Multistage Stochastic Mean-Variance Portfolio Analysis with Transaction Costs

Nalan Gülpınar, Berç Rustem,* Reuben Settergren

Department of Computing,
Imperial College of Science, Technology and Medicine
180 Queen's Gate, London, SW7 2RH, UK.
[ng2, br, reuben]@doc.ic.ac.uk

May 17, 2004

Abstract

Multistage stochastic programming is used to model the problem of financial portfolio management with transaction costs, given stochastic data provided in the form of a scenario tree. The mean or variance of total wealth at the end of the planning horizon can be optimized in view of the transaction costs by solving either a linear stochastic program or a quadratic stochastic program, respectively; solution of many almost identical quadratic stochastic programs yields points describing the Markowitz efficient frontier. The incorporation of proportional transaction costs leads to a model that reflects the effect of these costs on portfolio performance. It is also important with a quadratic model to ensure complementarity between buys and sells of the same asset. Numerical experiments backtesting the optimization strategies at different levels of risk and transaction cost are reported, as well as tests that do not optimize over the effected transaction costs.

Keywords Multistage portfolio optimization, transaction costs, stochastic programming, quadratic programming

1 Introduction

A rational framework for investment decisions is provided by the maximization of return for an acceptable level of risk. A fundamental example is the single-period Markovitz [20] model in which expected portfolio return is maximized and risk measured by the variance of portfolio return is minimized.

Consider n risky assets with random rates of return $r_1, r_2, \cdots, r_n$. Their expected values are denoted $E(r_i)$, $i = 1, \cdots, n$. A full description of our notation is given in Table 1. All quantities in boldface represent vectors in $\mathbb{R}^n$ unless otherwise noted. For example, $\mathbf{r}$ denotes the column vector whose ith element is $E(r_i)$. The transpose of a vector or matrix will be denoted with the symbol $'$.

The single period model of Markowitz considers a portfolio of n assets defined in terms of a set of weights w^i for $i = 1, \cdots, n$, which sum to unity. Many traditional portfolio analysis

*Published in *Innovations in Financial and Economic Networks*, Vol. 3, 2003, p.46–63, Ed. A.Nagurney, Edward Elgar Publishing Inc.

models seek only to maximize expected return. This can be achieved with a classical stochastic linear programming formulation which incorporates the mean term. This is a risk-neutral approach which does not take risk-attitudes into account. Stochastic linear programming is mathematical programming (with linear objective function and constraints) under uncertainty where one or more parameters are not known at the time of decision making. For further information on stochastic programming, the reader is referred to [2, 17, 24].

Given an expected rate of return $\bar{r}$, the optimal portfolio is defined in terms of the solution of the following quadratic programming problem:

$$\min_{\mathbf{w}} \{< \mathbf{w}, \Lambda \mathbf{w} > | \mathbf{w}'\mathbf{r} = \bar{r}, \;\; \mathbf{1}'\mathbf{w} = 1, \;\; \mathbf{w} \geq 0\}$$

where Λ is the covariance matrix of asset returns. The quadratic program yields the minimum variance portfolio. Note that the classical stochastic linear programming formulation maximizes the expected return but takes no account of risk.

The single period mean-variance optimization problem can be extended to multistage programming. In this case, after the initial investment, we can rebalance our portfolio (subject to any desired bounds) to maximize profit at the investment horizon and minimize the risk at discrete time periods and redeem at the end of the period. Uncertainty is represented by the mathematical methods such as mean-variance analysis, utility function analysis and arbitrage analysis, for example see [18, 23]. The mean-variance model has some inconsistencies such as the assumption of normal returns, quadratic utility and shifting efficient frontiers as the number of periods in the model changes. While these are serious issues, we believe that a framework that admits a longer term perspective and incorporates transaction costs is important. The issue of normality can be relaxed in due course by using higher order moments. The same is also partially true for the quadratic variance term. Including skewness and kurtosis is possibly an attractive extension of the present work, despite the added computational complexity. The sensitivity of the result to the length of horizon considered, on the other hand, is common to all dynamic optimization problems. It is somewhat alleviated by considering the sensitivity of the first-period decision to differing horizons. A discount factor for future risk is also a relevant option. The purpose of this paper is to discuss the basic tool and its use in view of transaction costs to overcome the well known suboptimality of the sequential application of single period optimal decisions.

The multistage decision problems under uncertainty are considered in [1, 5, 6, 25] for nonlinear problems. A mean-variance approach is adopted with a static transcription of the dynamic decision model. This utilises analytical second order approximations to the mean and variance functions. In [9, 28], however, a scenario-based discretisation is developed and a stochastic programming model formulated. Gassman and Ireland [12] presents scenario formulation in an algebraic modelling language. An alternative to the standard mean-variance approach for controlling risk that is based on a concept of wealth accumulation reflecting investor preferences was proposed in [10, 30].

In this paper, we consider the multistage extension of the mean-variance optimization problem. The first distinguishing feature of the mean-variance approach over expected value optimization, formulated as linear stochastic programming, is that the former takes into account the approximate nature of the set of discrete scenarios by considering the variance around each return scenario. The second feature of the variance term is that it does allow for the variability of returns over the scenario tree. Consequently, uncertainty on return values of instruments is represented by a discrete approximation of a multivariate continuous distribution as well as the variability due to the discrete approximation. This is discussed further in Section 2.

Another issue considered in the paper is the adoption of benchmark-relative computations in view of the transaction costs and the use of a powerful interior-point algorithm [21] to compute the overall solution. The existence of transaction costs makes it essential that the multistage decision problem is addressed directly. The application of single period optimization sequentially clearly leads to suboptimal results as overall investment performance depends on these costs. Single period optimization cannot incorporate the effect of future transaction of assets considered for the single period. Multistage optimization thus overcomes the suboptimality of myopic single period optimization.

A quadratic (linearly constrained) stochastic programming model is developed, and its implementation tested. Data from large-scale problems in the literature [8] are used to examine run-time. Backtesting is performed on historical data to examine the model's ability to compute Markowitz efficient investment strategies in a multistage context.

The rest of the paper is organised as follows. In Section 2, the problem statement is given. In Section 3, we present the multistage stochastic mean-variance portfolio optimization model (based on scenario tree). In section 4, the computational results are shown and conclusions are presented in Section 5.

2 Problem Statement

The central problem considered in the paper is the determination of multi-period discrete-time optimal portfolio strategies over a given finite investment horizon. Therefore, we start with the definition of returns and uncertainties. Subsequently, we present the model constraints, the expected return and risk formulations based on the scenario tree.

We consider n risky assets and construct a portfolio over an investment horizon T. The portfolio is restructured over a period in terms of both return and risk. After the initial investment ($t = 0$), the portfolio may be restructured at discrete times $t = 1, \dots, T-1$, and redeemed at the end of the period ($t = T$).

Scenario Tree

Let $\boldsymbol{\rho}^t \equiv \{\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_t\}$ be stochastic events at $t = 1, \dots, T$. The decision process is non-anticipative (i.e decision at a given stage does not depend on the future realization of the random events). The decision at t is dependent on $\boldsymbol{\rho}_{t-1}$. Thus, constraints on a decision at each stage involve past observations and decisions.

