Robust Modeling of Multi-Stage Portfolio Problems

Article · January 2000

DOI: 10.1007/978-1-4757-3216-0_12

Aharon Ben-Tal
Technion - Israel Institute of Technology

Arkadi Nemirovski
Georgia Institute of Technology

1 ROBUST MODELING OF MULTI-STAGE PORTFOLIO PROBLEMS

Aharon Ben-Tal, Tamar Margalit, and Arkadi Nemirovski

Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology
Technion City, Haifa 32000, Israel*

morbt@ie.technion.ac.il tammar@tx.technion.ac.il nemirovs@ie.technion.ac.il

Abstract: In the paper, we develop, discuss and illustrate by simulated numerical results a new model of multi-stage asset allocation problem. The model is given by a new methodology for optimization under uncertainty – the Robust Counterpart approach.

INTRODUCTION

The goal of this paper is to apply a novel modeling methodology aimed at treating data uncertainty in optimization problems – the *Robust Counterpart* approach [1, 2] – to the *multi-period asset allocation problem.* We start with the original model of Dantzig and Infanger [3] (Section 1) and apply to this model the Robust Counterpart approach [1, 2], thus coming to a new model of the problem (Section 1). The advantages/disadvantages of the resulting model as compared to the standard *Multistage Stochastic Programming* model of the Portfolio problem are discussed in Section 1. The concluding Section 1 presents some simulated numerical results.

The Original Problem

The *multi-period asset allocation problem* ("Portfolio problem" for short) as stated by Dantzig and Infanger [3] is as follows.

*Partially funded by the Fund for the Promotion of Research at Technion, the G.I.F. contract No. I-0455-214.06/95 and the Israel Ministry of Science grant # 9636-1-96

There are n types of assets $i = 1, ..., n$ plus cash ("asset $n+1$") and L time investment periods, and the problem is to control a portfolio of these assets. Let x_i^l be the amount ("dollar value") of asset i in the portfolio at the beginning of the investment period l, $l = 1, ..., L$.

The dynamics of the quantities x_i^l is given by the equations

A. $i \le n$ ("non-cash assets"):

$$x_i^l = r_i^{l-1} x_i^{l-1} - y_i^l + z_i^l, \tag{1.1}$$

where

• $r_i^{l-1} x_i^{l-1}$ is the amount coming from the preceding period (the coefficient $r_i^{l-1} > 0$ is the *asset return*)

• z_i^l is the amount of the asset we buy at the beginning of the period l

• y_i^l is the amount of the asset we sell at the beginning of the period l

B. $i = n+1$ ("cash"):

$$x_{n+1}^l = r_{n+1}^{l-1} x_{n+1}^{l-1} + \sum_{i=1}^{n} (1 - \mu_i^l) y_i^l - \sum_{i=1}^{n} (1 + \nu_i^l) z_i^l \tag{1.2}$$

where

• $r_{n+1}^{l-1} x_{n-1}^{l-1}$ is the cash coming from the previous period ($r_{n+1}^{l-1} > 0$ is the cash return)

• $(1 - \mu_i^l) y_i^l$ is the cash we get from selling amount y_i^l of the asset i at the beginning of the period l. Recall that the assets are measured by their "dollar value", so that in the case of costless transactions selling amount y_i^l of asset i we would get cash y_i^l; in fact the transactions are not costless, and the *transaction cost* $\mu_i^l \ge 0$ is the percent we pay for the transaction

• $(1 + \nu_i^l) z_i^l$ is the cash we pay to buy amount z_i^l of asset i ($\nu_i^l \ge 0$ is the corresponding transaction cost)

When making decision at instant l, we know all x_i^{l-1}, r_i^{l-1}, $i = 1, ..., n+1$. The decision is comprised of the quantities y_i^l, z_i^l, $i = 1, ..., n$ which should satisfy the restrictions

$$\begin{array}{rcccl} \underline{y}_i^l & \le & y_i^l & \le & \overline{y}_i^l, \ i = 1, ..., n, \\ \underline{z}_i^l & \le & z_i^l & \le & \overline{z}_i^l, \ i = 1, ..., n, \\ \underline{x}_i^l & \le & x_i^l & \le & \overline{x}_i^l, \ i = 1, ..., n+1, \end{array} \tag{1.3}$$

where $\underline{y}^l, \overline{y}^l, \underline{z}^l, \overline{z}^l, \underline{x}^l, \overline{x}^l$ are given vectors of bounds and x_i^l are defined according to (1.1) - (1.2).

From now on let us focus on the case of *simple bounds* – the lower ones are zero, the upper ones are $+\infty$.

The goal is to maximize the final total value of the assets

$$v = \sum_{i=1}^{n+1} r_i^L x_i^L \tag{1.4}$$

The case of complete information

Assume for a moment that the asset returns and transaction costs are known in advance. Then the situation in question can be modeled by the following Linear Programming problem:

$$
\begin{array}{lrcl}
(P_*) & & & \\
& & \max & \sum_{i=1}^{n+1} r_i^L x_i^L \\
\text{s.t.} & & & \\
& x_i^l & = & r_i^{l-1} x_i^{l-1} - y_i^l + z_i^l,\ i = 1,...,n, l = 1,...,L; \\
& x_{n+1}^l & = & r_{n+1}^{l-1} x_{n+1}^{l-1} + \sum_{i=1}^{n} (1-\mu_i^l) y_i^l - \sum_{i=1}^{n} (1+\nu_i^l) z_i^l \\
& y_i^l & \geq & 0,\ i = 1,...,n, l = 1,...,L; \\
& z_i^l & \geq & 0,\ i = 1,...,n, l = 1,...,L; \\
& x_i^l & \geq & 0,\ i = 1,...,n+1, l = 1,...,L.
\end{array}
\qquad (1.5)
$$

In this problem,

- x^l, y^l, z^l, $l = 1,...,L$, are decision vectors
- r_i^l, $l = 0,...,L$, same as μ_i^l, ν_i^l, $l = 1,...,L$, are data coefficients
- x^0 is a given initial state of the portfolio.

For our future purposes it makes sense to pass from the LP program (P_*) to an equivalent one, namely, to pass from the original design variables x_i^l, y_i^l, z_i^l to the new ones

$$
\begin{array}{rcl}
\xi_i^l & = & (R_i^l)^{-1} x_i^l, \\
\eta_i^l & = & (R_i^l)^{-1} y_i^l, \\
\zeta_i^l & = & (R_i^l)^{-1} z_i^l,
\end{array}
$$

where

$$R_i^l = r_i^0 r_i^1 ... r_i^{l-1}. \qquad (1.6)$$

In the new variables (P_*) becomes the program

$$
\begin{array}{lrcl}
(P^*) & & & \\
& & \max & \sum_{i=1}^{n+1} R_i^{L+1} \xi_i^L \\
\text{s.t.} & & & \\
& \xi_i^l & = & \xi_i^{l-1} - \eta_i^l + \zeta_i^l,\ i = 1,...,n, l = 1,...,L; \\
& \xi_{n+1}^l & = & \xi_{n+1}^{l-1} + \sum_{i=1}^{n} A_i^l \eta_i^l - \sum_{i=1}^{n} B_i^l \zeta_i^l,\ l = 1,...,L; \\
& \eta_i^l & \geq & 0,\ i = 1,...,n, l = 1,...,L; \\
& \zeta_i^l & \geq & 0,\ i = 1,...,n, l = 1,...,L; \\
& \xi_i^l & \geq & 0,\ i = 1,...,n+1, l = 1,...,L,
\end{array}
\qquad (1.7)
$$

where

$$
\begin{array}{rcl}
A_i^l & = & (1-\mu_i^l) R_i^l / R_{n+1}^l, \\
B_i^l & = & (1+\nu_i^l) R_i^l / R_{n+1}^l.
\end{array}
\qquad (1.8)
$$

