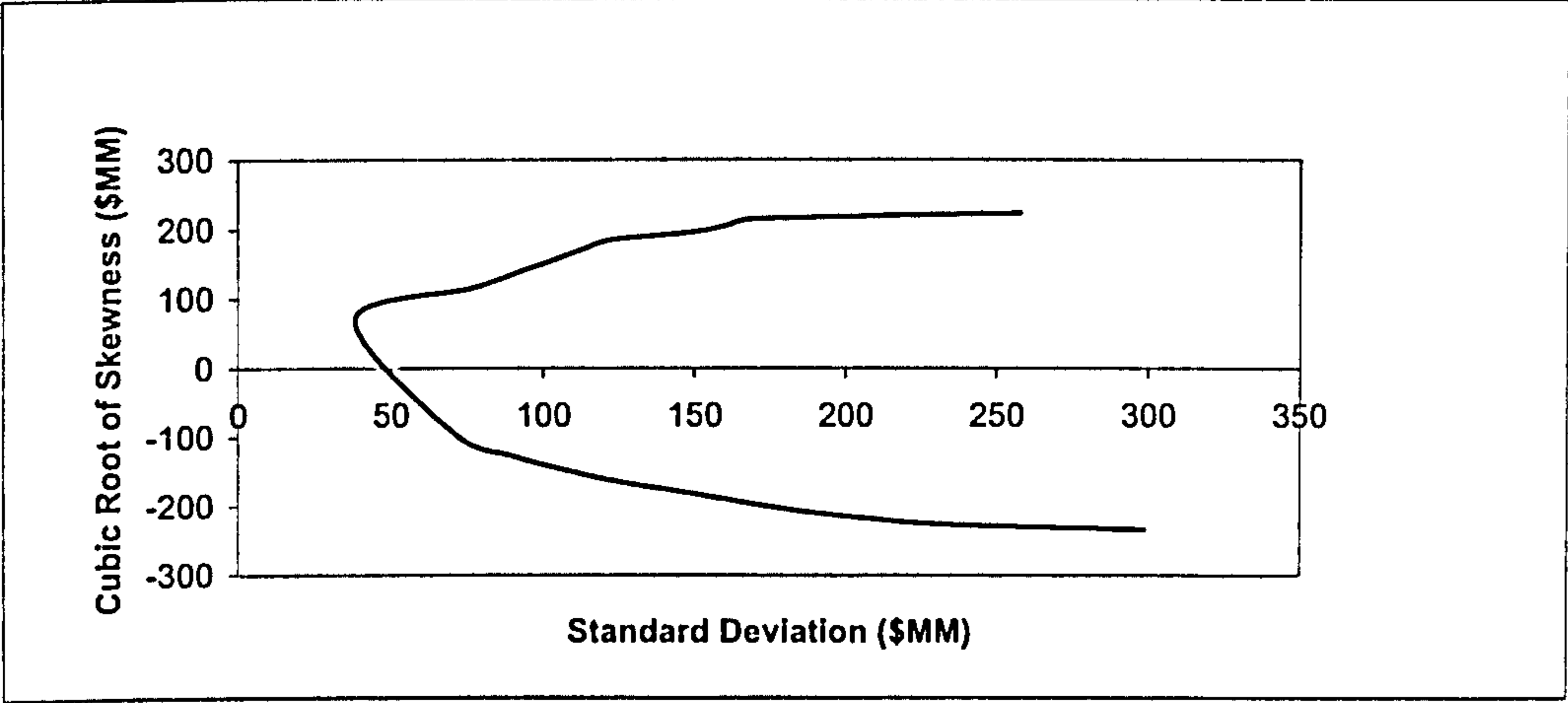


Figure 6.3: Variance-Skewness “Efficient” frontier. Guaranteed rate $g = 0\%$.

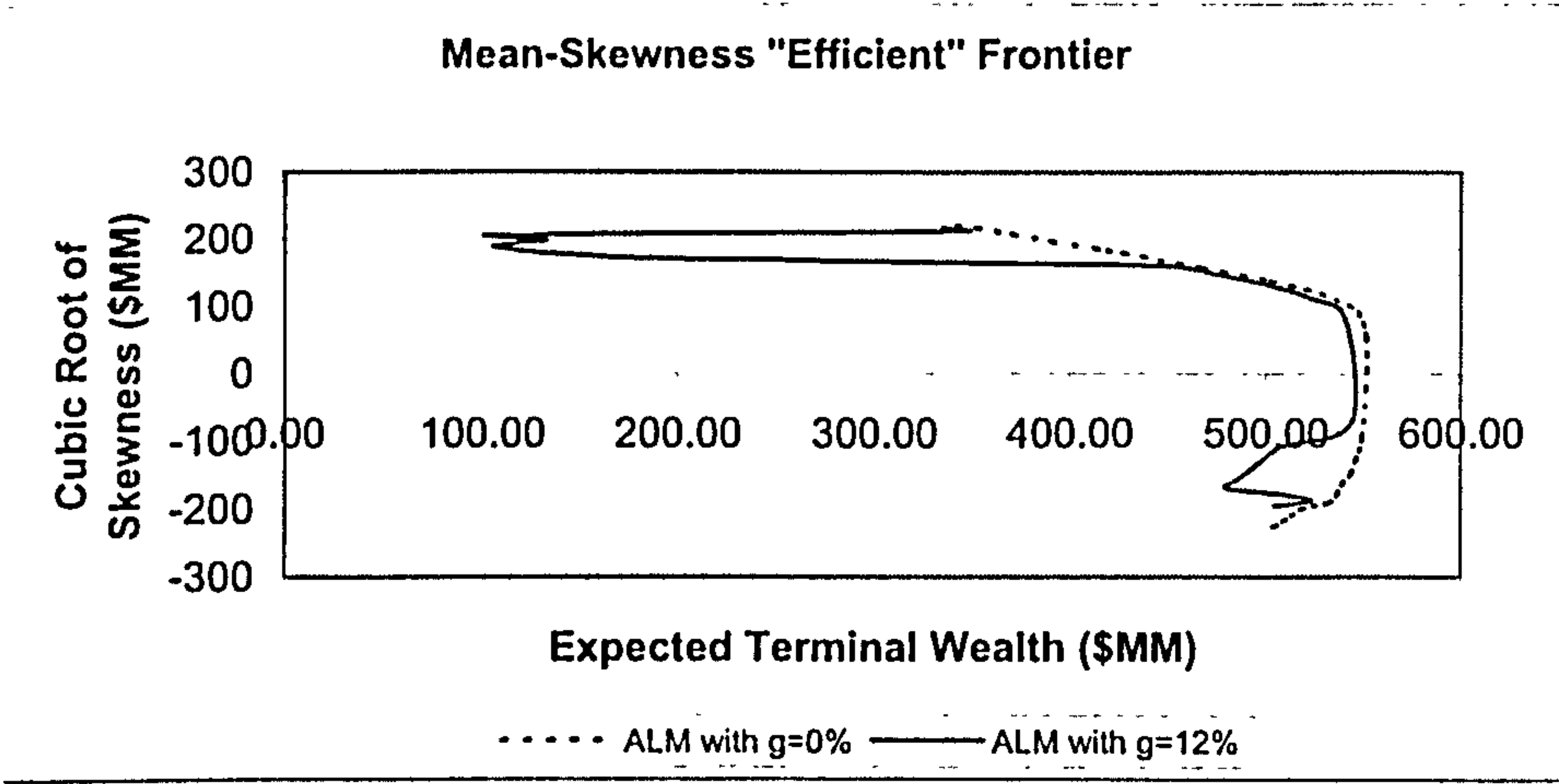


Figure 6.4: Mean-Skewness “Efficient” frontier for $g = 0\%$ and $g = 12\%$.

Chapter 7

Robust Asset Liability Management with Higher Order Moments

7.1 Introduction

In this chapter, we augment the higher moment ALM models of the previous section with the minimax method. We present the robust variants of the MVS and MDR models. The models have been implemented in AMPL, and back-tested using the MSCI regional data from

7.2 The Robust Mean-Variance-Skewness ALM Model

Here, we extend the mean-variance-skewness model to robust minimax optimisation over a set of rival return, variance and skewness scenarios describing the probabilistic distribution of the random variables. As with the mean-variance model, we assume rivalling sets of means, variances and skewnesses for the asset returns. Each rivalling set is defined by the pair $(\mathbf{r}_{je}, p_{je})$, where $e \in \mathcal{N}$, $j = 1, \dots, J$. We can obtain the rival means, variances and skewnesses from the pair $(\mathbf{r}_{je}, p_{je})$. The generation of rival scenario trees is discussed in section 4.2. Thus, we extend the mean-variance-skewness model to robust minimax optimisation over the set of rival mean, variance and skewnesses describing the probabilistic

distribution of the stochastic variables of the asset liability management models. Given the return vectors $r_j \in R^n$, $j = 1 \dots J$, the robust variant of the mean-variance-skewness problem can thus be written as below:

RobMVS:

$$\begin{aligned}
& \max_{\mathbf{H}, \mathbf{B}, \mathbf{S}} S \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_{je} + mb_e - ml_e - (1 + r^f_{jbe})mb_{a(e)} + \\
& (1 + r^f_{jle})ml_{a(e)} = 0, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, j = 1, \dots, J, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, j = 1, \dots, J, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, j = 1, \dots, J, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, j = 1, \dots, J, e \in \mathcal{N}, \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{ja(e)}(1 + \bar{g} + \pi_{je}^+), j = 1, \dots, J, \\
& \alpha R_{je}^P - \bar{g} = \pi_{je}^+ - \pi_{je}^-, j = 1, \dots, J, \\
& \pi_{je}^+ \geq 0, \pi_{je}^- \geq 0, \pi_{je}^+ \pi_{je}^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1, \dots, J, \\
& \sum_{e \in \mathcal{N}_T} P_{je} [W_{je} - \overline{W_{je}}]^2 \leq \mathcal{V}, j = 1, \dots, J, \\
& \sum_{e \in \mathcal{N}_T} P_{je} [W_{je} - \overline{W_{je}}]^3 \geq \mathcal{S}, j = 1, \dots, J, \\
& W_{je} = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W_{je}} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1, \dots, J.
\end{aligned}$$

7.3 Robust Mean-Downside Risk ALM Model

As has been done above for the mean-variance and mean-variance-skewness problems, we propose the robust counterpart of the mean-downside risk ALM problem:

$$\begin{aligned}
& \min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \max_{j \in J} [\max(0, \overline{W}_{je} - W_{je})]^2 \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_{je} + mb_e - ml_e - (1 + r^f_{jbe})mb_{a(e)} + \\
& (1 + r^f_{jle})ml_{a(e)} = 0, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, j = 1, \dots, J, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, j = 1, \dots, J, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, j = 1, \dots, J, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, j = 1, \dots, J, e \in \mathcal{N}, \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{ja(e)}(1 + \bar{g} + \pi_{je}^+), j = 1, \dots, J, \\
& \alpha R_{je}^P - \bar{g} = \pi_{je}^+ - \pi_{je}^-, j = 1, \dots, J, \\
& \pi_{je}^+ \geq 0, \pi_{je}^- \geq 0, \pi_{je}^+ \pi_{je}^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1, \dots, J, e \in \mathcal{N}, \\
& W_{je} = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \overline{W}_{je} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1, \dots, J. \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1 \dots J, e \in \mathcal{N}_T. \\
& \overline{W}_{je} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1 \dots J, e \in \mathcal{N}.
\end{aligned}$$

In the same manner as was done for the non-robust formulation, we can eliminate the discontinuous objective by the introduction of artificial variables v_{je}^- and v_{je}^+ and rewriting the problem as:

$$\begin{aligned}
& \min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \max_{j \in J} (v_{je}^-)^2 \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_{je} + mb_e - ml_e - (1 + r^f_{jbe})mb_{a(e)} + \\
& (1 + r^f_{jle})ml_{ja(e)} = 0, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, j = 1, \dots, J, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, j = 1, \dots, J, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, j = 1, \dots, J, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, j = 1, \dots, J, e \in \mathcal{N}, \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{ja(e)}(1 + \bar{g} + \pi_{je}^+), j = 1, \dots, J, \\
& \alpha R_{je}^P - \bar{g} = \pi_{je}^+ - \pi_{je}^-, j = 1, \dots, J, \\
& \pi_{je}^+ \geq 0, \pi_{je}^- \geq 0, \pi_{je}^+ \pi_{je}^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1, \dots, J, \\
& W_{je} = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1 \dots J, e \in \mathcal{N}_T. \\
& \overline{W}_{je} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1 \dots J, e \in \mathcal{N}. \\
& \sum_{e \in \mathcal{N}_T} P_e W_{je} \geq \mathcal{W}, j = 1 \dots J, e \in \mathcal{N}_T, \\
& \overline{W}_{je} - W_{je} = v_{je}^- - v_{je}^+, j = 1 \dots J, e \in \mathcal{N}_T, \\
& v_{je}^- \geq 0, v_{je}^+ \geq 0, j = 1 \dots J, e \in \mathcal{N}_T.
\end{aligned}$$

or using the NLP formulation for discrete minimax problems, as

RobDSR

$$\begin{aligned}
& \min_{\mathbf{H}, \mathbf{B}, \mathbf{S}} \mathcal{D} \\
& \text{subject to} \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_e + mb_e - ml_e + m^{ini} = 0, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \mathbf{H}^{ini} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_0, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)} + \mathbf{B}_e - \mathbf{S}_e, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{S}'_e(1 - \mathbf{c}_{se}) - \mathbf{B}'_e(1 + \mathbf{c}_{be}) - y_{je} + mb_e - ml_e - (1 + r^f_{jbe})mb_{a(e)} + \\
& (1 + r^f_{jle})ml_{ja(e)} = 0, e \in \mathcal{N}_I, j = 1, \dots, J, \\
& \mathbf{H}_e = \hat{\mathbf{r}}'_{je} \mathbf{H}_{a(e)}, e \in \mathcal{N}_T, j = 1, \dots, J, \\
& \mathbf{H}_e^L \leq \mathbf{H}_e \leq \mathbf{H}_e^U, j = 1, \dots, J, \\
& \mathbf{B}_e^L \leq \mathbf{B}_e \leq \mathbf{B}_e^U, j = 1, \dots, J, \\
& \mathbf{S}_e^L \leq \mathbf{S}_e \leq \mathbf{S}_e^U, j = 1, \dots, J, e \in \mathcal{N}, \\
& mb_e = 0, ml_e = 0, \mathbf{B}_e = 0, \mathbf{S}_e = 0, e \in \mathcal{N}_T, \\
& L_e = (1 - \kappa_e)L_{ja(e)}(1 + \bar{g} + \pi_{je}^+), j = 1, \dots, J, \\
& \alpha R_{je}^p - \bar{g} = \pi_{je}^+ - \pi_{je}^-, j = 1, \dots, J, \\
& \pi_{je}^+ \geq 0, \pi_{je}^- \geq 0, \pi_{je}^+ \pi_{je}^- = 0, e \in \mathcal{N}, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1, \dots, J, \\
& W_{je} = \mathbf{r}'_{je} \mathbf{H}_{a(e)} - L_e, e \in \mathcal{N}_T, \\
& \sum_{e \in \mathcal{N}_T} P_{je} W_{je} \geq \mathcal{W}, j = 1 \dots J, e \in \mathcal{N}_T. \\
& \overline{W}_{je} = \sum_{e \in \mathcal{N}_T} P_{je} W_{je}, j = 1 \dots J, e \in \mathcal{N}. \\
& \sum_{e \in \mathcal{N}_T} P_e W_{je} \geq \mathcal{W}, j = 1 \dots J, e \in \mathcal{N}_T, \\
& \overline{W}_{je} - W_{je} = v_{je}^- - v_{je}^+, j = 1 \dots J, e \in \mathcal{N}_T, \\
& v_{je}^- \geq 0, v_{je}^+ \geq 0, j = 1 \dots J, e \in \mathcal{N}_T. \\
& v_{je}^- \leq \mathcal{D}, j = 1 \dots J, e \in \mathcal{N}_T.
\end{aligned}$$

The resulting minimax problems are solved using the algorithm described in [59], which combines the interior point method in [2] with the discrete minimax algorithm in [68].