A scenario is defined as a possible realisation of the stochastic variables $\{\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_T\}$. Hence, the set of scenarios corresponds to the set of leaves of the scenario tree, $\mathcal{N}_T$, and nodes of the tree at level $t \geq 1$ (the set $\mathcal{N}_t$) correspond to possible realisations of $\boldsymbol{\rho}^t$. We denote a node of the tree (or event) by $\mathbf{e} = (s, t)$, where s is a scenario (path from root to leaf), and time period t specifies a particular node on that path. The root of the tree is $\mathbf{0} = (s, 0)$ (where s can be any scenario, since the root node is common to all scenarios). The ancestor (parent) of event $\mathbf{e} = (s, t)$ is denoted $a(\mathbf{e}) = (s, t-1)$, and the branching probability $p_{\mathbf{e}}$ is the conditional probability of event $\mathbf{e}$, given its parent event $a(\mathbf{e})$. The path to event $\mathbf{e}$ is a partial scenario with probability $P_{\mathbf{e}} = \prod p_{\mathbf{e}}$ along that path; since probabilities $p_{\mathbf{e}}$ must sum to one at each individual branching, probabilities $P_{\mathbf{e}}$ will sum up to one across each layer of tree-nodes $\mathcal{N}_t; t = 0, 1, \dots, T$. Each node at a level t corresponds to a decision $\{\mathbf{w}_t, \mathbf{b}_t, \mathbf{s}_t\}$ which must be determined at time t, and depends in general on $\boldsymbol{\rho}^t$, the initial wealth $\mathbf{w}_0$ and past decisions $\{\mathbf{w}_j, \mathbf{b}_j, \mathbf{s}_j\}, j = 1, \dots, t-1$. This process is adapted to $\boldsymbol{\rho}^t$ as $\mathbf{w}_t, \mathbf{b}_t, \mathbf{s}_t$ cannot depend on future events $\boldsymbol{\rho}_{t+1}, \dots, \boldsymbol{\rho}_T$ which are not yet realised. Hence $\mathbf{w}_t = \mathbf{w}_t(\boldsymbol{\rho}^t)$, $\mathbf{b}_t = \mathbf{b}_t(\boldsymbol{\rho}^t)$, and $\mathbf{s}_t = \mathbf{s}_t(\boldsymbol{\rho}^t)$. However, for simplicity, we shall use the terms $\mathbf{w}_t$, $\mathbf{b}_t$ and $\mathbf{s}_t$, and assume their implicit dependence on $\boldsymbol{\rho}^t$. Notice

Table 1: Notation

n	number of investment assets.
T	planning horizon. Initial investment is at $t = 0$, the portfolio is restructured at discrete times $t = 1, \dots, T-1$, and finally redeemed at $t = T$.
$\boldsymbol{\rho}_t$	vector of stochastic data observed at time t, $t = 0, \dots, T$.
$\boldsymbol{\rho}^t$	$\equiv \{\boldsymbol{\rho}_0, \dots, \boldsymbol{\rho}_t\}$ — history of stochastic data up to t.
$\mathcal{N}$	set of all nodes in the scenario tree.
$\mathcal{N}_t$	set of nodes of the scenario tree representing possible events at time t.
$\mathcal{N}_I$	$\equiv \mathcal{N} - (\mathcal{N}_0 \cup \mathcal{N}_T)$, i.e. set of all *interior* nodes of the scenario tree.
s	index denoting a scenario, i.e. path from root to leaf in the scenario tree.
$\mathbf{e} \equiv (s,t)$	index denoting an event (node of the scenario tree), which can be identified by an ordered pair of scenario and time period.
$a(\mathbf{e})$	ancestor of event $\mathbf{e} \in \mathcal{N}$ (parent in the scenario tree).
$p_{\mathbf{e}}$	branching probability of event $\mathbf{e}$: $p_{\mathbf{e}} = \text{Prob}[\mathbf{e} \vert a(\mathbf{e})]$
$P_{\mathbf{e}}$	probability of event $\mathbf{e}$: if $\mathbf{e} = (s,t)$, then $P_{\mathbf{e}} = \prod_{i=1\dots t} p_{(s,i)}$.
$\mathrm{E}\left[\cdot\right]$	Expectation with respect to ρ.
$\mathbf{p}$	Current portfolio position (i.e. at $t = 0$, before optimization).
$\mathbf{c}_b$	vector of unit transaction costs for buying.
$\mathbf{c}_s$	vector of unit transaction costs for selling.
$\mathbf{b}_*$	decision vector of "buy" transaction volumes.
$\mathbf{s}_*$	decision vector of "sell" transaction volumes.
$\mathbf{r}_t(\boldsymbol{\rho}^t)$	stochastic vector of return values for the n assets, $t = 1, \dots, T$.
$\Lambda \in \mathbb{R}^{n \times n}$	covariance matrix associated with return values
$\mathbf{r}_{\mathbf{e}}$	stochastic realisation of $\mathbf{r}_t$ in event $\mathbf{e}$: $\mathbf{r}_{\mathbf{e}} \sim N(\mathbf{r}_t(\boldsymbol{\rho}^t), \Lambda)$.
$\hat{\mathbf{r}}_{\mathbf{e}}$	$\equiv \mathrm{E}\left(\mathbf{r}_{\mathbf{e}}(\boldsymbol{\rho}_t \vert \boldsymbol{\rho}^{t-1})\right)$ — expectation of $\mathbf{r}_t(\boldsymbol{\rho}^t)$ for event $\mathbf{e}$, conditional on $\boldsymbol{\rho}^{t-1}$.
$\mathbf{w}_*$	decision vector indicating asset balances
$\overline{\mathbf{w}}_*$	market benchmark
$*$	Vectors $\mathbf{w}$, $\overline{\mathbf{w}}$, $\mathbf{b}$, and $\mathbf{s}$ can be indexed either with $t = 1, ..., T$ (in which case they represent stochastic quantities with an implied dependence on $\boldsymbol{\rho}^t$), or $\mathbf{e} \in \mathcal{N}$ (in which case they represent specific realisations of those quantities).
$\mathcal{W}_t$	expected total wealth at time $t = 1, \dots, T$.
$\mathbf{1}$	$\equiv (1, 1, 1, \dots, 1)'$
$\mathbf{u} \circ \mathbf{v}$	$\equiv (u_1 v_1,\ u_2 v_2,\ \dots, u_n v_n)'$

that $\boldsymbol{\rho}_t$ can take only finitely many values. Thus, the factors driving the risky events are approximated by a discrete set of scenarios or sequence of events. Given the event history up to a time t, $\boldsymbol{\rho}^t$, the uncertainty in the next period is characterised by finitely many possible outcomes for the next observation $\boldsymbol{\rho}_{t+1}$. This branching process is represented using a scenario tree. An example of scenario tree with 3 time periods and two-two-three branching structure is presented in Figure 1. We also model a continuous perturbation in addition to the discretised uncertainty (see also [9]), so that the vector of return values at time t has a multivariate normal distribution, with mean $\mathbf{r}_t(\boldsymbol{\rho}^t)$, and specified covariance matrix Λ.

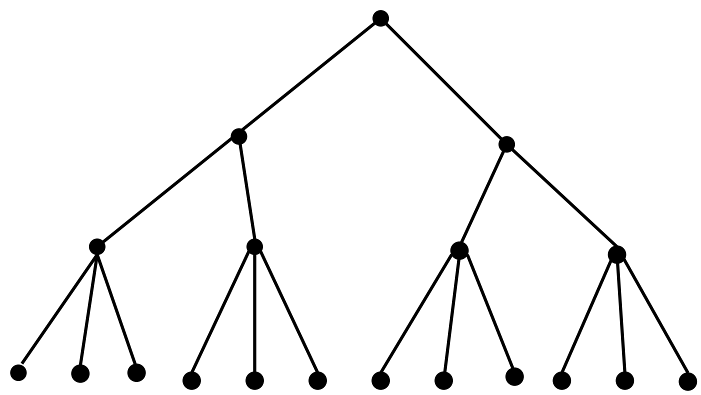


Figure 1: A scenario tree.