Now let us switch from (P^*) to its inequality constrained version, where the "cash flow equations" are replaced with "cash flow inequalities":

$$
\begin{array}{llcl}
(P) & & & \\
 & & \max & w \\
\text{s.t.} & & & \\
(I_{L+1}): & w & \leq & \sum\limits_{i=1}^{n+1} R_i^{L+1}\xi_i^L \\
 & \xi_i^l & = & \xi_i^{l-1} - \eta_i^l + \zeta_i^l,\ i=1,...,n, l=1,...,L \\
(I_l): & \xi_{n+1}^l & \leq & \xi_{n+1}^{l-1} + \sum\limits_{i=1}^{n} A_i^l\eta_i^l - \sum\limits_{i=1}^{n} B_i^l\zeta_i^l,\ l=1,...,L \\
 & \eta_i^l & \geq & 0,\ i=1,...,n, l=1,...,L \\
 & \zeta_i^l & \geq & 0,\ i=1,...,n, l=1,...,L \\
 & \xi_i^l & \geq & 0,\ i=1,...,n+1, l=1,...,L
\end{array}
\tag{1.9}
$$

Of course, all three problems (P_*), (P^*) and (P) are equivalent to each other.

Data uncertainty

The above models relate to the case of complete a priori information on the future values of the asset returns and transaction costs; in reality, of course, these quantities are not known exactly. The natural assumption is that

> **A.** the data which is known exactly at time instant l (i.e., at the beginning of time period l) are the past values of the asset returns $r_i^0, r_i^1, ..., r_i^{l-1}$, $i=1,...,n+1$, as well as the past and present values of the transaction costs $\mu_i^1, \mu_i^2, ..., \mu_i^l, \nu_i^1, ..., \nu_i^l$, $i=1,...,n$.

Let us denote the collection of data known at time instant l by ω^l, and let ω_l^L be the collection

$$\{r_i^t \mid 1 \leq i \leq n+1, l \leq t \leq L\} \cup \{\mu_l^t, \nu_l^t \mid 1 \leq i \leq n, l < t \leq L\}$$

of the future, w.r.t. time instant l, values of the asset returns and the transaction costs. At time instant l, the data ω_l^L are uncertain; following Dantzig and Infanger [3], we assume that ω_l^L is random, with known at time instant l probability distribution.

An immediate consequence of the fact that the data in (P) are uncertain is that the Linear Programming problem (P) is *not* anymore a valid model of the problem in question. The standard way to pass from (P) to a "usable" model is given by the Multistage Stochastic Programming approach. According to it, we should treat the decision variables $\xi_i^l, \eta_i^l, \zeta_i^l$ not as reals, but as measurable functions of ω^l – of the data known at the instant when the corresponding decisions should be implemented. These functions should almost surely satisfy the constraints of (P) (which now become inequalities between random variables), and the objective to be maximized is, e.g., the expectation, over ω^{L+1}, of the value of the original objective (for more details on the Multistage Stochastic Programming model of the Portfolio problem, see [3]).

If we believe that the uncertainty in the portfolio problem is of stochastic nature, then the Multistage Stochastic Programming model of the problem is completely adequate. This model, however, has a severe intrinsic drawback: it is "computationally intractable", provided that the number L of stages is > 2. Indeed, in the typical case of continuously distributed random data it is computationally intractable even to write down a candidate solution – a collection of functions of continuous argument without any analytical structure clear in advance. Multistage Stochastic Programming offers a number of techniques aimed at overcoming this intrinsic drawback of the approach – approximating continuous distributions of the data by discrete ones ("scenario approach"), importance sampling, etc., but even with all these techniques the Stochastic Programming approach in its computational aspects seems to be an "ad hoc skill" rather than a "ready-to-use" technique.

ROBUST COUNTERPART APPROACH TO THE PORTFOLIO PROBLEM

The goal of this paper is to develop an alternative model of the uncertain Portfolio problem, a model based on the methodology of *Robust Mathematical Programming* as developed in [1, 2]. In contrast to the Multistage Stochastic Programming approach, our primary goal is to end up with computationally efficiently tractable model, and to this end we are ready to be a bit conservative just from the beginning. Namely, let us look at (P) as at a program where we should choose all decisions $\{\xi_i^l, \eta_i^l, \zeta_i^l \mid 1 \le l \le L, 1 \le i \le n+1\}$ at the very first time instant $l = 1$ (and, consequently, these decisions are just reals, not functions of ω^l); the reader is kindly asked to suppress his natural reaction to this disastrous assumption and not to throw the paper away at least till the discussion in Section 1 is read.

After we have agreed to treat the decision variables in (P) as reals and not functions of ω^l, we can treat (P) as a usual Linear Programming with uncertain data and to apply to this uncertain optimization program the Robust Counterpart approach as presented in [1, 2].

The Robust Counterpart approach to uncertain Linear Programming problems. The approach is as follows. Consider an uncertain Linear Programming problem

$$(\mathcal{P}) \qquad \max_X \{c^T X \mid AX + b \ge 0\},$$

X being N-dimensional decision vector, c being an exactly known ("certain") objective and $[A; b] = \begin{bmatrix} a_1^T, b_1 \\ \cdots \\ a_m^T, b_m \end{bmatrix}$ being "uncertain" $m \times (N+1)$ constraint matrix; all known about this matrix is that it belongs to a given *uncertainty set* $\mathcal{U} \subset \mathbf{R}^{m \times (N+1)}$. With the Robust Counterpart approach, we treat as "robust feasible" solutions to the problem those X which satisfy the constraints whatever is the realization of "instance data" $[A; b]$ from the uncertainty set; in other words, a robust feasible solution to uncertain problem in question should satisfy the system of inequalities

$$AX + b \geq 0 \quad \forall([A; b] \in \mathcal{U})$$

with this approach, it is natural to define the *robust optimal* solution to $(\mathcal{P})$ as the robust feasible solution with the smallest possible value of the objective, i.e., the optimal solution of the *Robust Counterpart*

$$(\mathcal{P}^*) \qquad \max_X \{c^T X \mid AX + b \geq 0 \quad \forall [A; b] \in \mathcal{U}\}.$$

of the uncertain problem $(\mathcal{P})$.

For typical uncertainty sets, $(\mathcal{P}^*)$ is a semi-infinite optimization problem. There are, however, cases when $(\mathcal{P}^*)$ turns out to be an "explicit" convex optimization program. One of these cases which is of especial interest for us is when the projections $\mathcal{U}_i$ of the uncertainty set $\mathcal{U}$ on the subspaces of data (a_i^T, b_i) of i-th linear constraint of $(\mathcal{P})$ for all i are ellipsoids:

$$\mathcal{U}_i = \{(a_i^T, b_i) = ([a_i^0]^T, b_i^0) + \sum_{j=1}^{k_i} u_j([a_i^j]^T, b_i^j) \mid u^T u \leq 1\}.$$

In this case $(\mathcal{P}^*)$ clearly is the conic quadratic problem

$$\begin{array}{c} \max_X \{c^T X \mid [a_i^0]^T X + b_i^0 - \|\beta_i + \alpha_i X\|_2 \geq 0, i = 1, ..., m\}, \\ \beta_i = \begin{pmatrix} b_i^1 \\ b_i^2 \\ \cdots \\ b_i^{k_i} \end{pmatrix}, \ \alpha_i = \begin{pmatrix} [a_i^1]^T \\ [a_i^2]^T \\ \cdots \\ [a_i^{k_i}]^T \end{pmatrix}. \end{array} \tag{1.10}$$