The robust mean-variance profile in Figure 7.1 again shows the positive trade-off be-

tween the variance (or standard deviation) and the expected terminal wealth i.e. for any level of skewness, the higher the expected terminal wealth the higher the variance/standard deviation of the portfolio. The curve in Figure 7.2 again shows the negative trade-off between the expected terminal wealth and the skewness, same as with the normal mean-skewness frontier. That is, low skewness portfolios have a higher expected terminal wealth in the positive skewness range. In the negative skewness range, as the negative skewness decreases the expected terminal wealth increases. Figure 7.3 shows the positive trade-off between the skewness and the variance, similar to the mean-variance trade-off. Achieving greater positive skewness requires taking on greater risk in terms of variance or standard deviation. In the negative skewness range, higher levels of variance or standard deviation are associated with higher negative skewness.

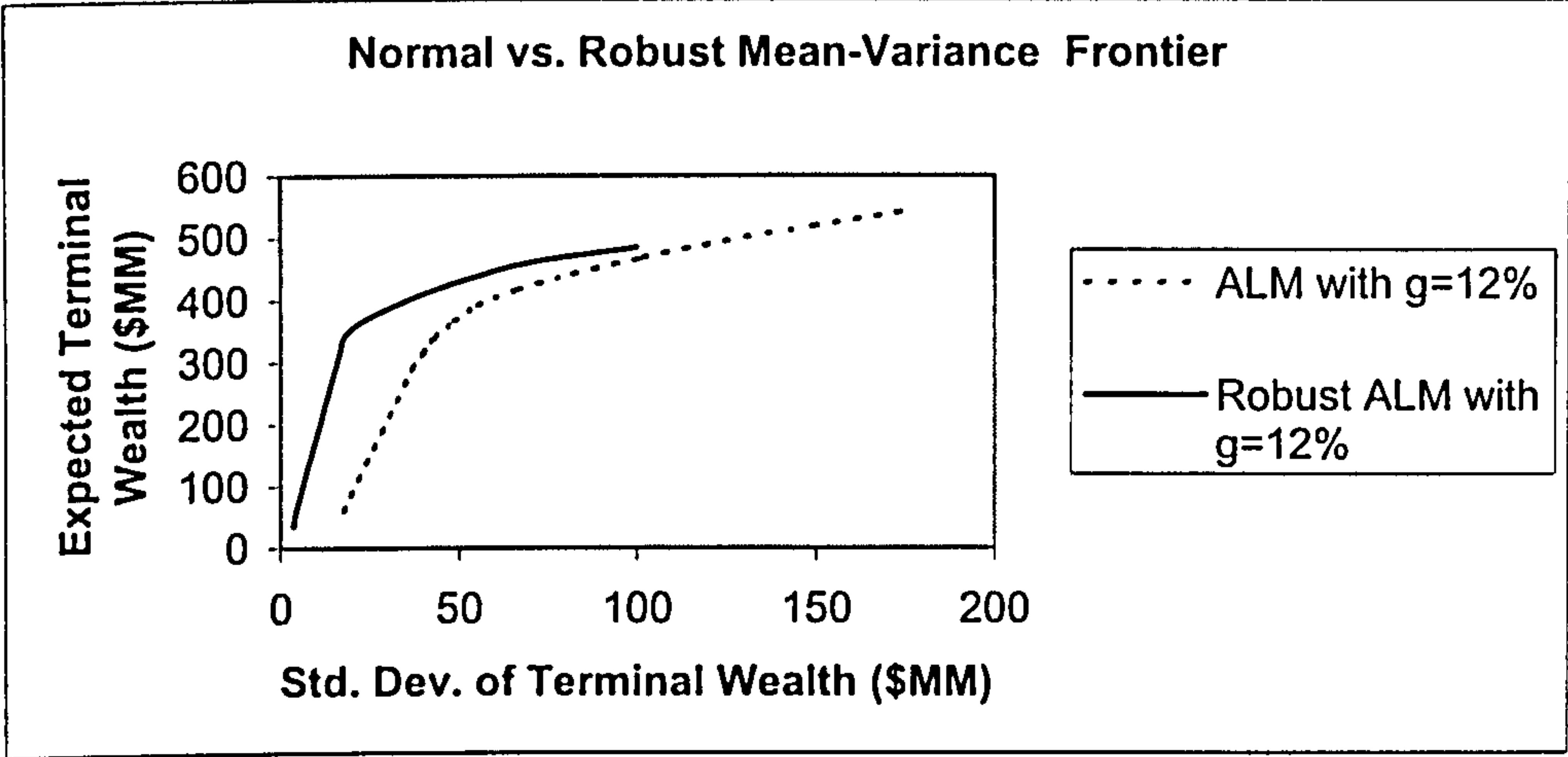


Figure 7.1: Mean-Variance Efficient frontier for the ALM problem. Guaranteed rate $g = 12\%$.

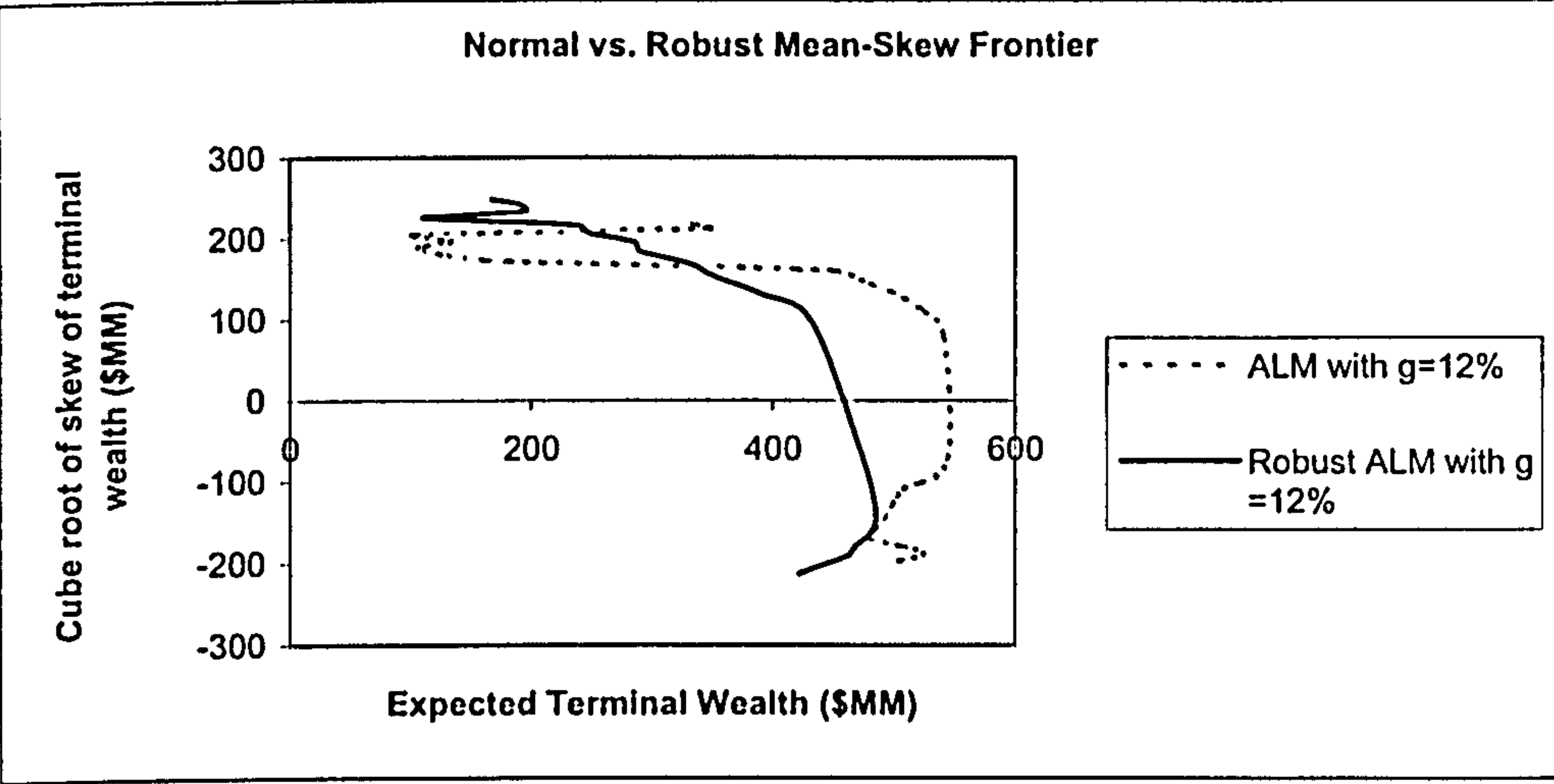


Figure 7.2: The normal and robust mean-skewness frontiers for the ALM problem. Guaranteed rate $g = 12\%$.

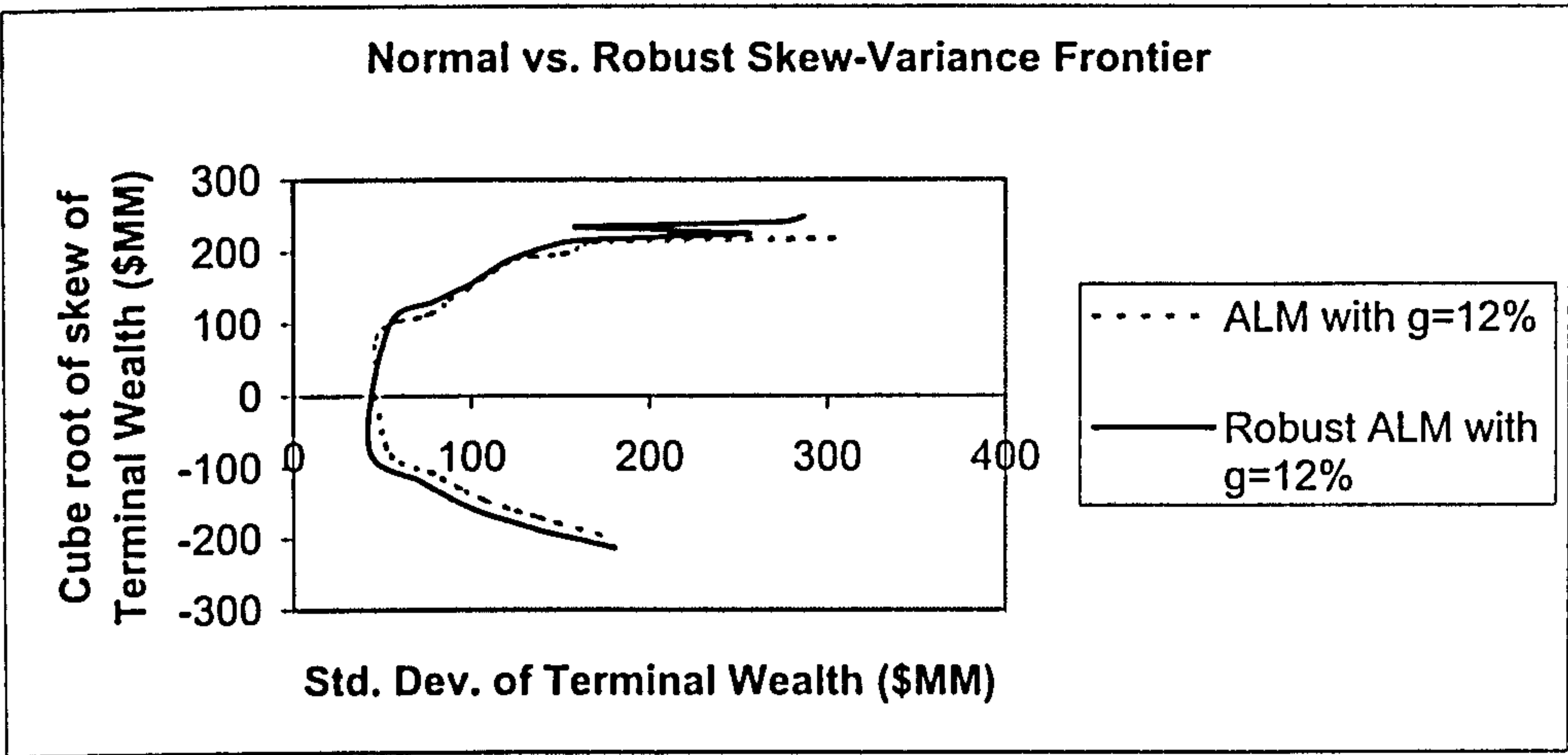


Figure 7.3: The normal and robust variance-skewness “Efficient” frontier. Guaranteed rate $g = 12\%$.

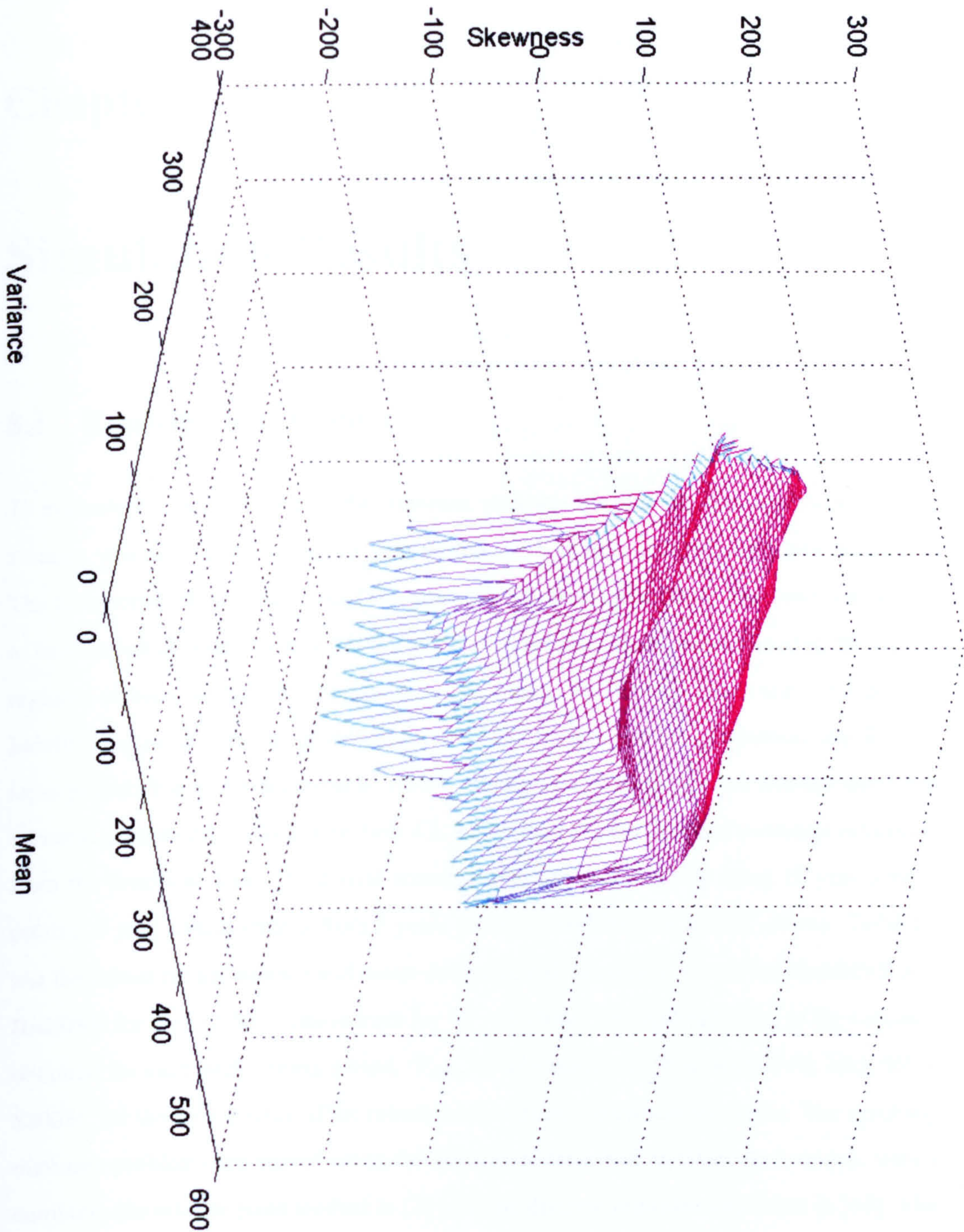


Figure 7.4: Mean-Variance-Skewness surface for $g=12\%$, m = mean, v =standard deviation, s =cubic root of skewness