Capital Allocation

As the problem is dynamic in nature, the wealth to be invested varies with time. Hence, the traditional approach used is the static formulation, where wealth is normalised to unity, needs to be appropriately extended. At $t = 0$, the initial budget is normalised to 1. If the investor currently has holdings of assets $1, \ldots, n$, then vector $\mathbf{p}$ (scaled so that $\mathbf{1}'\mathbf{p} = 1$) represents his current position. If the investor currently has no holdings (wishing to buy in at time $t = 0$), then $\mathbf{p} = \mathbf{0}$. If the investor wishes to add to a current portfolio for the initial time period, then $\mathbf{p}$ can be scaled so that $\mathbf{1}'\mathbf{p}$ is the appropriate value smaller than 1. Then, the allocation of the initial budget of 1 can be represented with the following constraints:

$$\mathbf{p} + (1 - \mathbf{c}_b)\mathbf{b}_0 - (1 + \mathbf{c}_s)\mathbf{s}_0 = \mathbf{w}_0 \quad (1)$$

$$\mathbf{1}'\mathbf{b}_0 - \mathbf{1}'\mathbf{s}_0 = 1 - \mathbf{1}'\mathbf{p} \quad (2)$$

Note the following consequences of transaction costs $\mathbf{c}_b$ and $\mathbf{c}_s$ in (1) which generalize beyond the initial time period:

- If any assets are initially transacted so that $\mathbf{w}_0 \neq \mathbf{p}$ (i.e. $\mathbf{b}_0 + \mathbf{s}_0 \neq \mathbf{0}$) the total value of the resulting initial portfolio will be less than the initial budget of 1, due to loss through transaction costs.
- The fact that transactions have cost ensures that for the same asset ($\mathbf{b}_0^j \cdot \mathbf{s}_0^j = 0$); buy and sell variables corresponding to the same asset can never simultaneously be nonzero.
- The incorporation of transaction costs in the model provides essential "friction" [11]; without this friction, the optimization has complete freedom in reallocating the portfolio every time period, which (if implemented) can result in significantly poorer

realised performance than forecast, due to excessive transaction costs. Hakansson explains that in the absence of transaction costs, myopic policies are sufficient to achieve optimality in [14, 15].

Constraint (2) enforces the initial budget of unity (whether it be new investment or reallocation). The net total amount of buying makes up the shortfall of the original portfolio beneath the budget. It is important to note that our model allows money to be added to the portfolio in this way only in the initial time period.

Transaction Constraints

An investor must pay a commission to the broker when buying or selling a stock. That commission is the transaction cost of the purchase or sale. The impact of transaction costs on performance of the mean variance models have been investigated by Perold and Konno et al [22, 19, 4].

The decision at time $t > 0$ is clearly dependent on $\boldsymbol{\rho}^t$ and yields, after observing $\boldsymbol{\rho}^t$, the investments in asset j as

$$\begin{aligned} w_t^j &= r_t^j(\boldsymbol{\rho}^t)w_{t-1}^j + (1-c_b^j)b_t^j - (1+c_s^j)s_t^j, \qquad && j = 1,\dots,n;\ t = 1,\dots,T-1 \\ w_T^j &= r_T^j(\boldsymbol{\rho}^T)w_{T-1}^j, && j = 1,\dots,n. \end{aligned}$$

These last two constraints can be written more concisely in vector form as

$$\mathbf{w}_t = \mathbf{r}_t(\boldsymbol{\rho}^t) \circ \mathbf{w}_{t-1} + (\mathbf{1}-\mathbf{c}_b) \circ \mathbf{b}_t - (\mathbf{1}+\mathbf{c}_s) \circ \mathbf{s}_t, \qquad t = 1,\dots,T-1 \tag{3}$$

$$\mathbf{w}_T = \mathbf{r}_T(\boldsymbol{\rho}^T) \circ \mathbf{w}_{T-1}. \tag{4}$$

As $\mathbf{w}_t(\boldsymbol{\rho}^t)$ depends on the past and not the future, the policy $\{\mathbf{w}_0, \dots, \mathbf{w}_t\}$ is non-anticipative. Furthermore, $\boldsymbol{\rho}_0$ is observed before the initial decision and is thus treated as deterministic information.

Balance Constraints

Assuming that there are no exogenous cash injections, the amount of stock bought at any time needs to be financed by the amount of stock sold in the same period. We also observe that the amount invested at each period is given by the amount inherited from the previous period, less transaction costs. This is reflected by the following equality

$$\mathbf{1}'\mathbf{w}_t = \mathbf{1}'(\mathbf{r}_t \circ \mathbf{w}_{t-1}) - \mathbf{1}'(\mathbf{c}_b \circ \mathbf{b}_t) - \mathbf{1}'(\mathbf{c}_s \circ \mathbf{s}_t) \tag{5}$$

In order to (5) to be satisfied along with (3) it suffices that we require subsequent transactions (buy $= \mathbf{b}_t$, sell $= \mathbf{s}_t$) not to alter the wealth within the period t. Hence, we have the condition

$$\mathbf{1}'\mathbf{b}_t - \mathbf{1}'\mathbf{s}_t = 0, \qquad t = 1,\dots,T-1. \tag{6}$$

In order to model additions to or withdrawals from the portfolio, (6) could be modified as $\mathbf{1}'\mathbf{b}_t - \mathbf{1}'\mathbf{s}_t = h$.

Benchmarks

A benchmark is a measure of performance for a predetermined set of securities. It is designed to reflect prevailing market conditions. Such sets may be based on published indexes or may be customized to suit an investment strategy. One aspect of this paper is the use of mean-variance optimization in relation to a specified benchmark portfolio. Benchmark tracking may be very important for portfolio management. A risk averse investor may prefer to remain as close to the benchmark as possible for all return scenarios.

An initial benchmark portfolio is specified as $\overline{\mathbf{w}}_0$ (possibly $= \mathbf{p}$), and benchmarks at later time periods derive from $\overline{\mathbf{w}}_0$ by accruing returns, but not allowing any reallocation:

$$\overline{\mathbf{w}}_t(\boldsymbol{\rho}^t) = \mathbf{r}_t(\boldsymbol{\rho}^t) \circ \overline{\mathbf{w}}_{t-1}(\boldsymbol{\rho}^{t-1}), \qquad t = 1, \ldots, T. \tag{7}$$

Henceforth, $\overline{\mathbf{w}}_t$ will be referred to without its implicit dependence on $\boldsymbol{\rho}^t$.

Expected Wealth

The objective of an investor is to minimize portfolio risk while maximizing expected portfolio return on investment, or achieving a prescribed expected return. The expected wealth at time t, arising from period $[t-1, t]$ is

$$\mathrm{E}\,[\mathbf{r}_t(\boldsymbol{\rho}^t)'\,\mathbf{w}_{t-1}].$$

Note that, due to transaction costs in (3), this could be slightly less than $\mathrm{E}\,[\mathbf{1}'\mathbf{w}_t]$; this is intended, as we want to measure the expectation and variance of the wealth before the accounts are reallocated with $\mathbf{b}_t$ and $\mathbf{s}_t$. Given the possible events $\mathbf{e} \in \mathcal{N}_t$ (the discretisation of $\boldsymbol{\rho}_t$), the expected wealth at time t, relative to benchmark $\overline{\mathbf{w}}_t$, is given by

$$\begin{aligned} \mathcal{W}_t &= \mathrm{E}\left[\mathbf{1}'\left(\mathbf{w}_t - \overline{\mathbf{w}}_t\right)\right] \\ &= \mathrm{E}\left[\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1}\right)\right] \\ &= \mathrm{E}\left[\sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}}\left(\mathbf{r}'_{\mathbf{e}}\left(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}\right)\right)\right] \\ &= \sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}}\left(\hat{\mathbf{r}}'_{\mathbf{e}}\left(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}\right)\right), \end{aligned} \tag{8}$$

where $\hat{\mathbf{r}}_{\mathbf{e}}$ and $\mathbf{w}_{\mathbf{e}}$ ($\mathbf{e} \in \mathcal{N}_t$) are the various realisations of stochastic quantities $\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})$ and $\mathbf{w}_{t-1}(\boldsymbol{\rho}^{t-1})$.