Now, the problem (P) we are interested in is a Linear Programming problem of the form $(\mathcal{P})$, and we may apply to it the outlined approach, and all we need is to specify somehow the uncertainty set $\mathcal{U}$ in the space of data $[A, b]$. To this end, consider the following reasoning. Let us denote by π^l, $l = 2, 3, ..., L+1$, the vector of design variables affected by uncertainty in the inequality (I_l):

$$\begin{array}{rcl} \pi_l & = & \begin{pmatrix} \eta^l \\ \zeta^l \end{pmatrix}, \ l = 2, ..., L, \\ \pi_{L+1} & = & \xi^L, \end{array} \tag{1.11}$$

and let P_l be the corresponding vectors of uncertain coefficients:

$$\begin{array}{rcl} P_l & = & \begin{pmatrix} A^l \\ -B^l \end{pmatrix}, \ l = 2, ..., L, \\ P_{L+1} & = & R^{L+1}. \end{array} \tag{1.12}$$

According to our initial assumption on the stochastic nature of the uncertain data, we know the distribution of the uncertain parameter vectors $P_2, ..., P_{L+1}$, in particular, their expectations p_l and their covariance matrices V^l. Now, the l-th uncertain inequality in (P) is of the form

$$a_l^T X + b_l \leq P_l^T \pi_l, \tag{1.13}$$

where X is the vector comprised of all design variables, a_l is certain coefficient vector and b_l is a certain constant; note that π_l is a known linear function of X. When X is fixed, the right hand side in (1.13) is a random variable with expectation $p_l^T\pi_l$ and variance $v_l(X) = \pi_l^T V^l \pi_l$. Now let us act as an engineer who assumes that a random real is "never" less than its mean minus θ times its standard deviation (an engineer would set $\theta = 3$, but we should not be that specific). With this "engineering" approach, a "safe" deterministic version of the constraint (1.13) is the usual ("certain") constraint

$$a_l^T X + b_l \le p_l^T \pi_l - \theta_l \sqrt{\pi_l^T V^l \pi_l}, \tag{1.14}$$

where $\theta_l > 0$ is a "safety parameter" we choose for the l-th of our original uncertain constraints.

Replacing all uncertain inequalities in (P) by their "safe versions", we end up with the following "safe version" of (P):

(P^+)

$$\max \quad w$$

s.t.

$$\begin{array}{rcl}
w + \theta_{L+1}\sqrt{[\xi^L]^T V^{L+1} [\xi^L]} & \le & \sum\limits_{i=1}^{n+1} \rho_i^{L+1} \xi_i^L \\
\xi_i^l & = & \xi_i^{l-1} - \eta_i^l + \zeta_i^l, \\
 & & \qquad i = 1, ..., n, l = 1, ..., L; \\
\xi_{n+1}^l + \theta_l \sqrt{\begin{pmatrix} \eta^l \\ \zeta^l \end{pmatrix}^T V^l \begin{pmatrix} \eta^l \\ \zeta^l \end{pmatrix}} & \le & \xi_{n+1}^{l-1} + \sum\limits_{i=1}^{n} \alpha_i^l \eta_i^l + \sum\limits_{i=1}^{n} \beta_i^l \zeta_i^l, \\
 & & \qquad l = 1, ..., L; \\
\eta_i^l & \ge & 0, \; i = 1, ..., n, l = 1, ..., L; \\
\zeta_i^l & \ge & 0, \; i = 1, ..., n, l = 1, ..., L; \\
\xi_i^l & \ge & 0, \; i = 1, ..., n+1, l = 1, ..., L.
\end{array} \tag{1.15}$$

Here:

• ρ_i^{L+1} are the expectations of R_i^{L+1}, and V^{L+1} is the covariance matrix of the random variables R_i^{L+1}

• for $l = 1, ..., L$, the vector $\begin{pmatrix} \alpha_i^l \\ \beta_i^l \end{pmatrix}$ is the expectation of the random vector $\begin{pmatrix} (1-\mu_1^l) R_1^l / R_{n+1}^l \\ \cdots \\ (1-\mu_n^l) R_n^l / R_{n+1}^l \\ -(1+\nu_1^l) R_1^l / R_{n+1}^l \\ \cdots \\ -(1+\nu_n^l) R_n^l / R_{n+1}^l \end{pmatrix}$, and the matrix V^l is the covariance matrix of this random vector (in fact this vector indeed is random only for $l > 1$) It can be easily demonstrated that (P^+) is nothing but the Robust Counterpart of (S) associated with the uncertainty set $\mathcal{U}$ which is the direct product of the ellipsoids

$$\{(P_l - p_l)^T (V^l)^{-1} (P_l - p_l) \le \theta_l^2\}, \; l = 1, ..., L.$$

The optimization program (P^+) is a ("certain") convex optimization program with nice analytical structure, and this is exactly the model of the uncertain Portfolio problem we propose.

Simplification of the Robust Portfolio model

In the "certain" problem (P^*) the variables η_i^l, ζ_i^l are, in a sense, redundant and can be eliminated. Indeed, we can write

$$\eta_i^l - \zeta_i^l = \xi_i^{l-1} - \xi_i^l;$$

for a given value of the right hand side in this equality, maximum contribution of $\eta_i^l \geq 0, \zeta_i^l \geq 0$ to the right hand side of the corresponding cash flow inequality

$$\xi_{n+1}^l = \xi_{n-1}^{l-1} + \sum_{i=1}^n A_i^l \eta_i^l - \sum_{i=1}^n B_i^l \zeta_i^l$$

is attained when

$$\begin{array}{rcl} \eta_i^l & = & (\xi_i^{l-1} - \xi_i^l)_+ \equiv \max\{\xi_i^{l-1} - \xi_i^l, 0\} \\ \zeta_i^l & = & (\xi_i^l - \xi_i^{l-1})_+ \equiv \max\{\xi_i^l - \xi_i^{l-1}, 0\} \end{array} \tag{1.16}$$

(note that by their origin $A_i < B_i$), and of course it is profitable to choose η_i^l and ζ_i^l accordingly. This observation reflects a completely evident fact: at every time instant, it never makes sense both to buy and to sell positive amounts of an asset – because of the transaction costs, we should either buy, or sell the asset, but not both simultaneously.

Relation (1.16) allows to eliminate from (P^*) all η- and ζ- variables, along with balance equations for non-cash assets. Of course, this elimination destroys the linear structure of the cash flow inequalities – they became convex nonlinear constraints.

Since the Robust Counterpart (P^+) of (P^*) from the very beginning is a nonlinear convex program, it might make sense to ask whether the possibility to eliminate η- and ζ-variables is inherited by (P^+). We were able to prove the corresponding statement only under additional, although not very restrictive, assumptions. Here is the result.

Lemma 1.0.1 *Let* $\psi_i^l = B_i^l - A_i^l \quad [= (\mu_i^l + \nu_i^l) R_i^l / R_{n+1}^l]$. *Assume that for some set* $\mathcal{I}$ *of pairs* (i, l) $(i \in \{1, ..., n\}, l \in \{1, ..., L\})$ *one has*

$$E\{(\psi_i^l)^2\} \leq (\theta_l^{-2} + 1)[E\{\psi_i^l\}]^2 \tag{1.17}$$

(E stands for expectation). Then there exists optimal solution to (P^+) *for which relations* (1.16) *are satisfied for all* $(i, l) \in \mathcal{I}$.

Proof. (P^+) clearly is feasible (a feasible solution is given by $\xi_i^l = \xi_i^0$, $\eta_i^l = \zeta_i^l = 0$ for all i and l), and the feasible domain of the problem clearly is bounded. It follows that the problem is solvable. Let $\{\xi_i^l, \eta_i^l, \zeta_i^l\}$ be an optimal solution in this problem, and let $(i,l) \in \mathcal{I}$. It suffices to verify that if we modify this solution, varying only the quantities η_i^l and ζ_i^l according to (cf. (1.16))

$$(\eta_i^l, \zeta_i^l) \mapsto ((\xi_i^{l-1} - \xi_i^l)_+, (\xi_i^l - \xi_i^{l-1})_+),$$

then the updated solution remains feasible (and then - optimal, since this updating clearly does not vary the value of the objective). Given the announced fact and applying sequentially to the initial optimal solution the indicated modifications for all $(i,l) \in \mathcal{I}$, we will end up with the optimal solution required in Lemma.