Chapter 8

Simulation Results

8.1 Experimental Results

To evaluate the performance of the minimax strategy dynamically, the application of the strategy was simulated over an 11 year period, with annual re-balancing when necessary. The time period chosen was from December 1992 to December 2003. The fund is assumed to have an initial endowment of \$500M to allocate across 5 assets, represented by the MSCI regional indices for the UK, Europe ex. UK, Japan, Pacific ex. Japan and the US. The liability is generated by the growth in the US MSCI index and the guaranteed rate, $\bar{g}$. The lapse probability, κ_t is assumed to be zero in every period. The returns are forecast using the scenario generator described in section 4.2, which matches the first four moments estimated from the historical data. Three rival scenario trees are generated by using 10 year annual returns, 5 year annual returns from 5 years previous, and 20 year annual returns. To back-test the robust mean-variance and mean-downside risk strategies, we solved **RobMVS** and **RobDSR** for each of the scenario trees for $\mathcal{W} = \$300M$. The optimal value of the variance obtained for each back-testing period, $\mathcal{V}_{min}$, is substituted for $\mathcal{V}$ in **RobMVS**, with $\mathcal{W} = \$300M$, for the back-testing of the robust mean-variance skewness strategies. The resulting minimax problems are solved using the algorithm described in Chapters 1 and 2, which combines the interior point method in [2] with the discrete minimax algorithm in [68]. The realised returns from the optimal holdings for each strategy are used to compute the total

asset values and liability values from which we calculate the surplus wealth. We assume that surpluses generated in any period are invested at a risk-free rate of 4%. Shortfalls are paid out from the money accrued in the surplus account if it contains sufficient funds, or borrowed at the rate of 4% otherwise. The standard mean-variance (MV), mean-variance-skewness (MVS) and mean-downside risk (MDR) strategies were also back-tested using **NLP1**, **ALMHOMSKEW** and **ALMHOMDSR** with $W = \$300M$. The optimal value of the variance from the solution of **NLP1** was substituted into **ALMHOMSKEW** to back-test the mean-variance-skewness strategy. Results and comparisons of the strategies are provided below.

In Figure 8.1, we show the evolution of the surplus account obtained by the application of the minimax mean-variance strategy starting from December 1992 to December 2003. In addition, we present the surplus accounts obtained from the mean-variance optimal holdings associated with each of the scenario trees. ScTree1 generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns. At the end of the back-testing period, the ScTree1 strategy generates a surplus of $\$193.92M$, ScTree2 strategy generates a surplus of $\$178M$. The best mean-variance strategy, the ScTree3 strategy, corresponds to a surplus of $\$199.89M$ while the Minimax strategy yields a surplus of $\$207.95M$.

Figure 8.2, we show the evolution of the surplus account obtained by the application of the minimax mean-variance-skewness strategy starting from December 1992 to December 2003. In addition, we present the surplus accounts obtained from the mean-variance-skewness optimal holdings associated with each of the scenario trees. At the end of the back-testing period, the ScTree1 MVS strategy generates a surplus of $\$234.48M$, ScTree2 MVS strategy generates a surplus of $\$206.90M$. The best mean-variance-skewness strategy, the ScTree3 strategy, corresponds to a surplus of $\$241.09M$ while the Minimax MVS strategy yields a surplus of $\$254.18M$.

In Figure 8.3, we show the evolution of the surplus account obtained by the application of the minimax mean-downside risk (MDR) strategy starting from December 1992 to December 2003. In addition, we present the surplus accounts obtained from the mean-

downside-risk optimal holdings associated with each of the scenario trees. At the end of the back-testing period, the ScTree2 MDR strategy generates a surplus of \$209.24M, ScTree3 MDR strategy generates a surplus of \$188.73M. The best MDR strategy, the ScTree1 strategy, corresponds to a surplus of \$223.63M while the Minimax MDR strategy yields a surplus of \$202.25M.

Figure 8.4 compares the performance of the three minimax strategies in back-testing. The minimax mean-variance-skewness approach provides the best performance over the 11 year period.

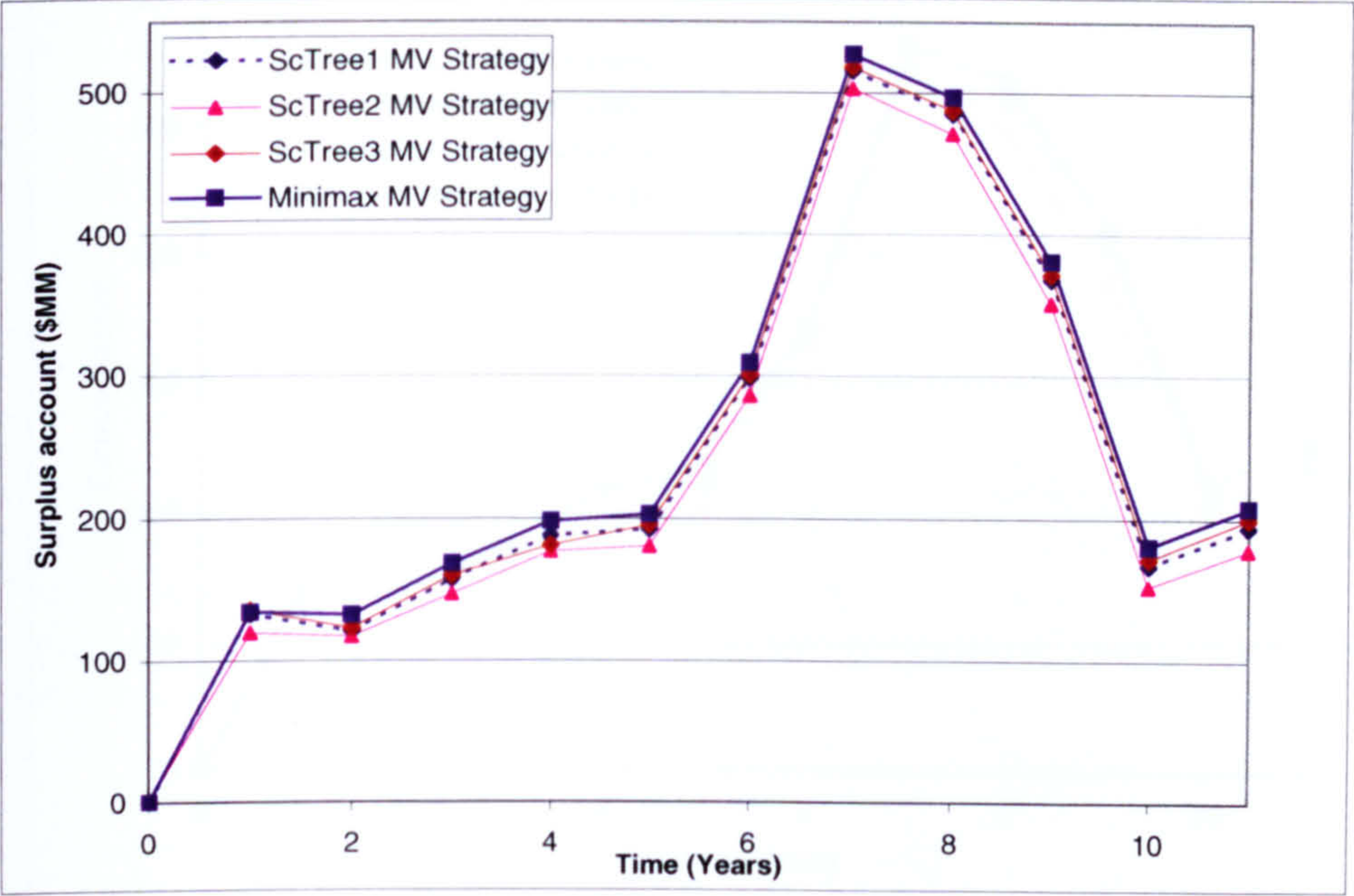


Figure 8.1: Evaluation of the minimax mean-variance strategy over 11 years and comparison with mean-variance efficient frontier strategies for scenario trees 1, 2 and 3. ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns

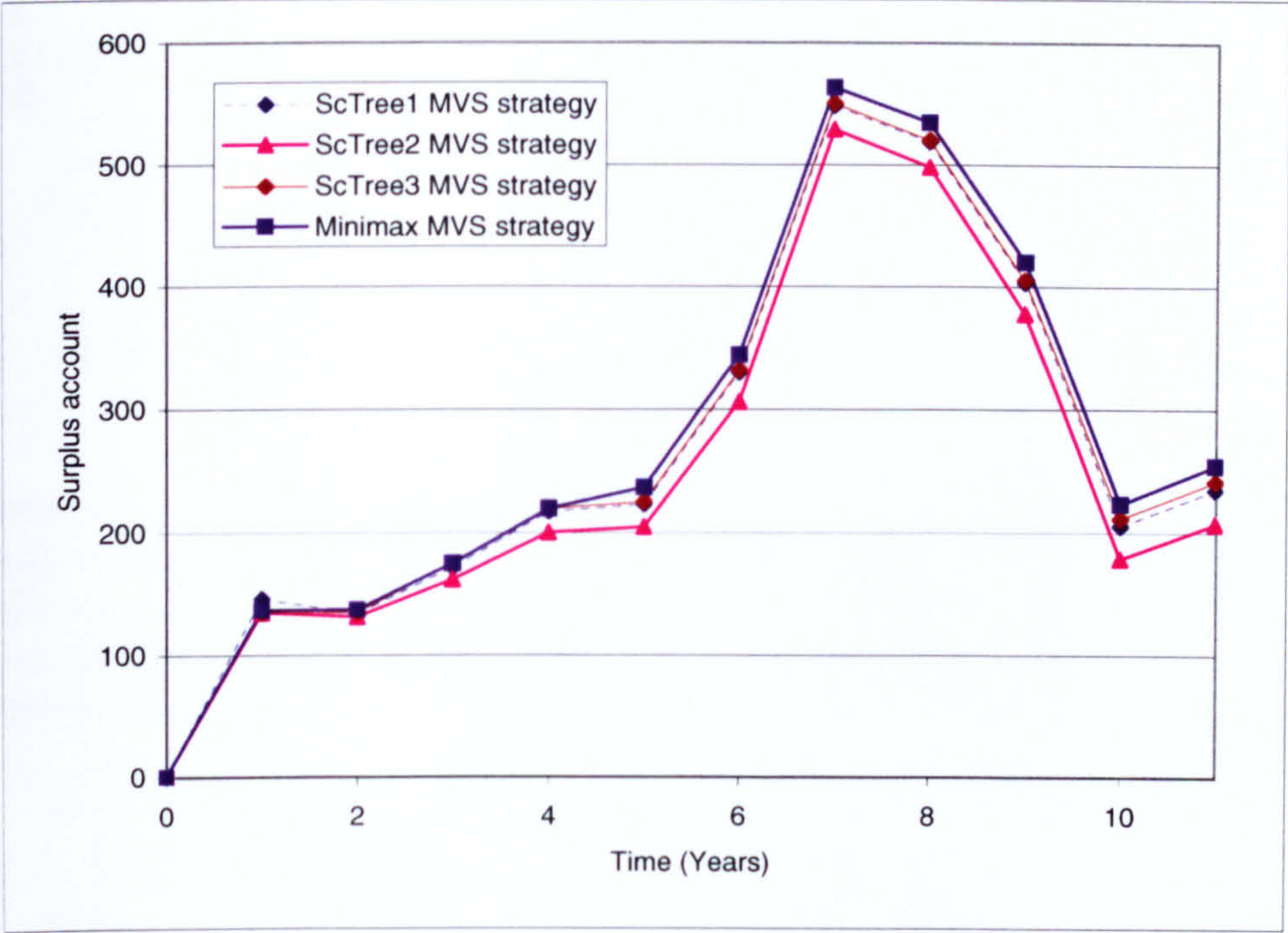


Figure 8.2: Evaluation of the minimax mean-variance-skewness strategy over 11 years and comparison with mean-variance-skewness efficient frontier strategies for scenario trees 1, 2 and 3. ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns

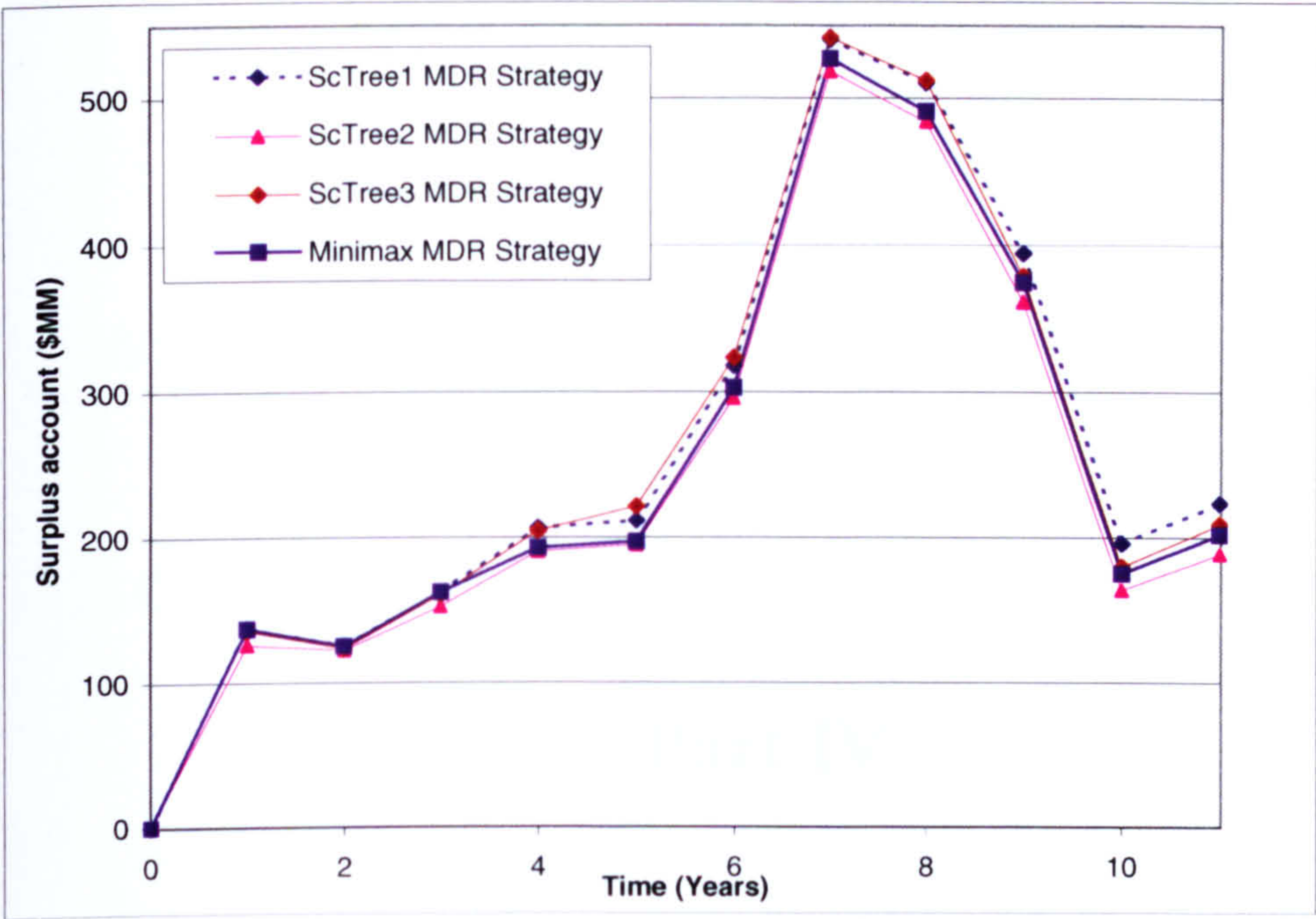


Figure 8.3: Evaluation of the minimax mean-downside-risk strategy over 11 years and comparison with mean-downside-risk efficient frontier strategies for scenario trees 1, 2 and 3. ScTree1 is generated using 10 year annual returns, ScTree2 is the scenario tree generated using the 5 year annual returns from 5 years previous, and ScTree3 is generated from 20 year annual returns

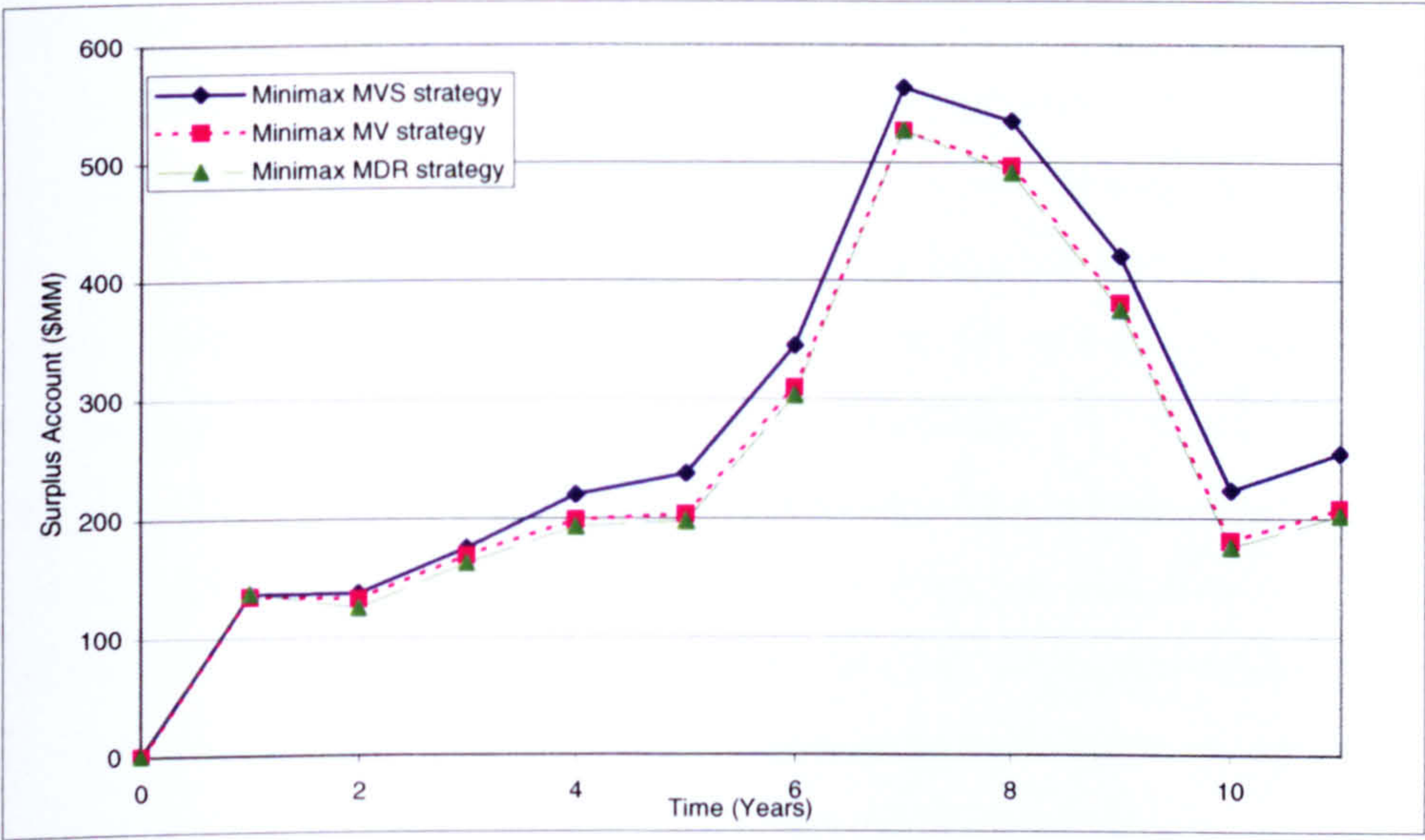


Figure 8.4: Comparison of the results obtained from the application of the robust mean-variance, mean-variance-skewness and mean-downside-risk strategies in back-testing.

Part IV

Conclusions

Chapter 9

Discussion

9.1 Summary

The robust strategy produced the best results in our back-testing for the MV and MVS framework. However, the robust strategy was not the best performing strategy for the MDR framework. Both the minimax and mean-downside risk methods are cautious approaches, and the combination of the two results in an overly cautious allocation strategy which impairs the performance of the portfolio.

Figures 9.1, 9.2, 9.3 show the surplus accounts obtained using the MV, MVS and MDR strategies with each of the scenario trees. ScTree3, which matches the moments using the previous 20 years of historical data, outperforms the other two trees which use the previous 5 years of historical data, and 5 years of historical data from 5 years previous respectively. For the back-testing period in our experimental results, the non-robust MVS strategy is the best performing non-robust strategy for each of the scenario trees. The non-robust MDR strategy yields the next best performance for all three scenario trees. In all three cases, the non-robust MV strategy is the poorest performer in our experiment. In conclusion, the minimax strategy over rival scenario trees was found to dominate the single-tree strategy for all three scenario trees considered for all but the MDR strategy.

For the MV and MDR frameworks, the robust strategies were not found to increase the average number of iterations required to solve the problems significantly, as can be seen

from table 9.1. The iteration counts for the MVS framework were more difficult to ascertain due to the globalisation scheme required.

Figures 9.4, 9.5 and 9.6 show the asset allocation structure of the mean-variance, mean-variance-skewness and mean-downside-risk minimax strategies respectively over the 11 year back-testing period. The changes in asset allocation suggested by the strategies are reasonable and implementable. The robust strategies are not guaranteed to perform better than non-robust strategies, as can be seen from our back-testing results for the MDR framework, but allow a way of incorporating a finite uncertainty structure into our data. The robust strategy gives a guaranteed minimum performance in view of the uncertainty set provided.

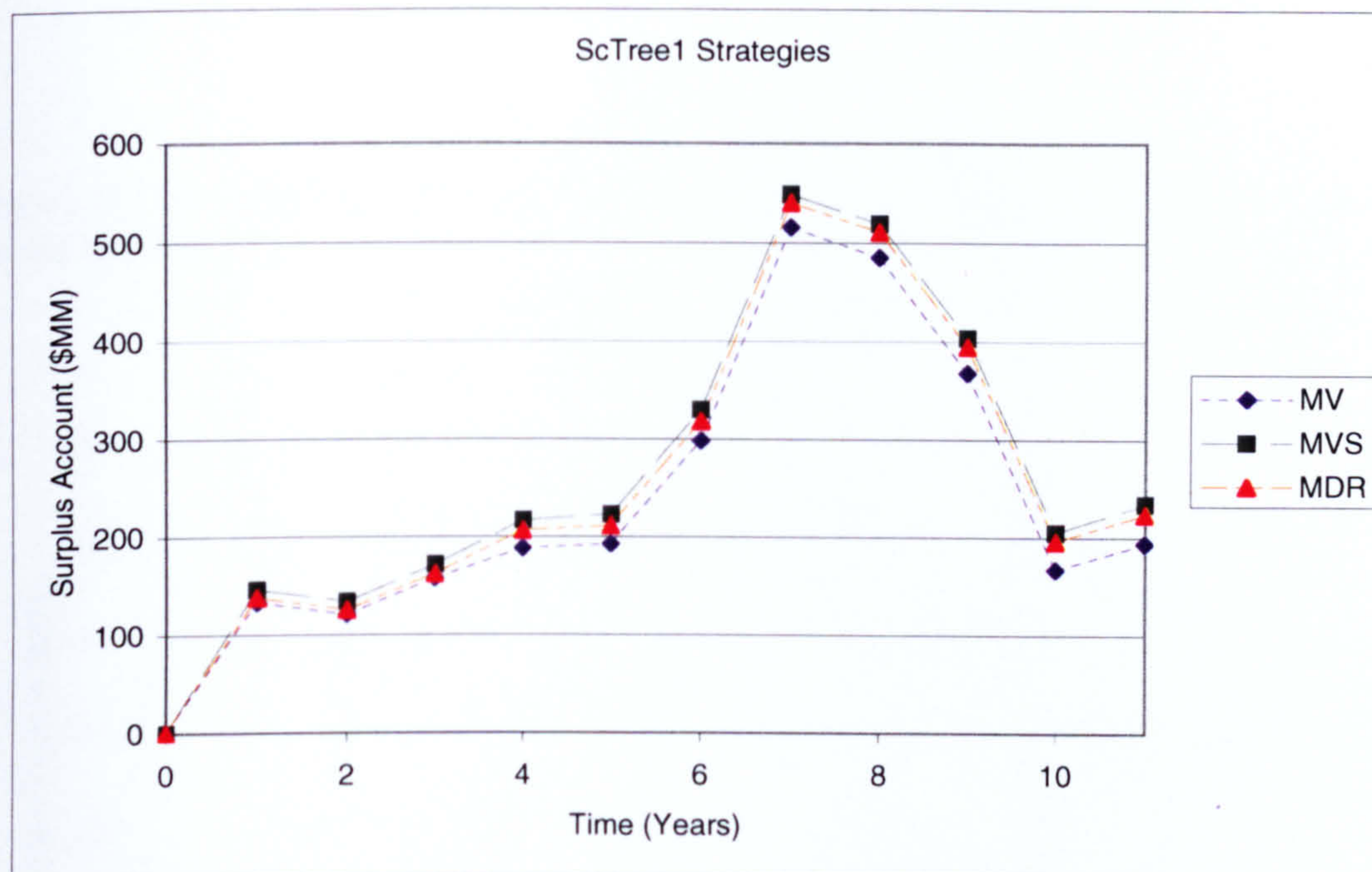


Figure 9.1: MV, MVS and MDR strategies with ScTree1 data (10 year annual returns) used in the back-testing.

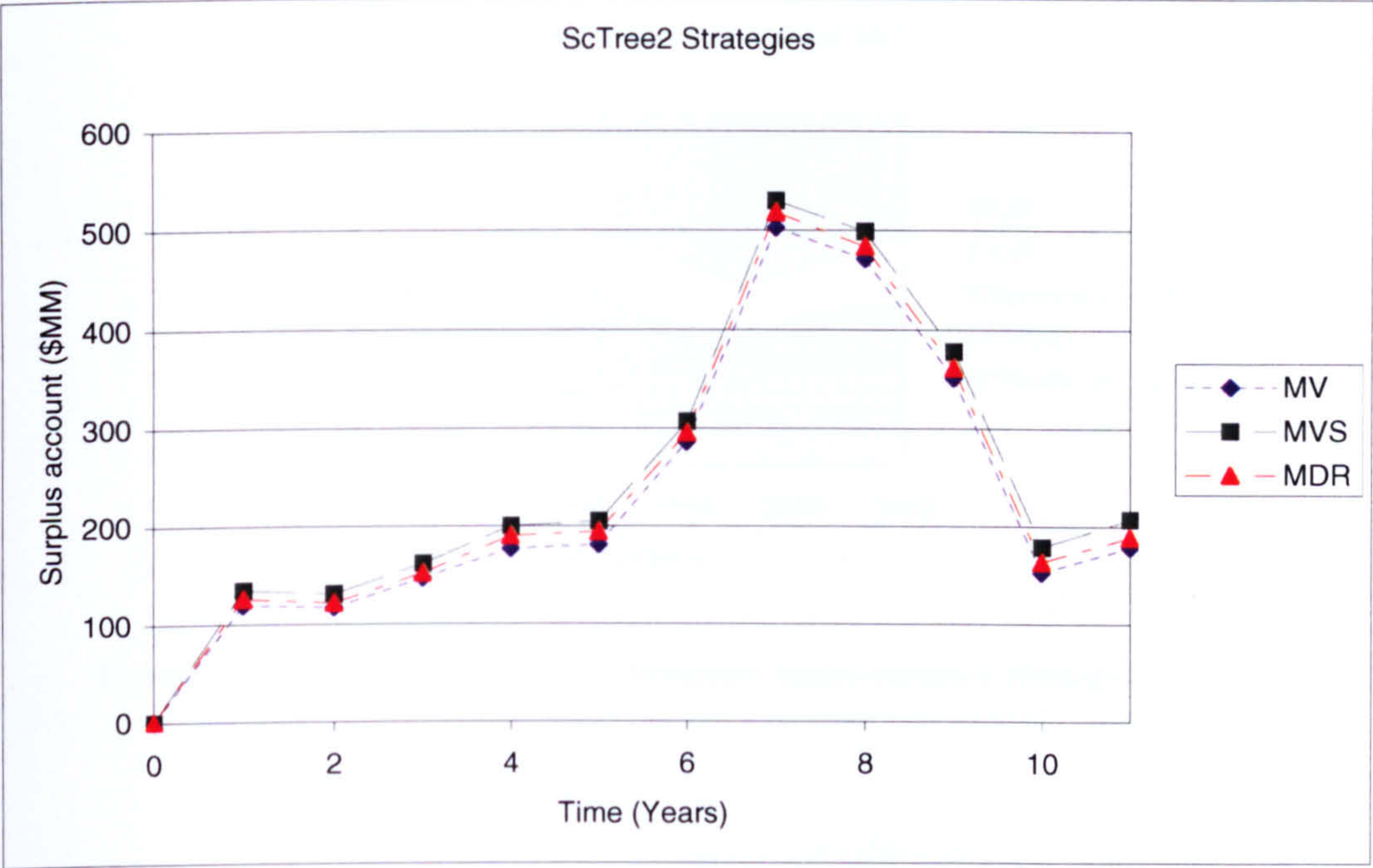


Figure 9.2: MV, MVS and MDR strategies with ScTree2 data (10 year annual returns) used in the back-testing.