Expected Risk

Risk, for any realisation of $\boldsymbol{\rho}^t$, is measured as the variance of the portfolio return relative to the benchmark $\overline{\mathbf{w}}$:

$$\begin{aligned} &\mathrm{Var}\left[\mathbf{r}_t(\boldsymbol{\rho}^t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1}\right)\right] \\ &= \mathrm{E}\left[\left(\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1}\right)\right)^2\right] - \left(\mathrm{E}\left[\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1}\right)\right]\right)^2 \\ &= \mathrm{E}\left[(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1})'(\Lambda + \mathbf{r}_t(\boldsymbol{\rho}^t|\boldsymbol{\rho}^{t-1})\mathbf{r}_t(\boldsymbol{\rho}^t|\boldsymbol{\rho}^{t-1})')(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1})\right] \\ &\quad - \left(\mathrm{E}\left[\mathbf{r}_t(\boldsymbol{\rho}^t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1} - \overline{\mathbf{w}}_{t-1}\right)\right]\right)^2 \end{aligned} \tag{9}$$

The two components $(\Lambda + \mathbf{r}_t\mathbf{r}'_t)$ in (9) reflect risk arising from two types of uncertainty. The term corresponding to the outer product $\mathbf{r}_t\mathbf{r}'_t$ represents the uncertainty due to the discrete distribution of realisations of $\boldsymbol{\rho}^t$ (various paths through the scenario tree, as discussed below). The term corresponding to Λ reflects continuous uncertainty due to the variability of the return at a given realisation of $\boldsymbol{\rho}$.

Once $\mathcal{W}_t$ is known, the variance of the wealth at time t (relative to the benchmark) can be similarly calculated:

$$\begin{aligned}
&\mathrm{Var}\left[\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1}-\overline{\mathbf{w}}_{t-1}\right)\right] \\
&\quad = \mathrm{E}\left[\left(\mathbf{r}_t(\boldsymbol{\rho}_t|\boldsymbol{\rho}^{t-1})'\left(\mathbf{w}_{t-1}-\overline{\mathbf{w}}_{t-1}\right)\right)^2\right] - (\mathcal{W}_t)^2 \\
&\quad = \mathrm{E}\left[\sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}}\left(\mathbf{r}_{\mathbf{e}}'\left(\mathbf{w}_{a(\mathbf{e})}-\overline{\mathbf{w}}_{a(\mathbf{e})}\right)\right)^2\right] - (\mathcal{W}_t)^2 \\
&\quad = \sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}}\Big((\mathbf{w}_{a(\mathbf{e})}-\overline{\mathbf{w}}_{a(\mathbf{e})})'(\Lambda+\hat{\mathbf{r}}_{\mathbf{e}}\hat{\mathbf{r}}_{\mathbf{e}}')(\mathbf{w}_{a(\mathbf{e})}-\overline{\mathbf{w}}_{a(\mathbf{e})})\Big) - (\mathcal{W}_t)^2. \qquad (10)
\end{aligned}$$

Bounds on Variables: No Short-sale

It is possible to sell an asset that you do not own through the process of short selling, or shorting the asset. In order to do this, the investor borrows the asset, then sells the borrowed asset. At a later date, the investor repays the loan by purchasing the asset and returning it to the lender. The short selling is profitable if the asset price declines. It is quite risky for many investors since potential for loss is unlimited. For this reason, short selling is prohibited within certain financial institutions, and it is purposely avoided as a policy by many individuals and institutions. However, it is not universally forbidden, and there is, in fact a considerable level of short selling of stock market securities. In the mean-variance optimization framework, bounds on decision variables prevent the short sale and enforce further restrictions imposed by the investor.

Let there be box constraints on $\mathbf{w}_{\mathbf{e}}$, $\mathbf{b}_{\mathbf{e}}$, $\mathbf{s}_{\mathbf{e}}$ for each event $\mathbf{e} \in \mathcal{N}$ such that

$$\mathbf{w}_{\mathbf{e}}^L \leq \mathbf{w}_{\mathbf{e}} \leq \mathbf{w}_{\mathbf{e}}^U$$

$$\mathbf{0} \leq \mathbf{b}_{\mathbf{e}} \leq \mathbf{b}_{\mathbf{e}}^U$$

$$\mathbf{0} \leq \mathbf{s}_{\mathbf{e}} \leq \mathbf{s}_{\mathbf{e}}^U$$

3 Multistage Stochastic Quadratic Programming Model

Let the expected returns $\hat{\mathbf{r}}_{\mathbf{e}}$, and covariance matrix $\Lambda_{\mathbf{e}}$ be given for $\mathbf{e} \in \mathcal{N}$. The multistage portfolio reallocation problem can be expressed as the minimization of risk subject to constraints given in (1)–(4) (which describe the growth of the portfolio along all the various scenarios), a performance constraint, and bounds on the variables.

The constant parameter $\mathcal{W}$ is supplied to constrain final expected wealth to a particular value (due to the minimization of risk, an equality constraint would be equivalent). The optimization will find the lowest-variance (least risky) investment strategy to achieve that specified expected wealth. Varying $\mathcal{W}$ and reoptimizing will generate points along the efficient frontier.

$$QP(\mathcal{W})$$

$$\min_{\mathbf{w},\mathbf{b},\mathbf{s}} \sum_{t=1}^{T} \alpha_t \sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}} \left[(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})})'(\Lambda_{\mathbf{e}} + \hat{\mathbf{r}}_{\mathbf{e}}\hat{\mathbf{r}}_{\mathbf{e}}')(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}) \right]$$

Subject to

$$\begin{array}{rcll}
\mathbf{p} + (1-\mathbf{c}_b)\mathbf{b}_0 - (1+\mathbf{c}_s)\mathbf{s}_0 & = & \mathbf{w}_0 & \\
\mathbf{1}'\mathbf{b}_0 - \mathbf{1}'\mathbf{s}_0 & = & 1 - \mathbf{1}'\mathbf{p} & \\
\hat{\mathbf{r}}_{\mathbf{e}} \circ \mathbf{w}_{a(\mathbf{e})} + (1-\mathbf{c}_b) \circ \mathbf{b}_{\mathbf{e}} - (1+\mathbf{c}_s) \circ \mathbf{s}_{\mathbf{e}} & = & \mathbf{w}_{\mathbf{e}} & \mathbf{e} \in \mathcal{N}_I \\
\mathbf{1}'\mathbf{b}_{\mathbf{e}} - \mathbf{1}'\mathbf{s}_{\mathbf{e}} & = & 0 & \mathbf{e} \in \mathcal{N}_I \\
\sum_{\mathbf{e}\in\mathcal{N}_T} P_{\mathbf{e}} \left(\hat{\mathbf{r}}_{\mathbf{e}}' (\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}) \right) & \geq & \mathcal{W} & \\
\mathbf{w}_{\mathbf{e}}^L \leq \mathbf{w}_{\mathbf{e}} \leq \mathbf{w}_{\mathbf{e}}^U & & & \mathbf{e} \in \mathcal{N} \\
\mathbf{0} \leq \mathbf{b}_{\mathbf{e}} \leq \mathbf{b}_{\mathbf{e}}^U & & & \mathbf{e} \in \mathcal{N}_I \cup \mathbf{0} \\
\mathbf{0} \leq \mathbf{s}_{\mathbf{e}} \leq \mathbf{s}_{\mathbf{e}}^U & & & \mathbf{e} \in \mathcal{N}_I \cup \mathbf{0}
\end{array}$$

Note that the classical stochastic linear programming formulation [7, 8] has the linear expression for $\mathrm{E}[\mathcal{W}]$ at the final time period as the objective function, and takes no account of risk. This formulation allows for calculation of efficient investment strategies which are not totally risk-seeking. There also exists stochastic linear programs which take account of risk such as utility models (which can be easily modelled as linear by piecewise approximation) and worst-case outcome, [29].