The modification in question affects only the balance equation for asset i at the time instant l and the cash flow inequality at this instant. The first of these equations clearly remains valid, so that all we need is to demonstrate that our modification does not decrease the right hand side of l-th cash flow inequality. This is exactly the same as to verify that the function

$$\begin{array}{rcl} g_i^l(t) &=& E\{A_i^l\}(t+\delta_i^l) - E\{B_i^l\}t - \theta_l\sqrt{E\{h^2(t)\}}, \\ \delta_i^l &=& \eta_i^l - \zeta_i^l = \xi_i^{l-1} - \xi_i^l, \\ h(t) &=& \sum\limits_{j=1}^n h_j(t), \\ h_j(t) &=& \left\{\begin{array}{ll} [A_i^l - E\{A_i^l\}](t+\delta_i^l) - [B_i^l - E\{B_i^l\}]t, & j = i \\ [A_j^l - E\{A_j^l\}]\eta_j^l - [B_j^l - E\{B_j^l\}]\zeta_i^l, & j \neq i \end{array}\right. \end{array}$$

is nonincreasing in t. Indeed, when we vary only η_i^l and ζ_i^l, preserving the balance equalities (i.e., replace η_i^l by $\delta_i^l + t$ and ζ_i^l by t, $t \in \Delta_i^l = [\max\{-\delta_i^l, 0\}, \infty)$) $g_i^l(t)$, up to independent of t additive term, is exactly the difference between the right and the left hand sides of the l-th cash flow inequality; this difference is nonnegative when $t = \zeta_i^l \in \Delta_i^l$, and given that it is nonincreasing in t, we could conclude it is also nonnegative when t is the left endpoint of Δ_i^l, and it would mean that the the modification of the solution we are interested in indeed is feasible).

To prove that $g_i^l(t)$ is nonincreasing in t, let us compute the derivative of the function:

$$\begin{array}{rcl} (g_i^l)'(t) &=& E\{A_i^l - B_i^l\} - \theta_l \frac{E\{h(t)h'(t)\}}{\sqrt{E\{h^2(t)\}}} \\ &\leq& E\{A_i^l - B_i^l\} + \theta_l\sqrt{E\{[h'(t)]^2\}} \\ && \text{[we have used the Cauchy inequality]} \\ &=& -E\{\psi_i^l\} + \theta_l\sqrt{E\{[-\psi_i^l + E\{\psi_i^l\}]^2\}} \\ && \text{[since } h'(t) = -\psi_i^l + E\{\psi_i^l\}] \\ &=& -E\{\psi_i^l\} + \theta_l\sqrt{E\{[\psi_i^l]^2\} - [E\{\psi_i^l\}]^2} \\ &\leq& 0 \quad \text{[see (1.17)]} \quad \blacksquare \end{array}$$

Remark 1.0.1 Note that (1.16) for sure is satisfied when $l = 1$, since then ψ_i^l are not random. It follows that when we use the robust counterpart approach in the rolling horizon mode (see below), our actual decisions never are both to buy and to sell a given asset at a given time instant.

Let us look whether (1.16) is satisfied for the case when μ_i^l and ν_i^l are certain positive constants and the vector comprised of $\ln r_i^l$ for all i,l has joint normal distribution. In this case the quantity $\ln \psi_i^l$ also has normal distribution, say, $N(\mu = \mu(i,l), \sigma^2 = \sigma^2(i,l))$. It follows that

$$\begin{array}{rcl} E\{[\psi_i^l]^p\} & = & (2\pi)^{-1/2}\sigma^{-1} \int \exp\{px - \frac{(x-\mu)^2}{2\sigma^2}\}dx \\ & = & (2\pi)^{-1/2}\sigma^{-1} \int \{exp\{\frac{(x-\mu-\sigma^2 p)^2}{2\sigma^2}\} \exp\{p\mu + \frac{p^2\sigma^2}{2}\}dx \\ & = & \exp\{(p\mu + \frac{p^2\sigma^2}{2}\}, \end{array} \tag{1.18}$$

whence

$$\frac{E\{[\psi_i^l]^2\}}{[E\{\psi_i^l\}]^2} = \exp\{\sigma^2\}.$$

Thus, in the case in question (1.17) is satisfied if

$$\theta_l \leq \frac{1}{\sqrt{\exp\{\sigma^2(i,l)\} - 1}} \quad \forall i,l. \tag{1.19}$$

Note also that whether it indeed makes sense to eliminate the η- and ζ-variables when possible, it depends on the numerical technique used to solve (P^+). With the interior point methods, this elimination hardly makes sense, since it complicates the analytical structure of the problem, and when solving the problem by interior point methods, we basically should reintroduce the eliminated variables. In contrast to this, it definitely makes sense to eliminate the variables in question when (P^+) is solved by nonsmooth optimization technique (the bundle methods).

Discussion

The reasoning which led us from (P) to (P^+) is, of course, a "common sense" reasoning, not a rigorous mathematical deduction; however, this is not a severe sin – we were building a mathematical model, and modeling of a real-world problem always lies beyond the bounds of mathematics. A sin, if any, is in treating (P) as a problem where all decisions should be made at the very first time instant, while in the actual portfolio management the decisions which should be implemented at time instant $l = 1, 2, ..., L$ may depend on the data which are unknown at the very first time instant, but become known at the time instant l. Whether the latter "sin of conservativeness" is sufficient to discard the model (P^+) in advance or not, it depends on whether we have in our disposal something better. An evident answer would be: "of course, we have something much better – the Multistage Stochastic Programming". We, however, would argue that the advantages of Stochastic Programming are not that evident at all. Indeed, it is true that in the Multistage Stochastic Programming approach we start with an adequate model of the actual process; but then we are supposed to carry out huge specific effort in order to approximate the initial computationally intractable model by a computationally tractable one. At this approximation stage, we have no a priori guarantees that the "tractable approximation" we end up with still will be relevant to the actual process, that we do not buy computational tractability at the price of lack of relevance. In a sense, the latter always is the case: the Multistage Stochastic Programming at best can provide us with good first stage ("here and now") components of the decisions, while the decisions of the subsequent stages – which, in the Stochastic Programming model of the Portfolio problem, are functions of continuous multidimensional arguments, cannot even be fully stored. As a result, the only possible way to apply Multistage Stochastic Programming in practice is to use it in the *rolling horizon* mode – at the first time instant, to approximate the multistage stochastic model by something computationally tractable and to implement the "here and now" part of the decisions, at the second time instant, to solve a new problem with reduced by 1 time horizon and to implement the "here and now" part of the decisions yielded by the solution, etc. There would be nothing bad in the "rolling horizon" scheme, if we were sure that the "here and now" part of the decisions given by the Stochastic Programming approach indeed comes from a nearly optimal solution to the Multistage Stochastic Programming model, but in fact these decisions come from computationally tractable approximation of the latter model, and we already have mentioned that there are no "ready-to-use" techniques capable to guarantee high-quality computationally efficient approximation of the optimal solution to the Multistage Stochastic Programming model.

In contrast to the Multistage Stochastic Programming model, the "Robust Counterpart" model (P^+) does not pretend to be completely adequate to the actual process; as a compensation, (P^+) is an explicit Convex Programming program with nice analytic structure, a program perfectly well suited for modern interior point methods, and therefore (P^+) can be routinely and efficiently built and processed computationally. Exploiting this model in the rolling horizon model, we may hope to eliminate to some extent the influence of the aforementioned "sin of conservativeness".