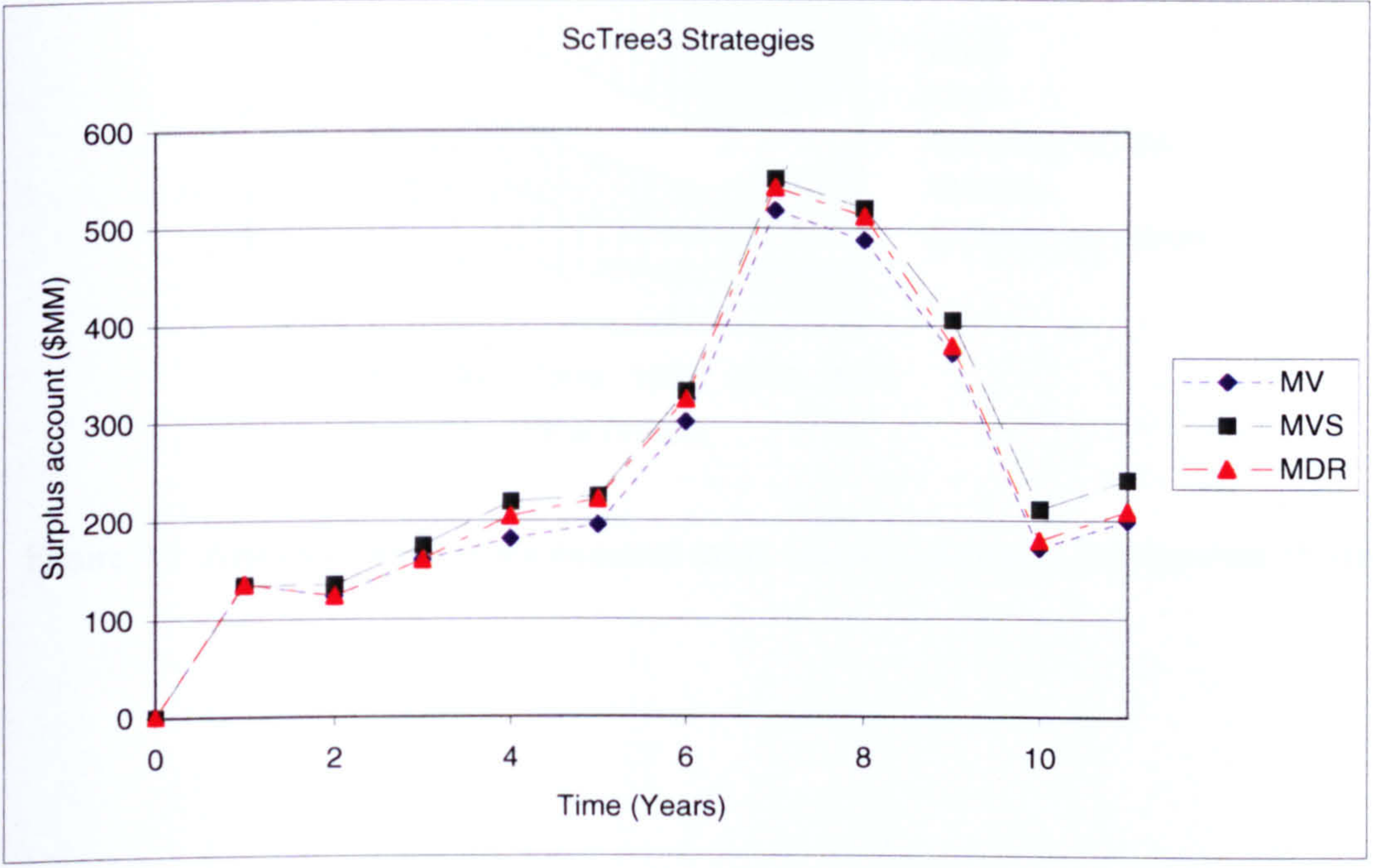


Figure 9.3: MV, MVS and MDR strategies with ScTree3 data (10 year annual returns) used in the back-testing.

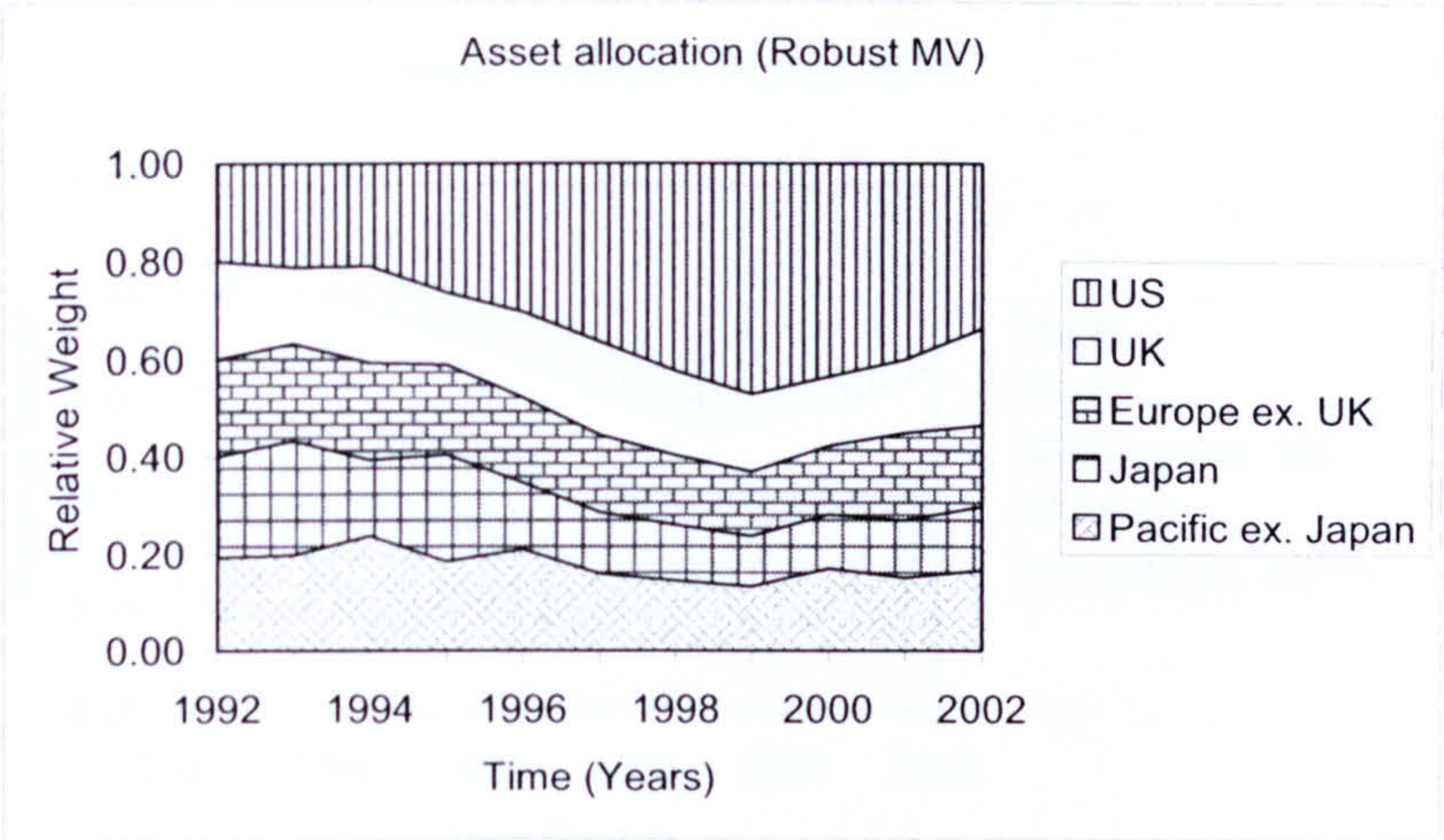


Figure 9.4: Asset allocation of the minimax mean-variance strategy over 11 years

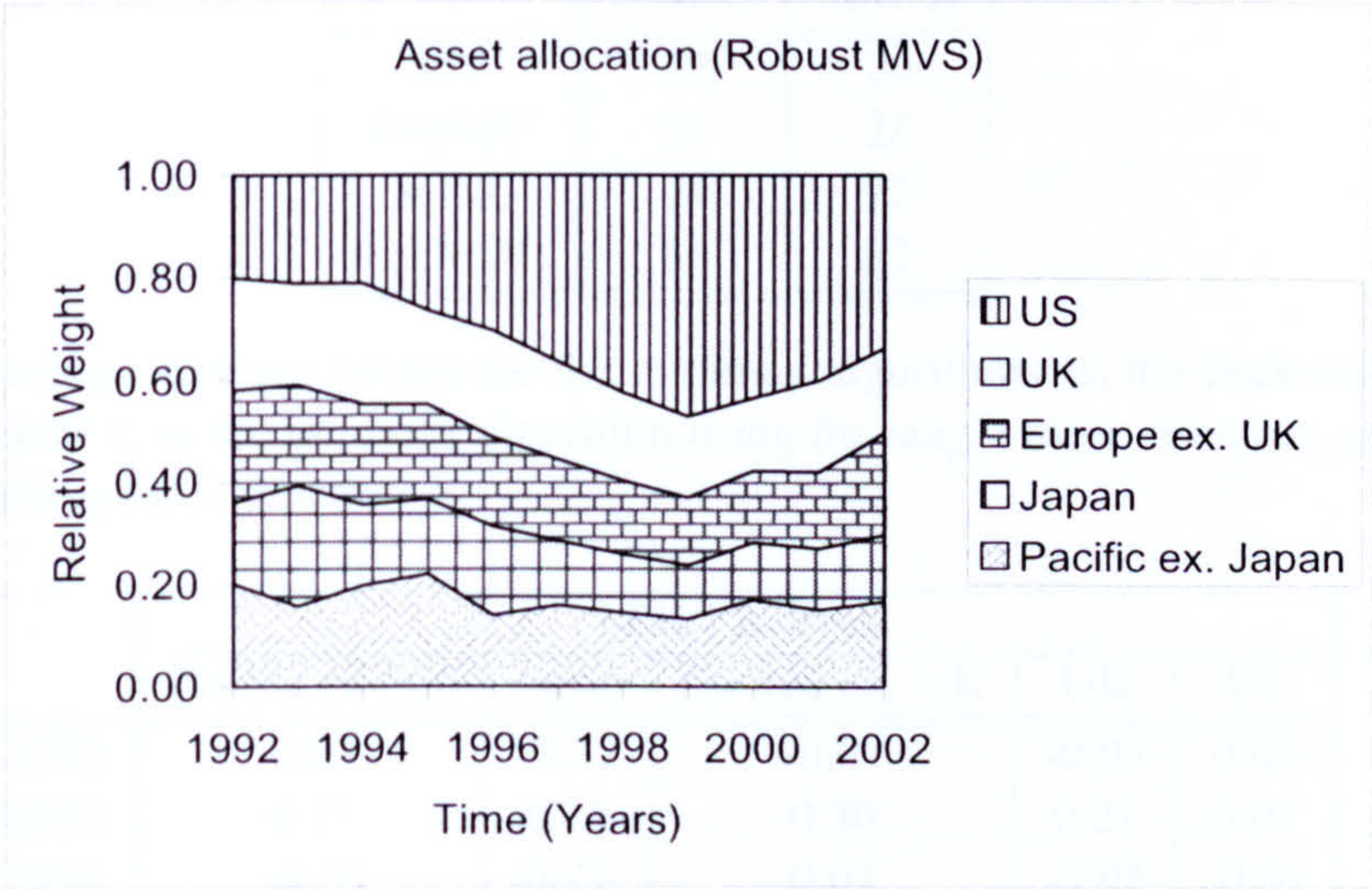


Figure 9.5: Asset allocation of the minimax mean-variance-skewness strategy over 11 years

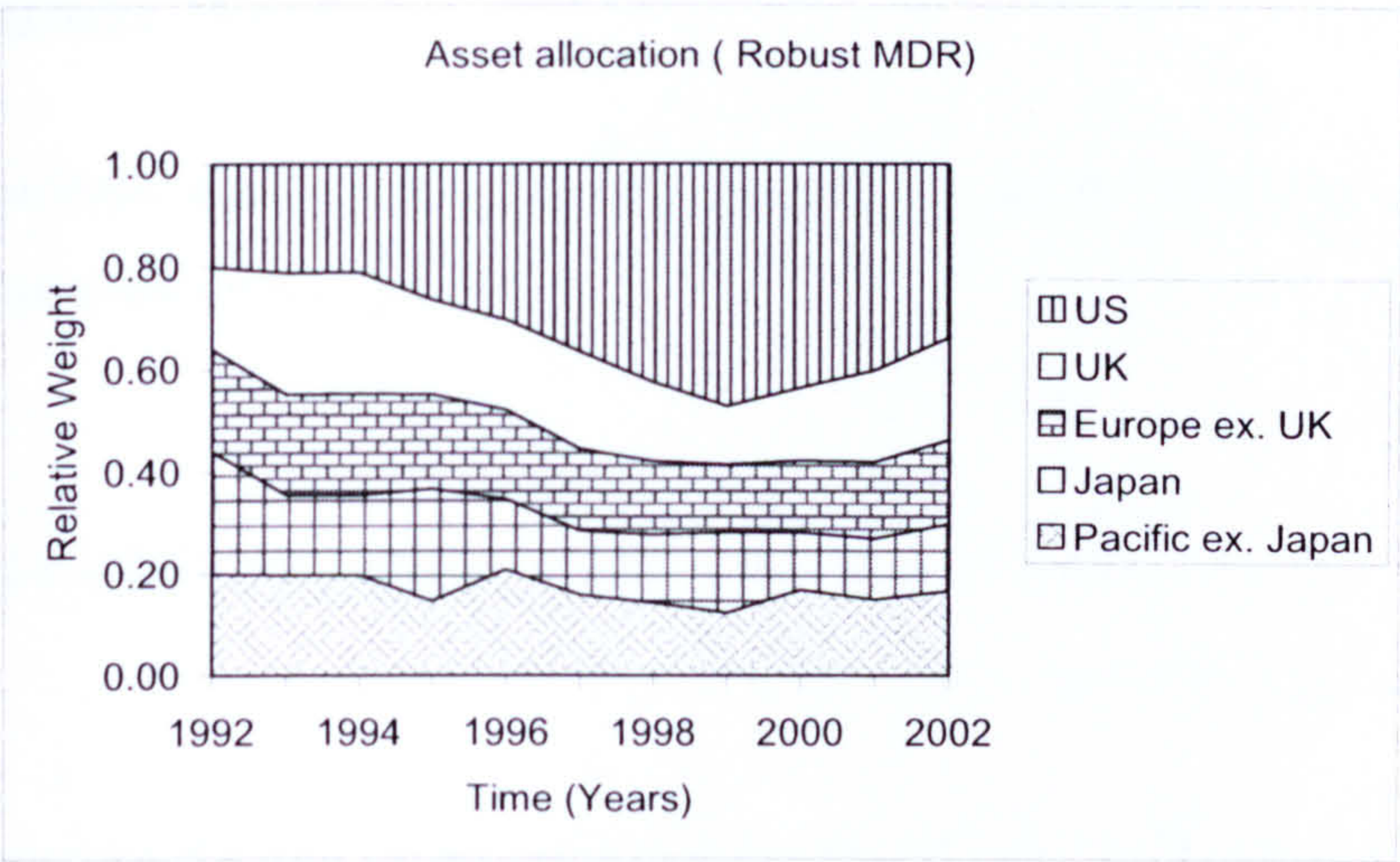


Figure 9.6: Asset allocation of the minimax mean-downside-risk strategy over 11 years

Strategy		
	MMX ₁	MMX ₂
MV	22	20
RobMV	31	28
MDR	74	72
RobMDR	86	82

Table 9.1: Average iteration counts for the minimax algorithms on the back-testing problems, where MMX_1 is the minimax algorithm using the single merit function, and MMX_2 uses the switching merit function.