3.1 Computation of Transaction Costs

The transaction cost computation requires further clarification. It is implicit in the formulation that buy and sell variables for a particular asset in any scenario cannot be both nonzero — otherwise more money is spent on transaction costs than is necessary. In linear programming formulations using the simplex algorithm, this is inherently assured. The problem becomes especially acute nearer the risk averse computations when portfolio performance is less important than risk. However, in quadratic programming formulations it has been observed that common buy and sell variables often are found to be both significantly nonzero. This seems to arise from numerical instabilities due to the near linear dependence of the constraint columns of buy and sell variables. To solve this problem in our quadratic model above, we penalize the quantity $b^i_{\mathbf{e}} s^i_{\mathbf{e}}$ in every instance. This yields the following augmented objective function:

$$\min_{\mathbf{w},\mathbf{b},\mathbf{s}} \sum_{t=1}^{T} \sum_{\mathbf{e}\in\mathcal{N}_t} \left(\alpha_t P_{\mathbf{e}} \left[(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})})'(\Lambda_{\mathbf{e}} + \hat{\mathbf{r}}_{\mathbf{e}}\hat{\mathbf{r}}_{\mathbf{e}}')(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}) \right] + \gamma \mathbf{b}_{\mathbf{e}} \circ \mathbf{s}_{\mathbf{e}} \right)$$

Too large of a penalty $\gamma \geq 0$ can introduce additional numerical instability that severely hinders convergence, but a penalty coefficient of $\gamma = 0.001$ was found empirically to enforce complementarity of buy and sell variables without significant detriment to optimization performance.

3.2 Constraining Intermediate Risk/Return

In addition to constraining the final wealth, constraints of the following form

$$\sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}} \left(\hat{\mathbf{r}}_{\mathbf{e}}' (\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}) \right) \geq \mathcal{W}_t \tag{11}$$

for any $t = 1\ldots, T-1$ can be added to ensure any desired intermediate (expected) performance.

The objective function of $QP(\mathcal{W})$ is a weighted sum of (10) for each time period, with constant terms $\mathcal{W}_t^2$ dropped. If $\alpha_t \equiv 1$, then risk in each time period is weighted equally, but allowing α_t to decrease with t (e.g. $\alpha_t = \gamma^t$, for some $\gamma \in (0,1)$) reflects a discount for risk further in the future. As another alternative, instead using an objective function which minimizes risk for all t, it may be preferable to have bounds on risk in the intermediate periods and just minimize final risk at $t = T$. This is reflected in the objective function by setting

$$\alpha_t = \begin{cases} 0 & t < T \\ 1 & t = T \end{cases} \tag{12}$$

and maximum intermediate risks would be enforced (in conjunction with (11)) by additional constraints of the form

$$\sum_{\mathbf{e}\in\mathcal{N}_t} P_{\mathbf{e}} \left[(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})})' (\Lambda_{\mathbf{e}} + \hat{\mathbf{r}}_{\mathbf{e}}\hat{\mathbf{r}}_{\mathbf{e}}')(\mathbf{w}_{a(\mathbf{e})} - \overline{\mathbf{w}}_{a(\mathbf{e})}) \right] - (\mathcal{W}_t)^2 \le \lambda_t \tag{13}$$

(with $\lambda_t \ge 0$) for any $t = 1, \ldots, T-1$. The intermediate performance (11) and intermediate risk (13)) constraint types may be applied at specific junctures and need not be present for all t.

3.3 Implementation Issues: Saving Space

In implementation, the model can be simplified to reduce the number of variables. First, all variables $\mathbf{w}_T$ may be substituted out using (4) (as they are in the performance constraint above). As the majority of nodes in any branching tree are nodes, this will result in significant space savings (which any good optimization preprocessor should do as well).

It could be unreasonable or impossible to specify bounds for the decision variables *a priori* as the size of the portfolio at any particular scenario node depends on values of variables in time periods, as well as on the quality of asset returns along the scenario path particular to that node. It is possible to express variable bounds in a percentage, rather than absolute sense; for instance, percentage bounds such as

$$\beta_{\mathbf{e}}^{iL} \le \frac{w_{\mathbf{e}}^i}{w_{\mathbf{e}}^1 + w_{\mathbf{e}}^2 + \ldots + w_{\mathbf{e}}^n} \le \beta_{\mathbf{e}}^{iU}$$

can easily be linearised, as

$$(1-\beta_{\mathbf{e}}^{iL})w_{\mathbf{e}}^i - \sum_{j\ne i} \beta_{\mathbf{e}}^{iL} w_{\mathbf{e}}^j \ge 0$$

$$(1-\beta_{\mathbf{e}}^{iU})w_{\mathbf{e}}^i - \sum_{j\ne i} \beta_{\mathbf{e}}^{iU} w_{\mathbf{e}}^j \le 0.$$

These constraints, unfortunately, are not as simple as the original box constraints (which are typically handled as a special case by any optimization algorithm), and could add significantly to the size of the problem, and its solution time. Percentage lower and upper bounds on $\mathbf{w}$, $\mathbf{b}$, and $\mathbf{s}$ variables will require $4n^2 + 2n$ non-zeroes per node of the problem (compared to $6n$ coefficients necessary for transaction and balance constraints).

A significant savings in space can be obtained with the introduction of an auxiliary total variable

$$T_{\mathbf{e}} = \mathbf{1}'\mathbf{w}_{\mathbf{e}},$$

allowing percentage constraints to be simplified to

$$\mathbf{w}_{\mathbf{e}}^i \ge \beta_{\mathbf{e}}^{iL} T$$

$$\mathbf{w}_{\mathbf{e}}^{i} \leq \beta_{\mathbf{e}}^{iU} T,$$

for a total of $9n + 1$ non-zeroes spent on percentage bounds. This compares favourably with the previous $4n^2 + 2n$ even with only $n = 2$ assets, and for higher n the move from a quadratic to a linear number of coefficients per node can reduce the problem to a fraction of its original size. In our implementation (using interior point quadratic optimizer BPMPD [21]), this reduction in space did not correspond to a reduction in solution time. This is presumably because the reduction in size of inequality constraints was offset by the complexity difference of additional equality constraints.

4 Performance

4.1 Implementation

The model $QP(\mathcal{W})$ was implemented as a software package called *foliage*. *Foliage*, coded in C++, is a financial software developed in the Department of Computing by the authors. It integrates with LP/QP solver BPMPD [21] which is based on special interior point algorithm developed for solving QP problems. An alternative solution strategy based on the interior point algorithm has been developed by Blomvall and Lindberg [16], where a Riccati based primal barrier method is specially fitted to solve large scale multistage stochastic programming with a quadratic objective.

Foliage computes the minimum-variance investment strategy to achieve that expected wealth by solving $QP(\mathcal{W})$ if a value of $\mathcal{W}$ is provided. Otherwise, *foliage* proceeds in the following manner:

- Accumulate the benchmark initial investment allocation through each scenario (from root to leaf), and calculate the expected final wealth generated by the non-active benchmark investment strategy. Call this amount $\mathcal{W}_{\min}$.
- Remove the final wealth constraint from $QP(\mathcal{W})$, and use its linear left-hand side to replace the quadratic objective function. Maximise the resulting linear program to obtain $\mathcal{W}_{\max}$.
- For a number of equally-spaced points $\mathcal{W} \in [\mathcal{W}_{\min}, \mathcal{W}_{\max}]$, solve $QP(\mathcal{W})$, and output the resulting points of the efficient frontier.