What is better for real world applications – to start with an adequate model which should be unavoidably "spoiled" in course of "ad hoc" numerical processing or to start with a rough model which for sure can be routinely and efficiently processed – this question can be resolved only in practice. Not trying to predict the answer, we, however, strongly believe that the approach we have presented is worthy of testing. In the remaining part of this paper, we present the results of a preliminary simulation-based testing of this type.

Before passing to numerical results, it makes sense to discuss an additional modeling issue – why the Robust Counterpart approach was applied to problem (P) rather than to the original problem (P_*). Of course, (P_*) itself cannot be treated via the Robust Counterpart approach, since this is a problem with uncertain *equality* constraints, and the straightforward Robust Counterpart to such a problem is normally infeasible. However, in the case of certain data (P_*) clearly is equivalent to its inequality constrained version (P^*) (all equalities are replaced with the inequalities $\leq$); why not to apply the Robust Counterpart approach to (P^*)? The answer is as follows: the problems (P^*) and (P) are equivalent as problems with certain data, and are *not* equivalent as uncertain problems to be processed by the Robust Counterpart approach. Indeed, a robust feasible solution to (P^*) is a once for ever prescribed sequence of *concrete actions*: at the first time instant, sell and buy these and these amounts of every asset, at the second time instant sell and by these and these amounts, and so on; the actions related to time instant l are completely independent of what happened with the market before this instant. In contrast to this, a robust feasible solution to (P) prescribes a behaviour which does depend, although in a simple way, on what goes on with the market: the amounts of assets to be bought and sold at time instant l are proportional to the quantities R_i^l, the proportionality coefficients being given by the solution in question. Although these coefficients are independent of what happens with the market, the quantities R_i^l do depend on market's behaviour, and so are the actions prescribed by a solution to (P). Now, since (P^*) and (P), treated as uncertain problems which we are going to process with the Robust Counterpart approach, are not equivalent to each other, the natural question is which one is better suited for this approach. The "common sense" answer is definitely in favour of (P) by the following reasons. Basically all constraints in (P^*) are uncertain, and most of them involve a single uncertain coefficient each. Applying the above reasoning to get a "safe" version of the constraints of this latter type, we end up with the original constraint with the uncertain coefficient being replaced with its nearly worst possible value, which is extremely conservative. In contrast to this, (P) involves just L uncertain constraints, and every one of them is affected by large number n of uncertain coefficients. If the dependencies between these coefficients are not too strong, we may hope that "bad" values of some of them will be to some extent compensated by "good" values of others, so that the robust version of the constraint will be not that conservative.

To illustrate this important point, consider the following "extreme" example: there is just one time slot, at the beginning of it we have \$ 1 in cash and no other assets; the problem is to distribute part of this cash between n assets in order to maximize the value of the resulting portfolio at the end of the time slot. In other words, we should solve the problem

$$\max_x \{y \sum_{i=1}^n \mid y \le \sum_{i=1}^n r_i x_i,\ x \ge 0, \sum_{i=1}^n x_i \le 1\} \tag{1.20}$$

As stated, the problem is of the type (P); the analogy of (P_*) in this case is the problem

$$\max_{x,y} \{\sum_{i=1}^n y_i \mid 0 \le y_i \le r_i x_i, x \ge 0, \sum_i x_i \le 1\}. \tag{1.21}$$

Now assume that r_i are, say, log-normal independent random variables with expectations ρ_i and standard deviations σ_i, all these quantities being of the same order of magnitude:

$$\rho_i, \sigma_i \in [1/\kappa, \kappa]$$

for a once for ever fixed $\kappa > 1$. As applied to (1.20), the Robust Counterpart approach yields the problem

$$\max_x \{\sum_{i=1}^n \rho_i x_i - \theta \sqrt{\sum_{i=1}^n \sigma_i^2 x_i^2} \mid x \geq 0, \sum_i x_i \leq 1\} \tag{1.22}$$

(note that this is, basically, the Markovitz model of the problem). For once for ever fixed small $\alpha > 0$, one can choose $\theta = \theta(\alpha)$ in such a way that, uniformly in n and in $x \in \mathbf{R}^n$, $x \geq 0$, $\sum\limits_{i=1}^n x_i \leq 1$, the probability of the event

$$\{\sum_{i=1}^n r_i x_i < \sum_{i=1}^n \rho_i x_i - \theta \sqrt{\sum_i \sigma_i^2 x_i^2}\}$$

will be less than α; thus, the actual portfolio value yielded by the optimal solution to (1.22) will be less than the optimal value of the latter problem with small ($\leq \alpha$) probability. On the other hand, for large n the optimal value in (1.22) clearly is at least

$$\gamma_n = \frac{1}{n} \sum_{i=1}^n \rho_i - \theta n^{-1} \sqrt{\sum_{i=1}^n \sigma_i^2} \geq (1 - o(1)) \frac{1}{n} \sum_{i=1}^n \rho_i$$

and corresponds to the optimal solution with $\sum\limits_i x_i = 1$. Thus, at least for large n the Robust Counterpart approach with properly chosen safety parameter θ enforces us to invest all our resources in the assets and guarantees, with probability $\geq 1 - \alpha$, yield at least $O(1)$. On the other hand, the same Robust Counterpart approach as applied to (1.21) results in the problem

$$\max\{\sum_{i=1}^n y_i \mid 0 \leq y_i \leq (\rho_i - \theta\sigma_i) x_i, x_i \geq 0, \sum_{i=1}^n x_i \leq 1\}.$$

Whenever $\theta \geq \kappa^2$, this approach yields the policy $x_i = 0,$, $i = 1, ..., n$, and results in the zero yield. Thus, (1.20) is incomparably better suited for the Robust Counterpart approach than (1.21).

SIMULATED NUMERICAL RESULTS

Since the approach in question deals with modeling issues, the only way to evaluate its actual potential is to look how it works in practice. To the moment we are able to report results coming from experiments with simulated market, and below we present in full details all issues related to our simulations.

Stochastic model of the market

The stochastic model of the data we use in our simulations is a simple factor model (cf. [3]) as follows:

$$\begin{array}{rcl} \ln r_i^l & = & \Omega_i^T[\kappa e + \sigma v^l], \\ & & l = 0, 1, ..., L, i = 1, ..., n; \\ \ln r_{n+1}^l & = & \kappa, \\ & & l = 0, 1, ..., L. \end{array} \tag{1.23}$$

where

• $\{v^0, v^1, ..., v^L\}$ are independent k-dimensional Gaussian random vectors with zero mean and the unit covariance matrix,

• $e = (1, ..., 1)^T \in \mathbf{R}^k$,

• $\Omega_i \in \mathbf{R}_+^k$ are fixed vectors,

and

• $\kappa, \sigma > 0$ are fixed reals.