	MSCI Indices				
	Pacific ex Jap.	Japan	Europe ex. UK	UK	US
1992	0.03	-0.22	-0.07	-0.07	0.04
1993	0.75	0.25	0.30	0.21	0.07
1994	-0.16	0.21	0.03	-0.05	-0.01
1995	0.10	0.00	0.20	0.17	0.35
1996	0.18	-0.16	0.16	0.23	0.21
1997	-0.33	-0.24	0.23	0.19	0.32
1998	-0.10	0.04	0.32	0.15	0.29
1999	0.39	0.61	0.16	0.10	0.21
2000	-0.17	-0.29	-0.08	-0.14	-0.14
2001	-0.12	-0.30	-0.23	-0.16	-0.13
2002	-0.09	-0.11	-0.22	-0.18	-0.24
2003	0.41	0.35	0.40	0.27	0.27

Table 9.2: Realised returns for the five indices over the back-testing period.

9.2 Future Work

Let the portfolio return at the end of the planning horizon be defined by R_T^P . Then the liability dynamics can be written as

$$\begin{aligned} L_T = L_0[1 + \max(\bar{g}, R_T^P)] &= L_0[1 + R_T^P] + L_0[1 + \max(\bar{g} - R_T^P, 0)] = \\ &L_0[1 + R_T^P] + L_0[1 + (\bar{g} - R_T^P)^+] \end{aligned} \quad (9.1)$$

The guaranteed product in the ALM problem provides a hybrid fixed-income and option-like pay-off structure for the policy-holder. Thus the product can be more or less synthesised by an investment in a fixed-interest bearing asset plus a call option on the MSCI index. The fixed-income asset would provide the guaranteed capital and the call option would provide the upside from rises in the index. This is often the practice by the financial institutions offering these products. The option can, however, also be synthesised by geared holdings in the underlying asset and some simple trading rules. For example, if the index fell substantially, the call option would be deeply out-of-the money. It could be synthesised by a small holding of MSCI index equities (in case of a rise in the index) funded by the surplus account or borrowing, if available. Conversely, if the MSCI index rose, greater exposure to the index would be required and more MSCI index equities would need to be bought. The option pay-off would be provided by the value of the MSCI holding in the portfolio. This is equivalent to a "buy high, sell low" strategy, and is demonstrated in the asset allocation displayed in our portfolio optimisation back-testing. In the periods of high returns in the US market, (1995-1999, as shown in table 9.2), where $(\bar{g} - R_T^P) \leq 0$, all the strategies increasingly allocate funds to the MSCI US equity index, as can be seen from figures 9.4, 9.5 and 9.6 below. Conversely, in the low return periods the portfolio exposure to the MSCI index is reduced in all the strategies. The synthetic option portfolio is particularly appealing to financial investors in a low-volatility environment, when the option will be cheaper. However, it would be interesting to compare the performance of both strategies. The asset universe for the optimisations performed in this thesis did not include the risk-free asset,

or options. Future work on the ALM problem will include the fixed-income products and options in the asset universe, to study the benefits of these instruments in the asset allocation. The asset allocation approach described here has the benefit of being able to include other asset classes which could substantially outperform the underlying asset, generating greater returns to the product issuer. A comparison of the performance of the synthetic portfolio approach with the asset allocation approach, considering different funding structures (i.e. from borrowing/gearing or the use of surplus funds), could therefore be very informative. It would also be interesting to include other asset classes, such as hedge funds and options, in the asset allocation universe. The inclusion of derivative instruments such as options in the stochastic optimisation problem requires the scenario generation method to exclude arbitrage opportunities in order to be consistent with financial asset pricing theory. Klaassen [43] demonstrated how arbitrage opportunities in a stochastic optimisation models could bias the solution, generating spurious profits. This issue was addressed by Vorst and Kouwenberg [44], who generated an arbitrage-free scenario tree while fitting the first two means of an underlying distribution. The addition of the arbitrage-free condition to the higher order moment matching scenario tree generator employed in this research is also an area of further development. In a mean-variance framework, alternative asset classes such as hedge funds and options have often been found to improve the performance of portfolios. L'Habitant [50] discusses the inadequacy of mean-variance based performance measures for portfolios with options. The extension of the higher order ALM model to fourth order moment measures would also be of interest. It would also be interesting to further investigate the relationship between varying the value for the guaranteed rate $\bar{g}$ gradually from 0

In this paper, in order to deal with the non-convexity of the skewness model, we employed a heuristic global optimisation scheme by using a warm-start approach. This approach does not guarantee a global optimum and lacks the rigour of more sophisticated techniques available to solve the global optimisation problem yielded by the skewness model (see, for example, [40],[41]). Further work is in progress to develop a minimax algorithm for global optimisation.

The stochastic models used in this thesis employed a three period scenario tree with three branches at each node. It would be of interest to apply the minimax robust optimisation method to more finely discretised scenario, which could arguably be more representative of reality. The minimax robust decision provides a worst-case optimal strategy which is guaranteed to improve if the worst-case does not occur. The combination of the minimax approach to large-scale stochastic programming problems with decomposition would also be an interesting area for further research.

9.3 Conclusion

The aim of this research was to develop an efficient tool for robust optimisation and successfully apply it to asset allocation problems with higher order moments under uncertainty. This has been achieved by the application of the minimax optimisation approach to asset allocation for robust asset liability management, with consideration also given to asymmetry and higher moments in return distributions. The asset allocation universe considered consists of the equity indices for five regions - Japan, Pacific excluding Japan, the US, the UK and Europe excluding the UK.

The efficacy of the minimax interior point algorithm has been demonstrated and the out-performance of the minimax robust optimisation methodology in some situations has been demonstrated through its application to an asset liability management problem. Though the size of the problems considered were relatively small, it has established the framework for solving such problems and set the foundation from some interesting areas of further research.

Part V

Appendix

Appendix 1: Examples of Performance of Merit functions 1 and 2

Example 1:

$$\begin{aligned} &\min f \\ &\text{s.t.} \\ &-x^3 \leq f, \\ &x^3 + 1 \leq f, \\ &x - 1 \leq f. \end{aligned}$$

Results:

- Initial value, $x_0 = (1.2)$.
 - l_2 norm merit function: $f^* = 0.99, x^* = -0.19$, iterations = 7. (Saddle Pt.)
 - Switching merit functions: $f^* = 0.5, x^* = -0.7937$, iterations = 9. (Minimum).

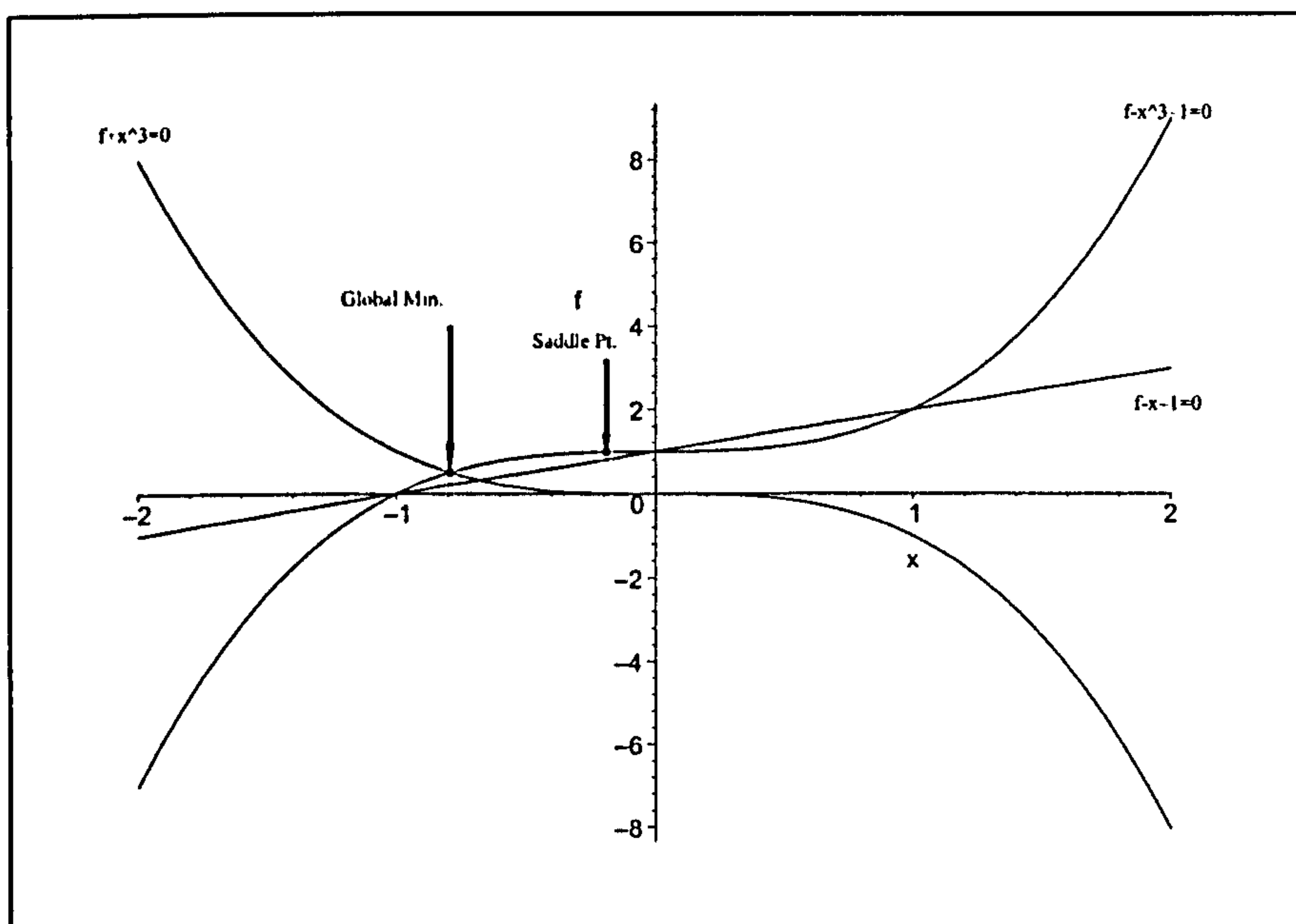


Figure 9.7: Example 1.

Example 2:

$$\begin{aligned} & \min f \\ & \text{s.t} \\ & x_1 * x_2 - 1 + x_1^2 + (x_2 - 1)^2 \leq f, \\ & x_1 * x_2 + 1 - x_1^2 - (x_2 - 1)^2 \leq f, \end{aligned}$$

Results:

- Initial value, $x_0 = (0.6, 1.3)$.
 - l_2 norm merit function: $f^* = 1.33333, x^* = (0.66667, 1.33333)$, iterations = 6. (Local Maximum).
 - Switching merit functions: $f^* = -1.29904, x^* = (-0.86603, 1.5)$, iterations = 13. (Global Minimum).

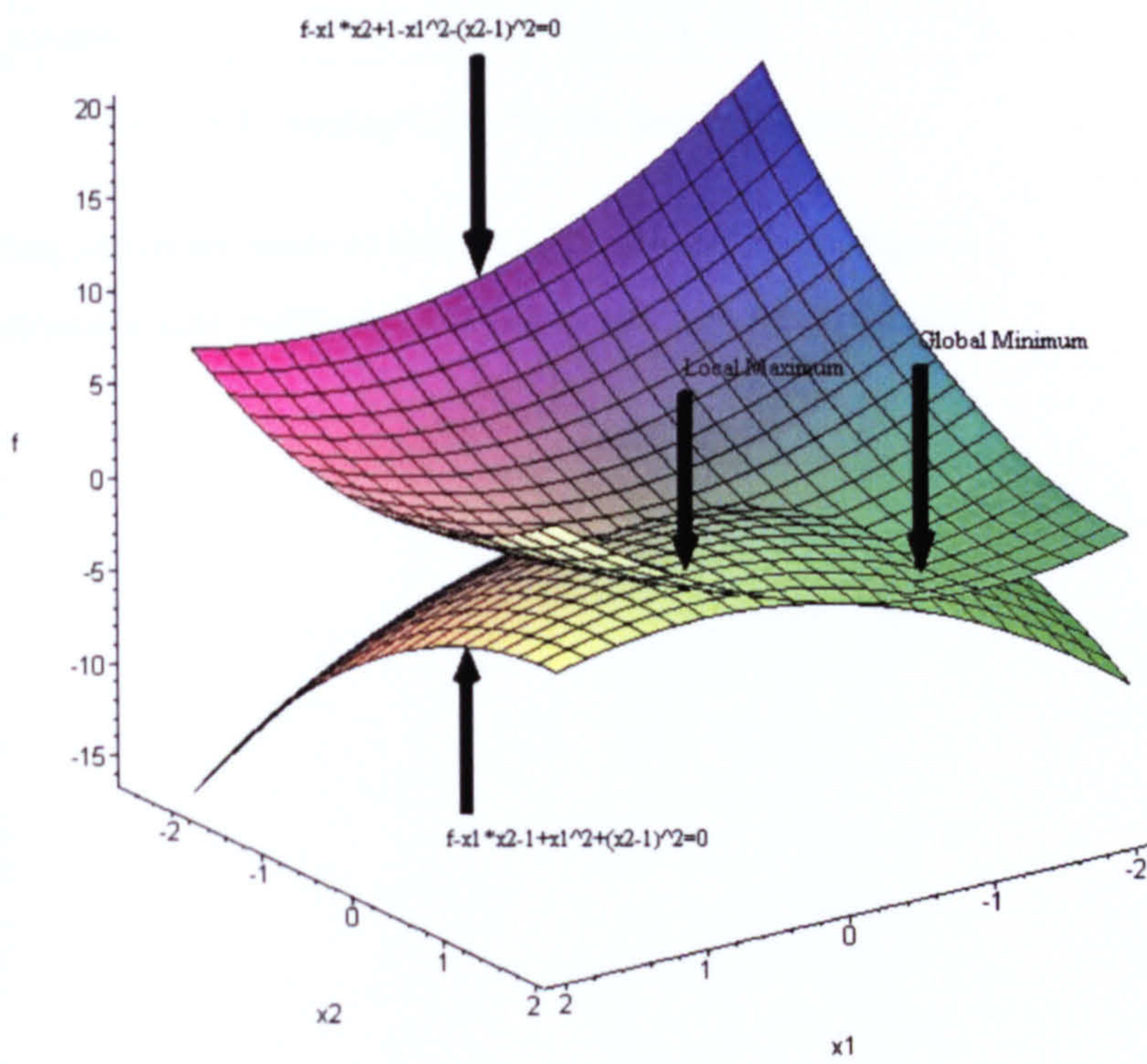


Figure 9.8: Example 2.

Appendix 2: Starting Values

Problem	X_0
goffin	$\mathbf{x}^0 = 0$
haifas	$\mathbf{x}^0 = 0$
madsen	$\mathbf{x}^0 = (0.3,1)$
makela4	$x_i^0 = 1$ for $i = 1, \dots 10$, $x_i^0 = -1$ for $i = 1, \dots 20$
mifflin2	$\mathbf{x}^0 = 0$
oct2	$\mathbf{x}^0 = 1$
polak1	$\mathbf{x}^0 = (1.0, 0.05)$
polak6	$\mathbf{x}^0 = (-1.0, 0.0, -1.0, 0.0)$

Table 9.3: Starting Points for the Test Problems.