Foliage has the ability to handle simple box constraints on the decision variables, as well as percentage constraints, as discussed above. The discounts can either be of the form $\alpha_t = \gamma^t$, or (12), which would ignore variance in intermediate time periods and optimize only the variance of the final wealth. The scenario trees input to the program can have arbitrary branching structure and depth (limited only by computer memory).

All of the problems were solved on a 500 MHz Pentium III, running Linux with 256Meg of RAM.

4.2 WATSON Data

The WATSON family of problems [8] forecast to a horizon of 10 years a portfolio of four asset classes, as well as number of liabilities, and riskless assets (bank deposits or borrowing). There are 24 data sets, ranging in size from 16 to 2688 scenarios. The data sets were generated in two ways, either with Conditional Scenarios (C), or Independent Scenarios (I). The CALM model [7] underlying the WATSON data sets incorporates more data than was needed by our model (i.e. liabilities in addition to assets, separate returns for the same asset bought at different times, etc.). We kept the same scenario tree structure, using

Table 2: Performance of *foliage* on WATSON data sets.

scenarios	branching	nonzeroes	secs.	(avg)	its.	(avg)
16	$(2^4 \cdot 1^6)$	10634	2.42	(1.2)	16	(6.6)
32	$(2^5 \cdot 1^5)$	18378	5.07	(2.4)	19	(7.2)
64	$(2^6 \cdot 1^4)$	30794	9.14	(4.5)	20	(7.9)
128	$(2^7 \cdot 1^3)$	49482	15.7	(8.1)	21	(8.9)
256	$(2^8 \cdot 1^2)$	74570	31.6	(12.9)	28	(9.0)
512	$(2^9 \cdot 1)$	100170	46.8	(20.2)	29	(10.1)
768	$(3 \cdot 2^8 \cdot 1)$	150250	70.2	(34.8)	28	(11.4)
1024	$(4 \cdot 2^8 \cdot 1)$	200330	124.5	(48.7)	38	(11.8)
1152	$(3^2 \cdot 2^7 \cdot 1)$	225226	112.9	(57.0)	30	(12.4)
1536	$(4 \cdot 3 \cdot 2^7 \cdot 1)$	300298	195.0	(85.2)	39	(13.7)
1920	$(5 \cdot 3 \cdot 2^7 \cdot 1)$	375370	239.4	(113.5)	38	(14.7)
2304	$(6 \cdot 3 \cdot 2^7 \cdot 1)$	450442	273.7	(144.5)	35	(15.2)
2688	$(7 \cdot 3 \cdot 2^7 \cdot 1)$	525514	335.1	(178.1)	37	(16.0)

return prices between times $t-1$ and t as our $\hat{\mathbf{r}}$, incorporated dividends into the returns, and discarded the remainder of the data.

For bounds, we specified that at all times each of the 4 assets must represent between 10–50% of the total portfolio, and that individual transactions must never be more than 50% of the portfolio (also in percentage). Use of the percentage bounds greatly increases the size of the problems (in these cases, typically threefold), due to the addition of $O(n^2)$ non-zeroes at each scenario node.

In Table 2, time and BPMPD iterations are reported for the solution of each problem, with $\mathcal{W} = 2.0$ (a value typically in the middle of the efficient frontier). In parentheses are the average time and iterations required to solve 20 equally-spaced points of the efficient frontier; the warm-start capabilities of BPMPD allow for significant time savings. Only the problems generated with Conditional Scenarios are provided, since the Independent Scenario problems have identical sizes and the performance of *foliage* was virtually identical.

4.3 Backtesting

In order to demonstrate the viability of the investment decisions produced by *foliage*, we did backtesting using historical stock price data, in the following manner:

Initialise the following parameters:

- Past horizon p
- Future horizon f
- Branching b
- Risk level r
- number of Simulations S

Set current time to $t = p$, and initial portfolio value to $V = 1$.

Analyze the past p time periods (from $t-p+1$ to t). Use logarithmic regression to estimate the growth rate of each stock. Calculate the covariance matrix of the residuals from this regression.

Forecast a scenario tree f time periods into the future, branching b at each time period. The root of the scenario tree incorporates the present data. To determine the branching of any scenario node to its children, simulate a large number S of single time period scenarios using the state of the present scenario and the growth and covariance data; then cluster those simulations into b groups, the "centers" of which will be the child scenario nodes. For more details on the forecasting of scenario trees through simulation, see [13].

Optimise investment decisions by running *foliage* on the generated scenario tree to find the efficient investment strategy for $\mathcal{W} = r\mathcal{W}_{\max} + (1-r)\mathcal{W}_{\min}$.

Invest portfolio value V among the various assets according to the initial allocation prescribed in the optimal solution.

Accrue balances in each asset according to the change in price from current time t to "tomorrow", the stock prices at $t+1$. Update V, the total value of the portfolio, and increment t.

Iterate back to step **Analyze** as long as there is more historical data.

Results are depicted in Figs. 2 and 3. The data arises from 10 British stocks (arbitrarily chosen from the FTSE 100), priced monthly from 1988 to early 2000. The "index" in all Figures charts the growth of an initially equally-weighted portfolio without rebalancing. Each of the other curves represent a fixed value of risk throughout the time horizon. It is interesting to note that the crash in month 120 reflects the crash of the FTSE in late 1998.

It should also be noted that, as the scenario trees involved in this experiment were generated using randomised procedures, repetition of the experiments can (and indeed do) yield different results. However, this effect was minimized by the use of low-discrepancy quasirandom (but deterministic) Sobol sequences ([27, 3], see [13] for comparison of quasirandom vs. pseudorandom numbers in scenario tree generation).

In Figs. 2 and 3 we can see the effect of transaction costs on the model, as well as the importance of an appropriate choice of benchmark. For Fig. 2, every scenario tree generated in the backtesting experiment used a constant benchmark of 1/10 per asset. Beyond being a natural arbitrary choice, we can justify this by noting that the index is based on an initially equally-weighted portfolio. The two graphs display the subtle effect of including transaction costs in the optimization model. In both backtesting experiments, a transaction fee of 0.2% was deducted for all purchases; however, only in the latter experiment was that transaction cost also included in the data input to optimizer *foliage*. The effect is a subtle improvement to the higher risk curves (75% and 99%), and subtle worsening of the lower risk curves (25% and 50%). This is because higher-risk strategies involve more active trading, so the cost of trading cannot be ignored. Lower-risk strategies, however, are already trading less, and inclusion of transaction costs in the optimization model further penalised trading, and prevented or delayed some necessary rebalancing.

Fig. 3, considers an alternative benchmark selection and alternative transaction cost value. The benchmark supplied as data to each optimization is the same as the current portfolio $\mathbf{p}$ prior to optimization. This means that the zero-risk (as measured by the model's quadratic objective function) strategy is to maintain the current portfolio. This ensures that a portfolio is never changed radically (unless high performance demands of risk-seeking strategies demand it), but allows the portfolio to migrate to the naturally lowest-risk strategy, instead of always being tied to an equal-weight benchmark (which might not be a particularly low-risk or high-performance strategy). The alternative benchmark is also motivated by an inertia-based approach to investment. The results indicate the clear superiority of taking transaction costs into account during optimization.

5 Conclusions

In this paper, we consider the optimization of the mean and variance of portfolio returns for a multiperiod investment problem with transaction costs. The basic investment model is adopted the benchmark relative computations in view of the transaction costs. A detailed study of transaction costs as well as computational approaches for addressing the issues arising from them is presented.

The paper is based on the observation that, in view of the transaction costs, it is essential to consider the multistage decision problem directly. This overcomes the suboptimality of myopic sequential application of single period optimization and provides a realistic cost conscious strategy.

We provide a general benchmark-relative return framework, and use an interior point optimizer to solve the overall quadratic programming problem. We generate scenario trees using low-discrepancy quasirandom Sobol sequences based on historical information to test the *ex ante* performance of the methodology. We also test the methodology using publicly available large-scale problems based on the CALM model. We report a number of computational backtesting experiments which illustrate the performance of the transaction cost model.