Note that the random vectors $r^l = \{r_i^l\}_{i=1}^n$, $l = 0, 1, ..., L$, are i.i.d., while the coordinates of every vector are dependent on each other. Note also that according to (1.23), the cash returns $r_{n+1}^l = \exp\{\kappa\}$ are deterministic and independent of time; this assumption is made with the only purpose to simplify the simulation. By the same reasons, the transaction costs also are assumed to be deterministic and independent of time and of asset's type:

$$\mu_i^l = \mu, \nu_i^l = \nu \quad \forall i = 1, ..., n, l = 1, ..., L. \tag{1.24}$$

Final form of the Robust Portfolio model

Given the stochastic data model (1.23), we can compute the expectations and covariance matrices involved into model (P^+). Note that according to Assumption **A**, when building and solving the latter model, we already know the returns r_i^0, $i = 1, ..., n$, so that the expectations and covariance matrices in question should be taken w.r.t. the corresponding conditional distribution of the data. Assuming that the matrix $n \times k$ matrix $\Omega = \begin{bmatrix} \Omega_1^T \\ ... \\ \Omega_n^T \end{bmatrix}$ is of rank k, we conclude that the conditional expectations/covariances in question are in fact expectations/covariances taken over the distribution of $v^1, v^2, ..., v^L$. A straightforward computation demonstrate that (P^+) is nothing but the program

$$
\begin{array}{lrcl}
(P^*) & & & \\
 & & \max & w \\
\text{s.t.} & & & \\
 & w + \theta_L \sqrt{[\xi^L]^T V^{L+1} [\xi^L]} & \leq & \sum\limits_{i=1}^{n+1} \rho_i^{L+1} \xi_i^L \\
 & \xi_{n+1}^l + \theta_l \sqrt{\begin{pmatrix} \eta^l(\xi) \\ \zeta^l(\xi) \end{pmatrix}^T V^l \begin{pmatrix} \eta^l(\xi) \\ \zeta^l(\xi) \end{pmatrix}} & \leq & \xi_{n+1}^{l-1} + \sum\limits_{i=1}^{n} \alpha_i^l \eta_i^l(\xi) \\
 & & & + \sum\limits_{i=1}^{n} \beta_i^l \zeta_i^l(\xi), \\
 & & & l = 1, ..., L; \\
 & \eta_i^l(\xi) & = & \max[\xi_i^{l-1} - \xi_i^l, 0], \\
 & & & i = 1, ..., n, l = 1, ..., L; \\
 & \zeta_i^l(\xi) & = & \max[\xi_i^l - \xi_i^{l-1}, 0], \\
 & & & i = 1, ..., n, l = 1, ..., L; \\
 & \xi_i^l & \geq & 0, \\
 & & & i = 1, ..., n+1, l = 1, ..., L.
\end{array}
\tag{1.25}
$$

In this problem,
- the design variables are $w, \{\xi_i^l \mid l = 1, ..., L, i = 1, ..., n+1\}$;
- the data are

– the quantities $\{\xi_i^0 \geq 0\}_{i=1}^{n+1}$ representing the initial state of the portfolio;
– the positive safety parameters $\theta_1, ..., \theta_L$;
– the reals α_i^l, β_i^l and the symmetric matrices V^l given by the parameters of the stochastic model (1.23) according to the relations

$$
\begin{array}{rcl}
\rho_i^{L+1} & = & \exp\left\{\Omega_i^T t^1 + (L+1)\omega_i \kappa + \frac{\lambda_i^2 L \sigma^2}{2}\right\} \\
 & & 1 \leq i \leq n, \\
\rho_{n+1}^{L+1} & = & \exp\{(L+1)\kappa\}; \\
\alpha_i & = & (1-\mu_i)\exp\left\{\Omega_i^T t^1 + l(\omega_i - 1)\kappa + \frac{\lambda_i^2 (l-1)\sigma^2}{2}\right\}, \\
\beta_i & = & -(1+\nu_i)\exp\left\{\Omega_i^T t^1 + l(\omega_i - 1)\kappa + \frac{\lambda_i^2 (l-1)\sigma^2}{2}\right\}, \\
 & & 1 \leq i \leq n, 1 \leq l \leq L;
\end{array}
\tag{1.26}
$$

$$
\begin{array}{rcl}
V_{ij}^l & = & \begin{cases} (1-\mu_i)(1-\mu_j)\Sigma_{i,j}^l, & 1 \leq i, j \leq n \\ -(1-\mu_i)(1+\nu_{j-n})\Sigma_{i,j-n}^l, & 1 \leq i \leq n, n < j \leq 2n \\ (1+\nu_{i-n})(1+\nu_{j-n})\Sigma_{i-n,j-n}^l, & n < i, j \leq 2n \end{cases}, \\
 & & 1 \leq i \leq j \leq 2n, 1 \leq l \leq L,
\end{array}
\tag{1.27}
$$

where for $1 \leq i, j \leq n$

$$
\begin{array}{rcl}
\Sigma_{i,j}^l & = & \exp\left\{l\kappa(\omega_i + \omega_j - 2) + (\Omega_i + \Omega_j)^T t^1 + \frac{(\lambda_i^2 + \lambda_j^2)(l-1)\sigma^2}{2}\right\} \\
 & & \times \left[\exp\{\Omega_i^T \Omega_j (l-1)\sigma^2\} - 1\right]
\end{array}
$$

(note that $V^1 = 0$);

$$\begin{array}{rcl} V_{ij}^{L+1} & = & \exp\left\{(\Omega_i+\Omega_j)^T t^1 + (L+1)\kappa(\omega_i+\omega_j) + \frac{(\lambda_i^2+\lambda_j^2)L\sigma^2}{2}\right\} \\ & & \times \left[\exp\{\Omega_i^T\Omega_j L\sigma^2\} - 1\right], \\ & & i \le i \le j \le n, \\ V_{i,n+1}^{L+1} & = & 0. \end{array} \tag{1.28}$$

Here for $1 \le i \le n$

$$\begin{array}{rcl} \omega_i & = & \Omega_i^T e, \\ \lambda_i & = & \sqrt{\Omega_i^T\Omega_i}, \end{array}$$

Note that in (1.25) we have already eliminated the η- and the ζ-variables; according to Lemma 1.0.1 and (1.19), it for sure is possible if

$$\theta_l \le \frac{1}{\sqrt{\max_{1\le i\le n}\exp\{\lambda_i^2(l-1)\sigma^2\} - 1}},\ 2 \le l \le L \tag{1.29}$$

(the value of θ_1 is unimportant, since $V^1 = 0$); from now on, we assume that (1.29) indeed takes place. In fact, in our simulations we were choosing θ_l, $2 \le l \le L$, according to

$$\theta_l = \min\left[\theta^*, \frac{\min(1,\theta^*)}{\sqrt{\max_i \exp\{\lambda_i^2(l-1)\sigma^2\} - 1}}\right], \tag{1.30}$$

θ^* being a setup parameter of the experiments.

Setup for market's model. In our simulations, the parameters of the stochastic data model (1.23) were specified according to a number of natural requirements as follows.

• Since the cash asset is risk-free, it is natural to ensure the (risky) non-cash assets to be more attractive than the cash, i.e., to ensure the expected returns $E\{r_i^l\}$, $i \le n$, to be $> \exp\{\kappa\}$.

Direct computation implies that (Mean denotes the expectation, and Std - the standard deviation of a random variable)

$$\begin{array}{rcl} \mathrm{Mean}(r_i^l) & = & \exp\{\omega_i\kappa + \lambda_i^2\sigma^2/2\}, \\ \mathrm{Std}(r_i^l) & = & \mathrm{Mean}(r_i^l)\sqrt{\exp\{\lambda_i^2\sigma^2\} - 1}. \end{array} \tag{1.31}$$

Assuming in accordance with reality that σ is of order of κ and both quantities are significantly less than 1 (i.e., that our time period is not that long; note that for the real economy rate of growth per year is few percents), we see that if ω_i is "significantly less" than 1, then $\mathrm{Mean}(r_i^l) < \exp\{\kappa\}$. Consequently, it makes sense to choose Ω_i in such a way that $\omega_i \ge 1$.

• The more attractive is a non-cash asset, the more " risky" it should be. From (1.31) we see that

$$\mathrm{Std}(r_i^l)/\mathrm{Mean}(r_i^l) = \exp\{\lambda_i^2\sigma^2\} - 1,$$

so that the "risk" – the left hand side ratio – grows with λ_i; this is close to what we need (in fact we are interested to have risk which grows with ω_i rather than with λ_i)

• In order to make the experiments more interesting, we should ensure that at least the most attractive assets – those with the largest ω_i – should be risky as compared to the cash, i.e., the corresponding returns should, with "significant" probability, be worse than $\exp\{\kappa\}$.