All other starting points are same as those used in the problems listed at <http://www.orfe.princeton.edu/rvdb/ampl/nlmodels/cute/problemname.mod>.

Appendix 3: AMPL Models

```
#####
# Mean-Variance ALM.model      #
# created by Enitan Obasanjo    #
# 19/09/03                      #
#####
```

```
param NOBJ;
```

```
param STAGES;
set I ; #asset types
```

```
### The stochastic process : The event tree ###
```

```
#Parameters provided in .dat file#
```

```
param T > 0; # No. of time stages
```

```
param f{0..T-1}; # Number of branches arising from each node in period t.
```

```
#Computed parameters#
```

```
param N{t in 0..T}:= if t=0 then 1 else N[t-1]*f[t-1]; #number of nodes in period t.
```

```
param p{t in 0..T, n in 1..N[t]}>=0;
```

```
#transition probabilities from any node in period t to its successors
```

```
#in t+1 -given in data file.
```

```
param a{t in 1..T, n in 1..N[t], k in 1..t}:=
```

```
if k=1 then if t=1 then 1
```

```
else ceil(n/f[t-1]) else a[t-1,a[t,n,1],k-1]; #k-period predecessor function
```

```
param l{t in 1..T, n in 1..N[t]}:= n-f[t-1]*(a[t,n,1]-1); #last transition index
```

```
param P{t in 0..T, n in 1..N[t]}:= if t=0 then 1
```

```
else p[t,n]*P[t-1,a[t,n,1]]; #node probability
```

```
param b{t in 0..T, n in 1..N[T]}:= if t<T then a[t+1,b[t+1,n],1] else n;
```

```
#index of node traversed at time t by
```

```
#scenario ending at (T,n)
```

```
### The portfolio process ###
```

```
#portfolio parameters#
```

```
param g:= 0.025; #Transaction costs ratio
```

```
param alpha := 0.8; #Participation rate of product
```

```
param beta := 1.0; #Percentage of total portfolio value allowed to be held in any one asset
```

```
param g_bar := 0.12; #guaranteed rate of return per period
```

```
param kappa{t in 0..T, n in 1..N[t]} = 0.0; #assuming probability of actuarial events is deterministic
```

```
param H0{I}; #Initial portfolio
```

```
param L0:= 100; #Initial liability payment
```

```
param F{t in 0..T}; #Funding available in time period
```

```
param return{t in 1..T,n in 1..N[t],i in I}; #asset returns
```

```
param LR{t in 0..T,n in 1..N[t]}; #Liability returns - no longer used
```

```
param L{t in 0..T,n in 1..N[t]}:= if t=0 then L0
```

```
else (LR[t,n]+0.01)*L[t-1,a[t,n,1]]; #Liability
```

```
param LT{t in T..T,n in 1..N[T]}:= if (LR[T,n]+0.01)*L[T-1,a[T,n,1]]>= L0*(1+g_bar) then (LR[T,n]+0.01)*L[T-1,a[T,n,1]]
```

```
else L0*(1+g_bar); #Terminal liability
```

```

#Variables#
var H{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=0;
var B{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=0;
var S{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=0;

var TermWealth{n in 1..N[T]};

var A{t in 0..T, n in 1..N[t]};
var ExpTermWealth;

var pi_plus{n in 1..N[T]} >= 0, :=0; #auxillary variable
var pi_minus{n in 1..N[T]} >= 0, :=0; #auxillary variable
var UpperHoldLimit{t in 0..T-1, n in 1..N[t]};

minimize obj: sum{n in 1..N[T]}P[T,n]*(TermWealth[n] - ExpTermWealth)^2;

##### ASSET HOLDING CONSTRAINTS #####
subject to assetholding0{n in 1..N[0],i in I}: H[0,n,i] = H0[i]+B[0,n,i]-S[0,n,i];
subject to assetholding1{i in I, t in 1..T-1, n in 1..N[t]}:
H[t,n,i] = return[t,n,i]*H[t-1,a[t,n,i],i]+B[t,n,i]-S[t,n,i];

##### LIABILITY DYNAMICS CONSTRAINTS #####
subject to assetvalue{n in 1..N[T]}: sum{i in I}(H[T-1,a[T,n,i],i]*return[T,n,i]) + alpha*L[T,n] = A[T,n];
subject to Wealth{n in 1..N[T]}: A[T,n]-LT[T,n] = TermWealth[n] ;
subject to ExpectedTermwealth: sum{n in 1..N[T]} P[T,n]*(A[T,n]-LT[T,n]) = ExpTermWealth ;
subject to ExpWealthCons: sum{n in 1..N[T]} P[T,n]*(A[T,n]-LT[T,n])>= 300;

##### CASH BALANCE CONSTRAINTS #####
## This equation assumes that cash for buying is from funds added into the account from      #
## premiums - change this to one period borrowing and lending? and no liabilities till end of period? #
#####
#subject to cashbalance1{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g)- sum{i in I}S[0,n,i]*(1-g) = L0;
subject to cashbalance1{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g) - sum{i in I}S[0,n,i]*(1-g) = 0;
subject to cashbalance2{t in 1..T-1, n in 1..N[t]}: sum{i in I} B[t,n,i]*(1+g)- sum{i in I}S[t,n,i]*(1-g) = 0;
subject to cashbalance3{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g) = (1-alpha)*L0;
subject to poscon1{t in 1..T-1, n in 1..N[t]}: sum{i in I}H[t,n,i] = UpperHoldLimit[t,n];
subject to poscon2{t in 1..T-1, n in 1..N[t], i in I}: H[t,n,i] <= beta*UpperHoldLimit[t,n];

```

Appendix: The Robust MV ALM Model in AMPL

```

#####
# Robust Mean-Variance ALM model      #
# created by Enitan Obasanjo          #
# 19/09/03                            #
#####

```

```
param STAGES;
```

```

set I ; #asset types
param NOBS; #no. of historical observations

### The stochastic process : The event tree ###
#Parameters provided in .dat file#
param T > 0; # No. of time stages
param f{0..T-1}; # Number of branches arising from each node in period t.
#param p{t in 0..T-1, k in 1..f[t]}>=0; #transition probabilities from any node in period t to its successors in t+1

#Computed parameters#
param N{t in 0..T}:= if t=0 then 1 else N[t-1]*f[t-1]; #number of nodes in period t.

param p_1{t in 0..T, n in 1..N[t]}>=0; #transition probabilities from any node in period t to its successors

param p_2{t in 0..T, n in 1..N[t]}>=0; #transition probabilities from any node in period t to its successors

param p_3{t in 0..T, n in 1..N[t]}>=0; #transition probabilities from any node in period t to its successors

#in t+1 -given in data file.
param a{t in 1..T, n in 1..N[t], k in 1..t}:=
if k=1 then if t=1 then 1
else ceil(n/f[t-1]) else a[t-1,a[t,n,1],k-1]; #k-period predecessor function
param l{t in 1..T, n in 1..N[t]}:= n-f[t-1]*(a[t,n,1]-1); #last transition index
param P_1{t in 0..T, n in 1..N[t]}:= if t=0 then 1
else p_1[t,n]*P_1[t-1,a[t,n,1]]; #node probability
param P_2{t in 0..T, n in 1..N[t]}:= if t=0 then 1
else p_2[t,n]*P_2[t-1,a[t,n,1]]; #node probability
param P_3{t in 0..T, n in 1..N[t]}:= if t=0 then 1
else p_3[t,n]*P_3[t-1,a[t,n,1]]; #node probability

param b{t in 0..T, n in 1..N[T]}:= if t<T then a[t+1,b[t+1,n],1] else n; #index of node traversed at time t by
#scenario ending at (T,n)

### The portfolio process ###
#portfolio parameters#
param g:= 0.025; #Transaction costs ratio
#param R:= 0.9 ; #Risk level
param alpha := 0.8; #Participation rate of product
param beta := 1.0; #Percentage of total portfolio value allowed to be held in any one asset
param g_bar := 0.12 ; #guaranteed rate of return per period
param kappa{t in 0..T, n in 1..N[t]} = 0.0;
#assuming probability of actuarial events is deterministic and 0
param H0{I}; #Initial portfolio
param L0:= 100; #Initial liability payment
param F{t in 0..T}; #Funding available in time period
#param A{t in 0..T}; #Target for time period
param return_1{t in 1..T,n in 1..N[t],i in I}; #asset returns
param return_2{t in 1..T,n in 1..N[t],i in I}; #asset returns
param return_3{t in 1..T,n in 1..N[t],i in I}; #asset returns
param LR_1{t in 0..T,n in 1..N[t]}; #Liability returns
param LR_2{t in 0..T,n in 1..N[t]}; #Liability returns
param LR_3{t in 0..T,n in 1..N[t]}; #Liability returns
param L_1{t in 0..T,n in 1..N[t]}:= if t=0 then L0
else (LR_1[t,n]+0.01)*L_1[t-1,a[t,n,1]];

```

```

param L_2{t in 0..T,n in 1..N[t]}:= if t=0 then L0
else (LR_2[t,n]+0.01)*L_2[t-1,a[t,n,1]];
param L_3{t in 0..T,n in 1..N[t]}:= if t=0 then L0
else (LR_3[t,n]+0.01)*L_3[t-1,a[t,n,1]];

param LT_1{t in T..T,n in 1..N[T]}:= if (LR_1[T,n]+0.01)*L_1[T-1,a[T,n,1]]>= L0*(1+g_bar) then (LR_1[T,n]+0.01)*L_1[T-1,a[T,n,1]]
else L0*(1+g_bar); #Terminal liability amount
param LT_2{t in T..T,n in 1..N[T]}:= if (LR_2[T,n]+0.01)*L_2[T-1,a[T,n,1]]>= L0*(1+g_bar) then (LR_2[T,n]+0.01)*L_2[T-1,a[T,n,1]]
else L0*(1+g_bar); #Terminal liability amount
param LT_3{t in T..T,n in 1..N[T]}:= if (LR_3[T,n]+0.01)*L_3[T-1,a[T,n,1]]>= L0*(1+g_bar) then (LR_3[T,n]+0.01)*L_3[T-1,a[T,n,1]]
else L0*(1+g_bar); #Terminal liability amount

param asset_res{j in 1..NOBS,i in I};
param liab_res{j in 1..NOBS};

param Hopt{t in 0..T-1,n in 1..N[t],i in I};
param Bopt{t in 0..T-1,n in 1..N[t],i in I};
param Sopt{t in 0..T-1,n in 1..N[t],i in I};

param varval;
#Variables#
var H{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=100;
var B{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=100;
var S{t in 0..T-1,n in 1..N[t],i in I} >= 0, :=100;

var TermWealth_1{n in 1..N[T]} >= 0, :=1;
var TermWealth_2{n in 1..N[T]} >= 0, :=1;
var TermWealth_3{n in 1..N[T]} >= 0, :=1;

## auxillary variables
var pi_plus_1{n in 1..N[T]} >= 0, :=1;
var pi_plus_2{n in 1..N[T]} >= 0, :=1;
var pi_plus_3{n in 1..N[T]} >= 0, :=1;

var pi_minus_1{n in 1..N[T]} >= 0, :=1;
var pi_minus_2{n in 1..N[T]} >= 0, :=1;
var pi_minus_3{n in 1..N[T]} >= 0, :=1;

var LTerm_1{n in 1..N[T]}>=0;
var LTerm_2{n in 1..N[T]}>=0;
var LTerm_3{n in 1..N[T]}>=0;

var A_1{t in 0..T, n in 1..N[t]};
var A_2{t in 0..T, n in 1..N[t]};
var A_3{t in 0..T, n in 1..N[t]};

var UpperHoldLimit_1{t in 0..T-1, n in 1..N[t]};
var UpperHoldLimit_2{t in 0..T-1, n in 1..N[t]};
var UpperHoldLimit_3{t in 0..T-1, n in 1..N[t]};
var ExpTermWealth_1;
var ExpTermWealth_2;
var ExpTermWealth_3;

```

```
var v; #>= 0;
```

```
minimize nmobj: v;
```

```
##### ASSET HOLDING CONSTRAINTS_1 #####
```

```
subject to assetholding0_1{n in 1..N[0], i in I}: H[0,n,i] = H0[i]+B[0,n,i]-S[0,n,i];
```

```
subject to assetholding1_1{i in I, t in 1..T-1, n in 1..N[t]}:
```

```
H[t,n,i] = return_1[t,n,i]*H[t-1,a[t,n,i],i]+B[t,n,i]-S[t,n,i];
```

```
##### LIABILITY DYNAMICS CONSTRAINTS_1 #####
```

```
subject to assetvalue_1{n in 1..N[T]}: sum{i in I}(H[T-1,a[T,n,i],i]*return_1[T,n,i]) + alpha*L_1[T,n] = A_1[T,n];
```

```
subject to Wealth_1{n in 1..N[T]}: A_1[T,n]-LT_1[T,n] = TermWealth_1[n];
```

```
subject to ExpectedTermwealth_1: sum{n in 1..N[T]} P_1[T,n]*(A_1[T,n]-LT_1[T,n]) = ExpTermWealth_1;
```

```
subject to ExpWealthCons_1: sum{n in 1..N[T]} P_1[T,n]*TermWealth_1[n] >=300;
```

```
subject to VarCons_1: sum{n in 1..N[T]} P_1[T,n]*((TermWealth_1[n])-ExpTermWealth_1)^2 <= v;
```

```
##### LIABILITY DYNAMICS CONSTRAINTS_2 #####
```

```
subject to assetvalue_2{n in 1..N[T]}: sum{i in I}(H[T-1,a[T,n,i],i]*return_2[T,n,i]) + alpha*L_2[T,n] = A_2[T,n];
```

```
subject to Wealth_2{n in 1..N[T]}: A_2[T,n]-LT_2[T,n] = TermWealth_2[n] ;
```

```
subject to ExpectedTermwealth_2: sum{n in 1..N[T]} P_2[T,n]*(A_2[T,n]-LT_2[T,n]) = ExpTermWealth_2;
```

```
subject to ExpWealthCons_2: sum{n in 1..N[T]} P_2[T,n]*TermWealth_2[n] >=300;
```

```
subject to VarCons_2: sum{n in 1..N[T]} P_2[T,n]*((TermWealth_2[n])-ExpTermWealth_2)^2 <= v;
```

```
##### LIABILITY DYNAMICS CONSTRAINTS_3 #####
```

```
subject to assetvalue_3{n in 1..N[T]}: sum{i in I}(H[T-1,a[T,n,i],i]*return_3[T,n,i]) + alpha*L_3[T,n] = A_3[T,n];
```

```
subject to Wealth_3{n in 1..N[T]}: A_3[T,n]-LT_3[T,n] = TermWealth_3[n];
```

```
subject to ExpectedTermwealth_3: sum{n in 1..N[T]} P_3[T,n]*(A_3[T,n]-LT_3[T,n]) = ExpTermWealth_3 ;
```

```
subject to ExpWealthCons_3: sum{n in 1..N[T]} P_3[T,n]*TermWealth_3[n] >=300;
```

```
subject to VarCons_3: sum{n in 1..N[T]} P_3[T,n]*((TermWealth_3[n])-ExpTermWealth_3)^2 <=v;
```

```
##### CASH BALANCE CONSTRAINTS_1 #####
```

```
## This equation assumes that cash for buying is from funds added into the account from
```

```
## premiums - change this to one period borrowing and lending? and no liabilities till end of period?
```

```
#subject to cashbalance1_1{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g)- sum{i in I}S[0,n,i]*(1-g) = L0;
```

```
subject to cashbalance3_1{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g) = (1-alpha)*L0;
```

```
subject to cashbalance1_1{n in 1..N[0]}: sum{i in I} B[0,n,i]*(1+g) - sum{i in I}S_1[0,n,i]*(1-g) = 0;
```

```
subject to cashbalance2_1{t in 1..T-1, n in 1..N[t]}: sum{i in I} B[t,n,i]*(1+g)- sum{i in I}S[t,n,i]*(1-g) = 0;
```

```
subject to poscon1_1{t in 1..T-1, n in 1..N[t]}: sum{i in I}H[t,n,i] = UpperHoldLimit_1[t,n];
```

```
subject to poscon2_1{t in 1..T-1, n in 1..N[t], i in I}: H[t,n,i] <= beta*UpperHoldLimit_1[t,n];
```