Acknowledgements

This research was supported by EPSRC grant GR/M41124. The authors are grateful to anonymous referees for helpful comments and suggestions.

References

[1] R. Becker, S. Hall and B. Rustem, Robust Optimal Control of Stochastic Nonlinear Economic Systems, *J. Economic Dynamics and Control*, **18** (1994) 125–148.

[2] J.R. Birge, F. Louveaux, *Introduction to Stochastic Programming* Springer-Verlag New York, (1997).

[3] P. Bratley, B.L. Fox, and H. Niederreiter, Implementation and Tests of Low-Discrepancy Sequences, *ACM Trans. on Modeling and Comp. Simulation*, **2**, 3 (1992), 195–213.

[4] V.K, Chopra, C.R. Hensel and A.L. Turner, Masaging Mean-Variance Inputs: Returns from Alternative Global Investment Strategies in the 1980s, *Management Science*, 39, 7, (1993), 845–855.

[5] J. Darlington, C.C. Pantelides, B. Rustem and B.A. Tanyi, An Algorithm for Constrained Nonlinear Optimisation Under Uncertainty, *Automatica*, 35, (1999), 217–228.

[6] J. Darlington, C.C. Pantelides, B. Rustem and B.A. Tanyi, Decreasing the Sensitivity of Open-loop Optimal Solutions in Decision Making Under Uncertainty, *European J. Operations Research*, 121, (2000) 330–342.

[7] M.A.H. Dempster, CALM: A Stochastic MIP model, Department of Mathematics, University of Essex, Colchester, UK (1993).

[8] Finance Research Group, Judge Institute of Management Studies, Cambridge, Watson Pension Fund Management Problem, available via the WWW at URL: `http://www-cfr.jims.cam.ac.uk/research/stprog.html`.

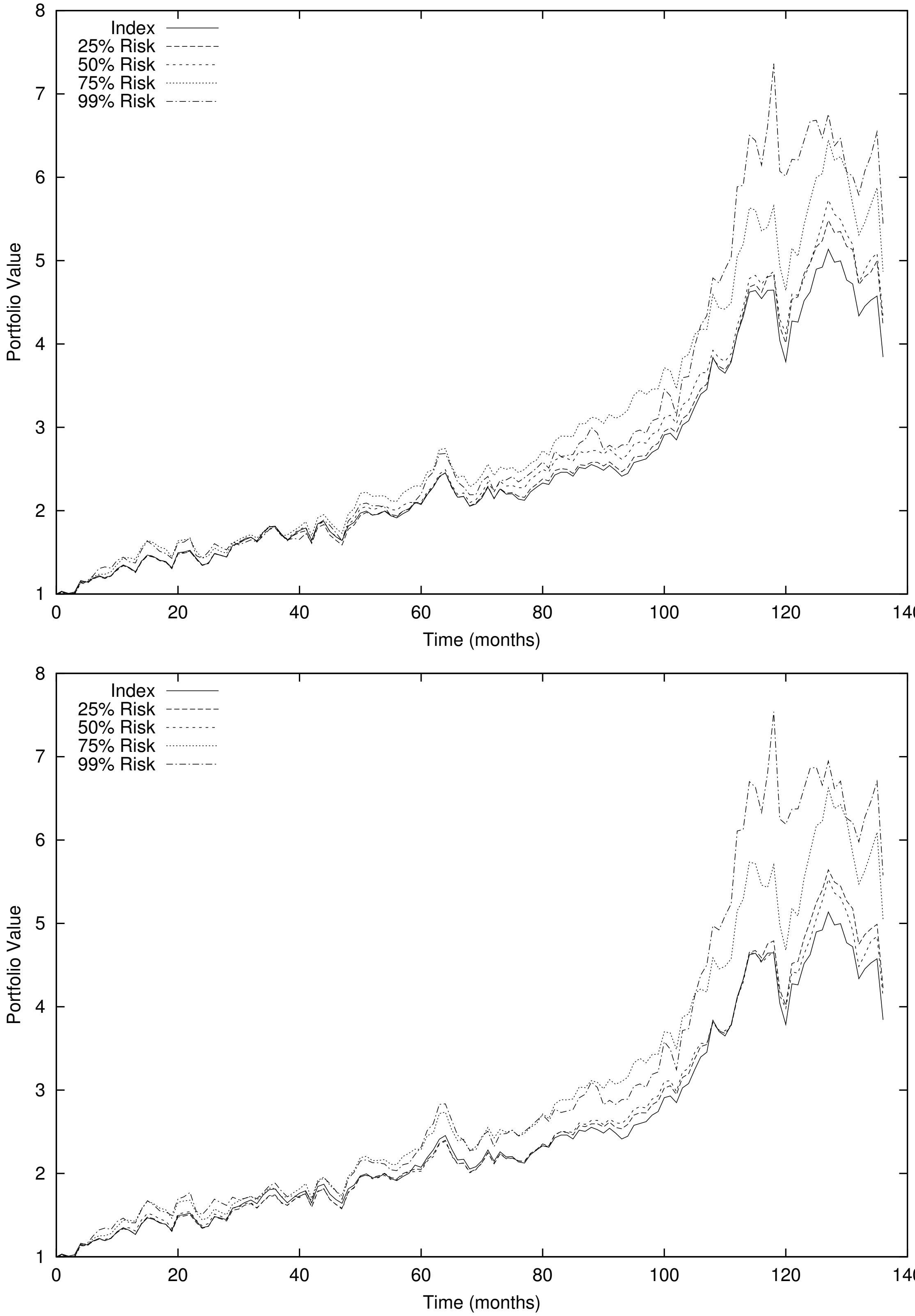
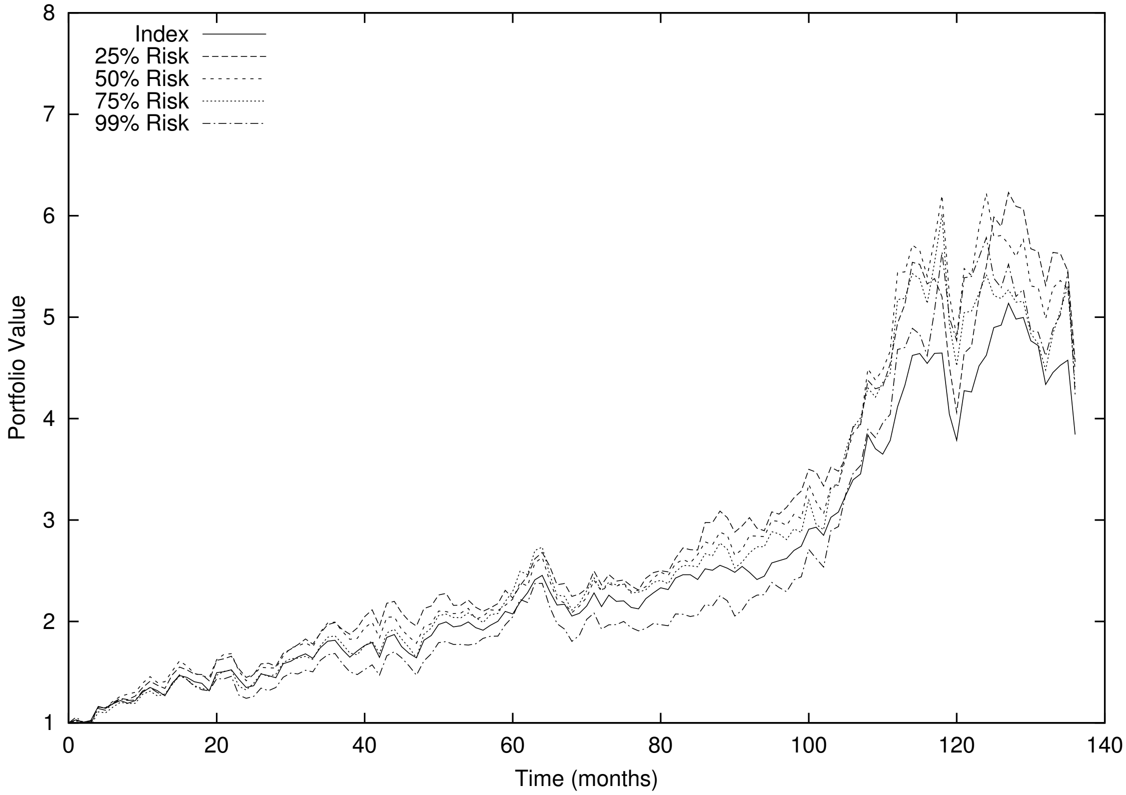


Figure 2: Backtesting results with a fixed, equal-weight benchmark, and 0.2% transaction cost for purchases. In the graph above, the optimizer had no knowledge of the transaction costs.