The random variable $\ln r_i^l$ $(i \le n)$ in our data model is Gaussian random variable with the mean $\omega_i\kappa$ and the standard deviation $\lambda_i\sigma$. In order for the probability of the event $\ln r_i^l < \kappa$ to be "significant", the ratio $\frac{(\omega_i-1)\kappa}{\lambda_i\sigma}$ should be at most a moderate constant γ, something like $0.5 - 1 - 1.5 - 2$. Now, by Cauchy's inequality $\omega_i \le \lambda_i\sqrt{k}$, so that the ratio in question is at most $\frac{(\omega_i-1)\kappa}{\omega_i\sigma}\sqrt{k}$. In order for this ratio to be $\le \gamma$, it suffices to have

$$k \le \gamma^2 \left(\frac{\omega_i}{\omega_i - 1}\right)^2 \frac{\sigma^2}{\kappa^2}. \tag{1.32}$$

In our experiments, we met the outlined requirements via parameterizing the model (1.23) by three "free parameters"

$$\kappa > 0, \quad \gamma \in [0.5, 1.5], \quad \omega_{\max} \in [1.5, 2]$$

in the following manner. First, we set

$$\begin{array}{rcl} \sigma &=& \kappa/\gamma; \\ k &=& \lfloor \left(\frac{\omega_{\max}}{\omega_{\max}-1}\right)^2 \rfloor; \\ k_i &=& i \min\left[\lfloor \frac{n-i}{n}k + \frac{i}{n} + 1\rfloor, k\right] \\ \omega_i &=& \frac{n-i}{n} + \frac{i}{n}\omega_{\max}, \\ && i = 1, ..., n. \end{array} \tag{1.33}$$

Second, for $1 \le i \le n$ we choose k-dimensional nonnegative vector Ω_i as follows:

– the number of nonzero entries in Ω_i is k_i, and the indices of these entries are chosen at random in $\{1, 2, ..., k\}$;

– the k_i-dimensional vector w_i comprised of nonzero entries of Ω_i is chosen at random in the simplex $\{w \in \mathbf{R}_+^{k_i} \mid \sum\limits_j w_j = \omega_i\}$.

With this setup, the quantities $\omega_i = \Omega_i^T e$ form a regular grid on $[1, \omega_{\max}]$, and the inequalities (1.32) indeed take place.

It is convenient to characterize the assets by their *risk indices* – probabilities

$$\mathcal{R}_i = \text{Prob}\{r_i^l < 1\}$$

to loose in dollar value of the asset in course of a single time period. Note that since in the model (1.23) the distribution of r_i^l is independent of l, the quantity $\mathcal{R}_i$ indeed depends solely on i. Note that with the outlined setup for the data model risk indices $\mathcal{R}_i$ typically grow with the "promise" ω_i of an asset. The maximum of risk indices of the assets

$$\mathcal{R} = \max_{1 \le i \le n} \mathcal{R}_i$$

is called *risk index* of the market; this indeed is a rough measure of how risky the market is.

Simulated policies and simulation scheme

In our experiments, we used the data model (1.23) with setup described in Section 1 in order to compare four portfolio management policies: the *Robust* (Rob), the *Stochastic Programming* (StP), the *Nominal* (Nom) and the *Conservative* (Cns) ones.

Robust policy. This policy is the one given by the Robust Portfolio model (P^*).

Stochastic Programmingi policy. This policy is given by a straightforward implementation of the Multistage Stochastic Programming approach. Namely, we

A. Fix positive integers $m_1, ..., m_{L-1}$ – *numbers of scenarios* for the periods starting at time instances $1, ..., L-1$. In order to save notation, we add to this collection also $m_0 = 1$.

B. For $l = 0, 2, ..., L-1$, we generate independently samples $\mathcal{T}_l$ comprised of m_l *scenarios* each; a scenario is a realization of random Gaussian k-dimensional vector with zero mean and unit covariance matrix. Let j-th scenario of l-th sample be denoted by $\tau^l[j]$, and let the vectors $r^l[j]$ be given by the relation (cf. (1.23))

$$r_i^l[j] = \begin{cases} \exp\{\omega_i \kappa + \sigma \Omega_i^T \tau^l[j]\}, & i = 1, ..., n \\ \exp\{\kappa\}, & i = n+1 \end{cases}$$

C. We replace the original continuous distribution (1.23) of random vectors of returns r^l by the distribution given by the same formula, but now the random vector v^l, instead of being Gaussian, takes values in the sample $\mathcal{T}_l$ with equal probabilities $1/m_l$, and the vectors τ^l for different l are independent of each other.

After replacing the true – continuous – distribution of the data with the indicated discrete distribution, we straightforwardly use the Multistage Stochastic Programming approach as outlined in Section 1. It is easily seen that in the case in question this approach results in a usual Linear Programming program ("the deterministic equivalent of the discretized multi-stage Stochastic Programming problem") with

$$N = (3n+1) \sum_{l=0}^{L-1} m_0 m_1 ... m_l$$

nonnegative design variables and

$$M = (n+1)\sum_{l=0}^{L-1} m_0 m_1 ... m_l$$

linear equality constraints.

With the outlined implementation of the Multistage Stochastic Programming approach, the main question which should be resolved is how to choose the numbers of scenarios $m_1, ..., m_{L-1}$. The limiting factor is, of course, the size of the resulting LP program. As it will become clear in a while, in our experiments we were supposed to solve hundreds of these LP's; in order to make experimentation not too time-consuming (days, not weeks), we have restricted the design dimension N of the LP's not to exceed 10,000. For 3-stage problem ($L = 3$) with $n = 30$ assets – these were the sizes we used in most of experiments – this convention implies the bound

$$m_1 + m_1 m_2 \leq 109,$$

which is very restrictive; e.g., in the "equi-discretized" case $m_1 = m_2$ we have to use not more than 9 scenarios per stage (and this should represent distribution of 30-dimensional random vector!). Note that even if we were ready to increase the design dimension of the resulting LP to 100,000, the upper bound on $m_1 = m_2$ would become 32. By the indicated reasons, in our experiments we use 2-stage approximation of the 3-stage program, i.e., set $m_2 = 1$, which under the restriction $N \leq 10,000$ yields $m_1 = 54$. The resulting LP program has 9,919 design variables and 3,379 constraints. For comparison: the Robust Portfolio model (P^*) in the case in question is a convex optimization program with 94 nonnegative design variables, linear objective and 3 nonlinear constraints.

Nominal policy. This policy is very simple and rough: here we model the situation by the usual LP program (P), where all uncertain data are replaced with their expected values.

Conservative policy. This policy is motivated by the fact that in our stochastic data model (1.23) there is a risk-free asset – the cash; the policy in question is just to sell all other assets at the very first time instant and never buy them again.

Rolling horizon simulation. All outlined policies were tested in the Rolling horizon mode. In other words, to test, say, the Robust policy as applied to a 3-stage problem, we start with building and solving the 3-stage Robust Portfolio model (P^*) and update the initial portfolio according to the optimal, from the viewpoint of (P^*), decisions related to the very first time instant (here and below all time instants/periods relate to "absolute" time scale). Then we simulate the returns r_i^1 according to (1.23), thus imitating the the behaviour of the market during the first time period. After this period is passed, we build and solve the 2-stage Robust Portfolio model corresponding to the current (corresponding to time instant 2 on the absolute time scale) state of portfolio/market and the remaining two time periods. The optimal decisions related to the first stage of this 2-stage problem define the updating of the portfolio at time instant 2. This updating, along with simulated behaviour of the market during the second time period, define the state of portfolio/market at the third time instant, where we should for the last time update the portfolio. This updating is given by the 1-stage Robust Portfolio problem associated with the state of the portfolio/market at the third time instant. Simulating the behaviour of the market at the third time period, we can compute the value of the resulting portfolio at the fourth time instant, thus getting a realization of the yield of the Robust policy as applied to the 3-stage Portfolio Management problem.