Appendix: Global Optimisation Routine in AMPL

```
option solver nmxipl;
```

```
param nprove;
```

```
param ctr;

param count := BigNum;

param switch;

let switch := 0;

param globalopt;

let globalopt := Inf;

param Hinit{t in 0..T-1,n in 1..N[t],i in I}:=100;
param Binit{t in 0..T-1,n in 1..N[t],i in I}:=100;
param Sinit{t in 0..T-1,n in 1..N[t],i in I}:=100;
param MinHold := 0;

let{t in 0..T-1,n in 1..N[t],i in I}Hopt[t,n,i] := Hinit[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Bopt[t,n,i] := Binit[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Sopt[t,n,i] := Sinit[t,n,i];

let nprove := 0;

repeat while nprove < count {
  let nprove := nprove + 1;
  display nprove;
  if (switch = 1) then # use discretised in feasible region
  {
    let {t in 0..T-1,n in 1..N[t],i in I} H[t,n,i] := Hinit[t,n,i]+((100*UpperHoldLimit[t,n]-MinHold)/count)*nprove;
    let {t in 0..T-1,n in 1..N[t],i in I} B[t,n,i] := Binit[t,n,i]+((100*UpperHoldLimit[t,n]-MinHold)/count)*nprove;
    let {t in 0..T-1,n in 1..N[t],i in I} S[t,n,i] := Sinit[t,n,i]+((100*UpperHoldLimit[t,n]-MinHold)/count)*nprove;

  }
  else # use warmstarts
  {
    let{t in 0..T-1,n in 1..N[t],i in I}H[t,n,i] := Hopt[t,n,i];
    let{t in 0..T-1,n in 1..N[t],i in I}B[t,n,i] := Bopt[t,n,i];
    let{t in 0..T-1,n in 1..N[t],i in I}S[t,n,i] := Sopt[t,n,i];
  }

  solve;

  display ExpTermWealth;
  display sum{n in 1..N[T]} P[T,n]*(TermWealth[n]-ExpTermWealth)^3;
  display sum{n in 1..N[T]} P[T,n]*(TermWealth[n]-ExpTermWealth)^2;
  display sum{n in 1..N[T]} P[T,n]*(TermWealth[n]);
  display {n in 1..N[T]} P[T,n];
```

```

display {n in 1..N[T]} L[T,n];
display {n in 1..N[T]} LT[T,n];

display solve_result_num;

let{t in 0..T-1,n in 1..N[t],i in I}Hopt[t,n,i] := H[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Bopt[t,n,i] := B[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Sopt[t,n,i] := S[t,n,i];

##if (obj > globalopt ) then {

if ( globalopt > obj ) then {    #use this for minimisation problems
if (solve_result_num < 200) then { #check that solver completes without a fail

let{t in 0..T-1,n in 1..N[t],i in I}Hopt[t,n,i] := H[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Bopt[t,n,i] := B[t,n,i];
let{t in 0..T-1,n in 1..N[t],i in I}Sopt[t,n,i] := S[t,n,i];

let globalopt := obj;
let switch := 1;    #once the problem is solved to optimality, use increments instead of warmstarts
}

}

#display TerminalWealth, globalopt;

display obj, globalopt;
}
;

solve;

```

Bibliography

- [1] I. Akrotirianakis and B. Rustem. Design and implementation of a primal-dual interior point algorithm for nonlinear programming. Technical Report 2000-08, Department of Computing, Imperial College of Science and Technology, 2000.
- [2] I. Akrotirianakis and B. Rustem. A globally convergent interior point algorithm for non-linear programming problems. Technical report, Department of Computing, Imperial College of Science and Technology, 2005.
- [3] F. D. Arditti. Risk and the required return on equity. *Journal of Finance*, 22 (1):19–36, 1967.
- [4] F. D. Arditti and H. Levy. Portfolio efficiency analysis in three moments: The multi-period case. *Journal of Finance*, 30 (3):797–809, 1975.
- [5] V. S. Bawa. Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2(1):95–121, 1975.
- [6] V. S. Bawa. Safety first, stochastic dominance and optimal portfolio choice. *Journal of Financial and Quantitative Analysis*, 3(1):225–271, 1978.
- [7] A. Ben-Tal, T. Margalit, and A. Nemirovski. Robust modeling of multi-stage portfolio problems. In *High Performance Optimization*, pages 303–328, Dordrecht, 2000. Kluwer Acad. Publ.
- [8] H. Benson. The cute test set in ampl format. www.orfe.princeton.edu/rvdb/ampl/nlmodels/cute/, 1996.
- [9] H. Benson, R. Vanderbei, and D. Shanno. Interior point methods for nonconvex nonlinear programming: Jammimg and comparative numerical testing. Technical Report ORFE-00-02, Operations Research and Financial Engineering, Princeton University, Princeton, N. J., 2000.
- [10] X. Cai, K.L. Teo, X. Q. Yang, and X. Y. Zhou. Portfolio optimization under a minimax rule. *Management Science*, 46:957–972, 2000.

- [11] X. Cai, K.L. Teo, X. Q. Yang, and X. Y. Zhou. Minimax portfolio optimization: Empirical study. *Journal of the Operations Research Society*, 55:65–72, 2004.
- [12] D. Carino and T. Kent et al. The russell-yasuda kasai model: An asset/liability model for a japanese insurance company using multistage stochastic programming. *Interfaces*, 23:29–49, 1994.
- [13] R. M. Chamberlain, C. Lemarechal, H. C. Pedersen, and M. J. D. Powell. The watchdog technique for forcing convergence in algorithms for constrained optimization. *Mathematical Programming Study*, 16:1–17, 1982.
- [14] V. K. Chopra. Improving optimization. *Journal of Investing*, 8:51–59, 1993.
- [15] P. Chunhachinda, K. Dandapani, S. Hamid, and A. Prakash. Portfolio selection and skewness: Evidence from international stock markets. *Journal of Banking and Finance*, 21:143–167, 1997.
- [16] G. Consigli and M. A. H. Dempster. Dynamic stochastic programming for asset liability management. *Annals of Operations Research*, 81(1):131–162, 1998.
- [17] A. Consiglio, F. Cocco, and S. A. Zenios. Scenario optimization asset and liability modeling for endowments with guarantees. Technical Report 00-41 B, Wharton School, University of Pennsylvania, February 2001.
- [18] G. M. de Athayde. A few more moments to finance theory: Introducing higher moments in investment theory and econometrics, 2001.
- [19] J. Dupacova, G. Consigli, and S. Wallace. Scenarios for multistage stochastic programs. *Annals of OR*, 100:25–53, 2001.
- [20] S. Zenios (editor). *Financial Optimization*. Cambridge University Press, 1993.
- [21] A. El-Bakry, R. Tapia, T. Tsuchiya, and Y. Zhang. On the formulation and theory of the newton interior point method for nonlinear programming. *Journal of Optimisation Theory and Applications*, 89(3):507–541, 1996.

- [22] P. C. Fishburn. Mean-risk analysis with risk associated with below-target returns. *American Economic Review* Vol, 67:116–126, 1977.
- [23] R. Fourer, D. Gay, and B. Kernighan. *AMPL: A Modelling Language for Mathematical Programming*. Boyd and Fraser Publishing, 1993.
- [24] K. Frauendorfer. The stochastic programming extension of the markowitz approach. *J. Mass-Parallel Computing Information Systems*, 53:449–460, 1995.
- [25] P. Frost and J. Savarino. For better performance constrain portfolio weights. *Journal of Portfolio Management*, 15:29–34, 1988.
- [26] L. Garalppi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty:a multi-prior approach. *submitted to the Review of Financial Studies*, 2004.
- [27] D. M. Gay, M. L. Overton, and M. L. Wright. A primal-dual interior point method for non-convex programming. *Advances in Nonlinear Programming*, pages 31–56, 1998.
- [28] I. Gilboa and D. Schmeidler. Maximin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 16:65–88, 1989.
- [29] P. E. Gill and J. W. Murray. *Practical Optimization*. Academic Press, 1981.
- [30] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28:1–38, 2003.
- [31] J. Gondzio and R. Kouwenberg. High performance computing for asset-liability management. *Operations Research*, 49(6):879–891, 2001.
- [32] P. E. Green. Risk attitudes and chemical investment decisions. *Chemical Engineering Progress*, 59:35–40, 1963.

-
- [33] M. Guidolin and A. Timmerman. International asset allocation under regime switching, skew and kurtosis preferences. Technical Report 2005-034, Federal Reserve Bank of St. Louis, 2005.
- [34] N. Gulpinar and B. Rustem. Optimisation and simulation approaches to scenario tree generation. *Journal of Economics Dynamics and Control*, 28 (3):1291–1315, 2004.
- [35] N. Gulpinar, B. Rustem, and R. Settergren. Multistage stochastic mean-variance portfolio analysis with transaction cost. *Innovations in Financial and Economic Networks*, 3:46–63, 2003.
- [36] L. P. Hansen and T. J. Sargent. Robust control and economic model uncertainty.
- [37] C. R. Harvey, J. C. Liechty, M. Liechty, and P. Muller. Portfolio selection with higher moments. Technical Report W70, Duke University, Durham, North Carolina, 2003.
- [38] S. J. Jobst and Stavros Zenios. Extending credit risk models for the simulation and valuation of portfolios of interest rate and credit risk sensitive securities. Technical report, HERMES Centre for Finance, University of Cyprus, 2001.
- [39] N. K. Kamarkar. A new polynomial time algorithm for linear programming. *Combinatorica*, 4(3):373–395, 1984.
- [40] A. H. G. Rinnooy Kan and G. T. Timmer. Stochastic global optimisation methods, part i: clustering. *Math. Prog*, 39:27–56, 1987.
- [41] A. H. G. Rinnooy Kan and G. T. Timmer. Stochastic global optimisation methods, part ii: Multilevel methods. *Math. Prog*, 39:57–78, 1987.
- [42] M. Kaut and S. Wallace. Evaluation of scenario-generation methods for stochastic programming. Technical Report SPEPS 14-2003, Molde University College, Norway, 2003.

- [43] P. Klaassen. Discretized reality and spurious profits in stochastic programming models for asset/liability management. *European Journal of Operational Research*, 101(2):374–392, 1997.
- [44] R. Klaassen and T. Vorst. Dynamic portfolio insurance: A stochastic programming approach. Technical report, Erasmus University, 1998.
- [45] M. Kojima, S. Mizuno, and A. Yoshise. Primal-dual interior point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming: Interior Point and Related Methods*, pages 29–47. Springer-Verlag, 1989.
- [46] H. Konno and K. Suzuki. Mean-variance-skewness portfolio optimisation model. *Journal of the Operations Research Society of Japan*, 38:173–187, 1995.
- [47] R. Kouwenberg and S. Zenios. Stochastic programming models for asset liability management. In *Handbook of Asset and Liability Modeling*, 2001.
- [48] H. W. Kuhn and A. W. Tucker. Non-linear programming. In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, 1951.
- [49] T. Y. Lai. Portfolio selection with skewness: A multiobjective approach. *Review of Quantitative Finance and Accounting*, 1:293–305, 1991.
- [50] F. L'Habitant. Enhancing portfolio performance using options strategies. why beating the market is easy. Technical report, University of Lausanne, 1998.
- [51] H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.
- [52] J. Medanic and M. Andjelic. On a class of differentiable games without saddle point solutions. *Journal of Optimisation Theory and Applications*, 8:413–430, 1971.
- [53] J. Medanic and M. Andjelic. Minmax solution of the multiple target problem. *IEEE Transactions on Automatic Control AC*, 17:597–604, 1992.
- [54] R. Michaud. *Efficient Asset Management*. Harvard Business School Press, 1998.

-
- [55] J. M. Mulvey and A. E. Thorlacius. The towers perrin global capital market scenario generation system. In *Worldwide Asset Liability Modelling*, 1998.
- [56] D. Nawrocki. The characteristics of portfolios selected by n -degree lower partial moment. *International Review of Financial Analysis*, 1(3):195–209, 1992.
- [57] S. Nielsen and S. A. Zenios. A stochastic programming model for funding single-premium deferred annuities. *Mathematical Programming Series B*, 75:177–200, 1996.
- [58] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, New York, 1999.
- [59] E. Obasanjo and B. Rustem. An interior point method for nonlinear minimax problems.
- [60] R. B. Porter. Semivariance and stochastic dominance: A comparison. *The American Review*, 64:200–204, 1974.
- [61] A. Prakash, C-H. Chang, and T. Pactwa. Selecting a portfolio with skewness: Recent evidence from u.s., european, and latin american equity markets. *Journal of Banking and Finance*, 27:1375–1390, 2003.
- [62] M. Rubinstein. The fundamental theorem of parameter preference and security valuation. *Journal of Financial and Quantitative Analysis*, 8:61–69, 1973.
- [63] B. Rustem. Convergent stepsizes for constrained optimization algorithms. *Journal of Optimisation Theory and Applications*, 49:136–160, 1986.
- [64] B. Rustem. *Algorithms for Nonlinear programming and Multiple-Objective decisions*. Wiley, 1992.
- [65] B. Rustem. A constrained min-max algorithm for rival models of the same economic system. *Mathematical Programming*, 53:279–295, 1992.

- [66] B. Rustem. Equality and inequality constrained optimization algorithms with convergent stepsizes. *Journal of Optimisation Theory and Applications*, 76:429–453, 1993.
- [67] B. Rustem and Q. Nguyen. An algorithm for the inequality-constrained discrete min-max problem. *SIAM Journal of Optimisation*, 8(2):265–283, 1998.
- [68] B. Rustem and Q. Nguyen. An algorithm for the inequality constrained discrete minimax problem. *SIAM Journal of Optimisation*, 8(2):265–283, 1998.
- [69] B. Rustem and R. Settergren. Scenario specification for robust analysis. In *Computational Methods in Decision Making, Economics and Finance: Optimization Models*. Kluwer Academic Publishers, 2002.
- [70] P. Samuelson. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Review of Economic Studies*, 37 (1/2):537–542, 1970.
- [71] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52:591–611, 1965.
- [72] S. S. Shapiro, M. B. Wilk, and Chen. A comparative study of various tests for normality. *Journal of American Statistical Association*, 63:1343–1372, 1968.
- [73] M. Simkowitz and W. Beedles. Diversification in a three moment world. *Journal of Financial and Quantitative Analysis*, 13:927–941, 1978.
- [74] J. C. Singleton and J. Wingender. Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analysis*, 21 (3):335–341, 1986.
- [75] Q. Sun and Y. Yan. Skewness persistence with optimal portfolio selection. *Journal of Banking and Finance*, 27:1111–1121, 2003.
- [76] R. H. Tutuncu and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132:157–187, 2004.