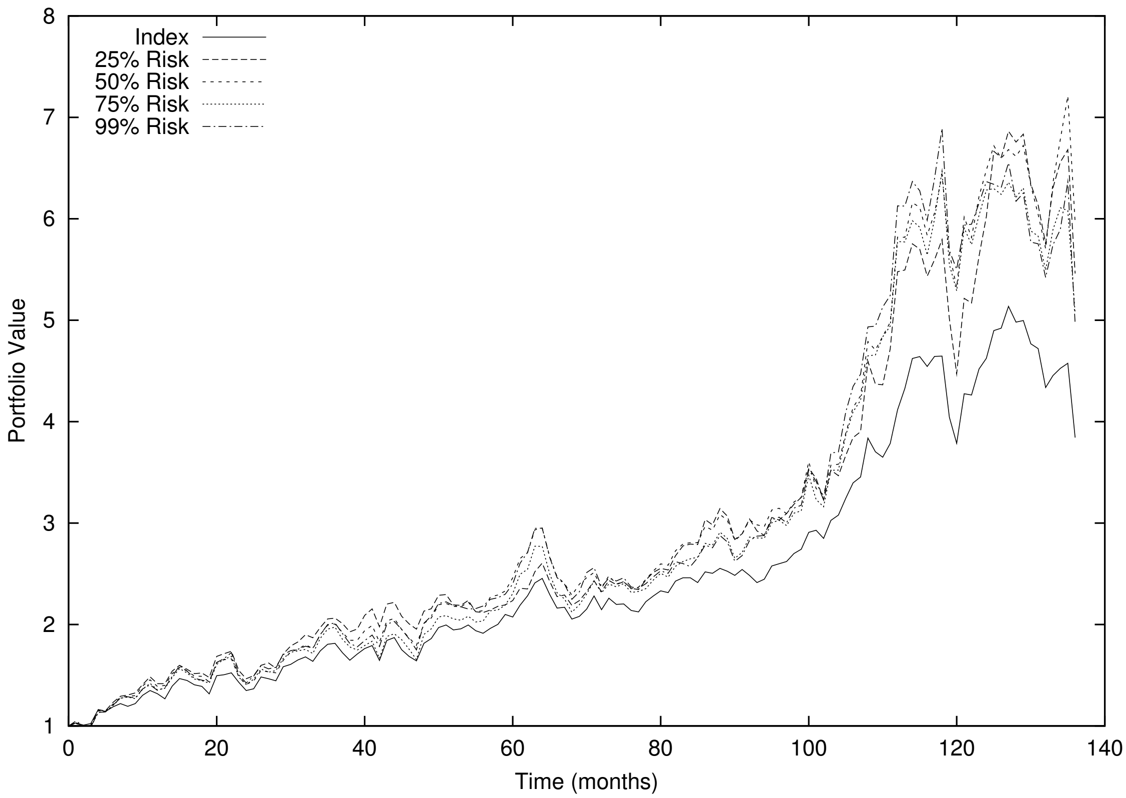


Figure 3: Backtesting results with a dynamic (current portfolio) benchmark, and 0.5% transaction cost for purchases. In the graph above, the optimizer had no knowledge of the transaction costs.

[9] K. Frauendorfer, The stochastic Programming Extension of the Markowitz Approach, *J. Mass-Parallel Comp. Inform. Sys.*, 5 (1995), 449–460.

[10] K. Frauendorfer, H. Siede, Portfolio Selection Using Multistage Stochastic Programming, *Central European Journal of Operations Research*, 7, (4), (2000), 277–289.

[11] Fragniere, Emmanuel, U. Lausanne, Private Communicationi, (2001).

[12] H. Gassman, A.M. Ireland, Scenario Formulation in an Algebraic Modelling Language, *Annals of Operations Research*, 59, (1995), 45–75.

[13] N. Gulpinar, B. Rustem, R. Settergren, Optimisation and Simulation Approaches to Scenario Tree Generation, Forthcoming (2004), *Journal of Economic Dynamics and Control*.

[14] N.H. Hakansson, On Optimal Myopic Portfolio Policies, with and without Serial Correlations of Yields, *Journal of Business*, 44, (1971), 324–334.

[15] N.H. Hakansson, Convergence to Isoelastic Utility and Policy in Multiperiod Portfolio Choice, *Journal of Financial Economics*, 1, (1974), 201–224.

[16] J. Blomvall, P.O. Lindberg, A Riccati-Based Primal Interior Point Solver for Multistage Stochastic Programming-Extensions, *Optimization Methods and Software*, 17 (3), (2002), 383–407.

[17] P. Kall, W. Wallace, Stochastic Programming, Wiley, (1994), Chichester.

[18] J.G. Kallberg and W.T. Ziemba, Comparison of Alternative Utility Functions in Portfolio Selection Problems, *Management Science*, 29, (11), (1983), 1257–1276.

[19] H. Konno and H. Yamazaki, Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokya Stock Market, *Management Science*, 37, 5, (1991), 519–531

[20] H. Markowitz, Portfolio Selection, *Journal of Finance*, 7, (1952), 77–91.

[21] C. Mészáros, BPMPD User's Manual Version 2.20, Department of Computing, Imperial College, *Research Report*, July 1997.

[22] A.F. Perold, Large Scale Portfolio Optimization, *Management Science*, 30, 10, (1984), 1143–1160

[23] L.B. Pulley, A General Mean-Variance Approximation to Expected Utility for Short Holding Periods, *Journal of Financial and Quantitative Analysis*, 16, (1981), 361–373.

[24] A. Prekopa, Stochastic Programming, Kluwer Academic Publisher, (1995), Dordrecht.

[25] B. Rustem, Stochastic and Robust Control of Nonlinear Economic Systems, *European Journal of Operations Research*, **73** (1994) 304–318.

[26] B. Rustem, R. Becker and W. Marty, Robust Min-max Portfolio Strategies for Rival Forecast and Risk Scenarios, *Journal of Economic Dynamics and Control*, 24, (2000), 1591–1623.

[27] I.M. Sobol, The Distribution of Points in a Cube and the Approximate Evaluation of Integrals, *USSR Comput. Math. Phys.* **7**, 4 (1967), 86–112.

[28] M. Steinbach, Recursive Direct Optimisation and Successive Refinement in Multistage Stochastic Programs, Preprint SC 98-27, Konrad-Zuse Centrum für Informationstechnik Berlin (1998).

[29] M.R. Young, A Minimax Portfolio Selection Rule with Linear Programming Solution, *Management Science*, 44, (5), 1998, 673–683

[30] W.T. Ziemba, L.C. MacLean, Growth versus Security Tradeoffs in Dynamic Investment Analysis, *Annals of Operations Research*, 85, (1999), 193–225.