In our experiments, all four policies were tested in the rolling horizon mode (in fact, of course, there is no actual necessity to test in this mode the Conservative policy – the result can be predicted at the very first time instant).

Running the experiments. In our experimentation, we dealt with 3-stage problems ($L = 3$) and $n = 30$ assets. A single *experiment* starts with setting up the "free" parameters $\theta^*, \kappa, \gamma, \omega_{\max}$ and generating the "number of independent factors" k and vectors Ω_i, $i = 1, ..., n$, as explained in Section 1. Then 50 *major simulations* are run. At a single simulation, we first generate (at random) the initial state of the portfolio $\{x_i^0\}_{i=0}^{n+1}$ and the "trajectory" of the market – a sample v^0, v^1, v^2 of independent Gaussian k-dimensional vectors which, via (1.23), determines the asset returns r_i^l.

A major simulation itself consists of rolling horizon testing of all our four policies, for the chosen initial state of the portfolio and the trajectory of the market. Note that the final value of the portfolio, for given management policy and market trajectory, still is a random variable: its value depend on random asset returns $r^4 = \{r_i^4\}_{i=1}^n$. Thus, a single major simulation gives rise to 4 random final values of the portfolio, one per policy (in fact, of course, the value of the portfolio for the Conservative policy is not random). "Within" every major simulation, we generate 100 realizations of random vector r^4 and compare to each other the corresponding 4 final values of the portfolio.

Robust Portfolio models (P^*) were processed by the *Bundle-Level* method for nonsmooth constrained convex optimization [4]; the LP problems responsible for the Stochastic Programming and Nominal policies were solved by CPLEX.

Computational results

We are about to present the results of four typical experiments. In all these experiments, the transaction costs were set to 0.1; the parameter θ^* used to specify the safety parameters in the Robust Portfolio model (see (1.30) was set to 1.

The setups for the stochastic data model (1.23) are given in Table 1, and the results are represented by Tables 1.1 – 1.2.

Table 1.1 Setup of the stochastic data model (1.23).

Exp #	κ	γ	$\omega_{\max}$	*Risk index*
1	0.1	0.330	1.2	0.339
2	0.1	0.250	1.2	0.377
3	0.1	0.216	1.2	0.393
4	0.1	0.200	1.2	0.401

Table 1.1 represents the *gain* in the portfolio value, i.e., the ratio of the final (at time instant 4) dollar value of the portfolio to its initial dollar value $\sum\limits_{i=1}^{n+1} x_i^0$. The gain is, of course, a random variable, and we give, for each experiment/policy, the empirical maximum, minimum, average and standard deviation of the gain, along with empirical probabilities of "loss" (the gain is < 1) and "big loss" (the gain is < 0.8).

Table 1.2 represents "pair dominance" of the four policies in question. i.e., empirical probabilities (in percents) for the "row" policy to yield better gain than the "column" policy.

Conclusions. The conclusions of the outlined experiments can be formulated as follows.

1. The policy based on Robust Portfolio model indeed is robust – for risky market, the corresponding standard deviation of the gain in the portfolio value is by factor 5 - 8 less than for the Nominal policy and by factor 4.7 - 5 less than for the SP (Stochastic Programming) policy. The Robust policy in the reported experiments never results in losses, while for the Nominal and the SP policies the probabilities of losses (big losses), for the most risky market, are quite significant (15 - 20 %).

2. From the viewpoint of average gain, the Robust policy is nearly optimal, except the case of the most risky market, where the Robust policy is slightly dominated by the SP and significantly – by the Nominal ones. Note, however, that the distributions in question are asymmetric, so that dominance in the mean is not that informative. E.g., from Table 3 it is seen that the probability to get better results than with every one of the competitors, is $> 1/2$ (except the least risky market, where the Robust policy is slightly dominated in this sense by the Nominal one).

3. Surprisingly enough, the SP policy, which traditionally is supposed to be the most adequate one, seems to have no advantages at all: as far as the expected gain is concerned, the SP policy is not better than the Nominal policy (and in 3 of our four experiments – by the Robust policy as well). As about risk, here the SP policy clearly looses to the Robust one.

Table 1.2 Gain in the portfolio value.

Quantity	*Policy*	*Exp. # 1*	*Exp # 2*	*Exp. # 3*	*Exp. # 4*
Market Risk Index		0.339	0.377	0.393	0.401
Gain in the value, Max	*Cns*	1.576	1.851	1.667	1.749
	Rob	2.346	3.122	2.514	2.530
	Nom	2.405	12.447	14.332	22.927
	StP	4.912	15.254	9.150	15.423
Gain in the value, Min	*Cns*	1.173	1.197	1.118	1.076
	Rob	1.159	1.078	1.036	1.096
	Nom	1.109	0.175	0.218	0.163
	StP	0.491	0.269	0.110	0.100
Gain in the value, Mean	*Cns*	1.384	1.393	1.360	1.413
	Rob	1.659	1.652	1.553	1.651
	Nom	1.664	1.823	1.652	2.118
	StP	1.571	1.845	1.537	1.863
Gain in the value, StD	*Cns*	0.077	0.116	0.139	0.129
	Rob	0.193	0.259	0.231	0.233
	Nom	0.205	1.418	1.130	1.949
	StP	0.498	1.235	0.951	1.245
Loss probability, %	*Cns*	0.00	0.00	0.00	0.00
	Rob	0.00	0.00	0.00	0.00
	Nom	0.00	31.62	32.00	27.02
	StP	9.84	21.48	31.44	21.36
Big loss probability, %	*Cns*	0.00	0.00	0.00	0.00
	Rob	0.00	0.00	0.00	0.00
	Nom	0.00	20.70	20.70	17.90
	StP	4.12	11.26	20.16	13.34

We would conclude that the outlined experiments demonstrate high potential of the Robust Counterpart approach to the Portfolio problem, especially taking into account that the computational effort required to implement the policy is incomparably less than the one for Multistage Stochastic Programming approach.

References

[1] Ben-Tal, A., A. Nemirovski. (1996). "Robust solutions to uncertain linear programs," accepted to *Operations Research Letters*

[2] Ben-Tal, A., and A. Nemirovski. (1997). "Robust Convex Optimization," *Mathematics of Operations Research*, November 1998.

Table 1.3 Pair dominance of the policies

	Exp. # 1				*Exp. # 2*			
	Rob	*Nom*	*StP*	*Cns*	*Rob*	*Nom*	*StP*	*Cns*
Rob		48.9	62.2	95.4		58.0	55.9	88.9
Nom	51.1		62.8	94.5	42.0		45.2	50.7
StP	37.8	37.2		65.0	44.1	54.8		56.6
Cns	4.6	5.5	35.0		11.1	49.3	43.4	

	Exp. # 3				Exp. # 4			
	Rob	*Nom*	*StP*	*Cns*	*Rob*	*Nom*	*StP*	*Cns*
Rob		57.6	62.0	86.8		51.9	51.5	90.4
Nom	42.4		52.6	50.7	48.1		51.4	56.0
StP	38.0	47.4		46.9	48.5	48.6		59.9
Cns	13.2	49.3	53.2		9.6	43.9	40.1	

[3] Dantzig, G.B., and G. Infanger. “Multi-stage stochastic linear programs for portfolio optimization,” *Annals of Operations Research* v. 45 (1993), 59–76.

[4] Lemarechal, C., A. Nemirovski, and Yu. Nesterov. “New variants of bundle methods,” *Mathematical Programming Series B*, v. 69 (1995) No. 1, 111–148.