EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 3, No. 3, 2010, 572-592
ISSN 1307-5543 – www.ejpam.com

SPECIAL ISSUE ON
GRANGER ECONOMETRICS AND STATISTICAL MODELING
DEDICATED TO THE MEMORY OF PROF. SIR CLIVE W.J. GRANGER

Approximating Expectation Functionals for Financial Optimization

N.C.P. Edirisinghe

Department of Statistics, Operations and Management Science, College of Business
University of Tennessee, Knoxville, TN 37996, USA.

Abstract. Numerical evaluation of the expectation of a function of a random vector is often difficult because either the knowledge of the underlying probability distribution is not complete, or the probability space is continuous and each function evaluation is expensive. Such difficulties often arise in financial optimization where a risk measure is expressed as an expectation functional of random (asset) returns. Not only does the latter expectation depends on investment positions created in the underlying assets, but also it requires the solution of a mathematical program. First, the basic results from generalized moment problems are presented to establish tightness properties of approximations. Then, first and second moment approximations are presented for the expectation. These results are applied within a financial optimization problem to illustrate the efficiency of the approximations for determining optimal positions in a portfolio of the Standard and Poors 100 stocks.

2000 Mathematics Subject Classifications: 46N10, 44A60, 91G99

Key Words and Phrases: Approximating expectation, lower bounds, moment problems, portfolio optimization.

---

1. Introduction

Determining the expectation of a function of a random vector arises in many applications, including agriculture, economics, engineering, and finance. Numerical evaluation of such expectation is often difficult because either the knowledge of the underlying probability distribution is not complete, or the probability space is continuous and each function evaluation is expensive. For instance, the function evaluation may involve simulation or solution of a mathematical program. These difficulties are further compounded when the

Email address: chanaka@utk.edu (N.C.P. Edirisinghe)

underlying random variables are stochastically dependent and such an expectation functional is embedded in an (outer) decision optimization problem. Consequently, it may become necessary to evaluate the expectation functional as many times as needed in an iterative search designed to determine an optimal decision for the problem at hand. Multidimensional numerical integration as a computational strategy is prohibitively expensive even for a modest number of random variables.

To motivate, consider the following financial optimization problem. A portfolio manager wishes to determine a set of (risky) assets, such as stocks, for investment for a certain future period during which asset returns are uncertain. Having created the portfolio positions, and upon observing the realized asset returns, the portfolio manager may revise (or rebalance) her portfolio in order to control the expected deviation of portfolio value from a prescribed wealth target. In this context, the portfolio manager wishes to pick an initial portfolio that would hedge well against all possible realizations of random asset returns with respect to a desired portfolio return and any risk thereof. Such decision problems can be typically formulated as two-stage stochastic optimization models, e.g. see Edirisinghe [8], Kall [11], Wets [20], in which an optimized current allocation is made to maximize the portfolio return, less an appropriately measured risk associated with the allocation, subject to policy and other constraints. Let x denote the vector of (anticipatory) allocative decisions, which yields an (expected) profit function $c(x)$, associated with a given risk function $\psi(x)$. Consider the following model:

$$\begin{aligned} Z^* := \max_x \quad & c(x) - \lambda\, \psi(x) \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0, \end{aligned} \tag{1}$$

where $\lambda \geq 0$ is a risk-aversion parameter. Upon realization of the random vector ξ, the specific risk consequence $\phi(x, \xi)$ is determined under adaptive decisions y (such as portfolio rebalancing or financing missed wealth targets), and it is modeled by

$$\begin{aligned} \phi(x, \xi) := \min_y \quad & q(y) \\ \text{s.t.} \quad & Wy = h(\xi) - T(\xi)x \\ & y \geq 0. \end{aligned} \tag{2}$$

Note that a random matrix T transforms the decisions x into risk constraints to measure against the observed state of nature (vector) h. Then, the risk function is given by $\psi(x) = E[\phi(x, \xi)]$, where $E[.] \equiv E_{P^{tr}}[.]$ represents the mathematical expectation with respect to the true probability measure P^{tr} on Ξ, the domain of the random K-vector ξ. P^{tr} is assumed to be nondegenerate, and Ξ is assumed to be a convex subset in $\Re^K$. The matrices A $(\in \Re^{m_1 \times n_1})$ and W $(\in \Re^{m_2 \times n_2})$ are deterministic, while the matrix $T : \Xi \to \Re^{m_2 \times n_1}$ and the right-hand side $h : \Xi \to \Re^{m_2}$ are linear affine in the random vector ξ. The objective functions $c(.)$ and $q(.)$ are concave and convex, respectively.

The model in (1)-(2) admits several practical investment decision-making situations and a variety of risk descriptions. Also, note that depending on the measure of risk assessment, the optimization in (2) may become *fictitious*. As an example, consider the case when $W = I$, $h = 0$, $T = diag(\xi_1 - E[\xi_1], \ldots, \xi_K - E[\xi_K])$, and $q(y) = \left(\sum_{j=1}^{n_2} y_j\right)^2$. Then, $\phi(x, \xi) = x'(\xi - E[\xi])(\xi - E[\xi])'x$, where prime denotes transposition of a vector. Thus, it follows that the risk function is $\psi(x) = x'Mx$, where M is the variance-covariance matrix of ξ. In this special case, hence, the model in (1) achieves a mean-variance trade off in asset selection, see Markowitz [15].

In general, the optimization in (2) must be carried out and the expectation risk functional is difficult to evaluate. It is assumed that the feasible set X of (1) is nonempty, i.e.,

$$X := \{x \in \Re^{n_1} \; : \; Ax = b, \; x \geq 0\} \neq \emptyset. \tag{3}$$

Also, it is assumed that the risk-defining (second stage) problem (2) is feasible and bounded on Ξ for $x \in X$. Thus, ϕ is a (proper) convex function in ξ and x, separately. Since integration with respect to a probability measure preserves order, the risk function $\psi(x)$ is convex on X, see Wets [19].

It is generally accepted that econometric modeling of financial time series is of paramount importance for successful investment decision making in the stock markets. Financial time series are often co-integrated and this phenomenon is useful in constructing tests of stock market volatility. The implications of excess volatility resulting from Granger's co-integration approach is highly valuable to financial managers. Rather than focusing on time series modeling, this paper aims at controlling risks directly via portfolio optimization, recognizing that scenarios of the future cannot be known with certainty. This is done through incorporating risk functions, of type $\psi(x)$, that involve computing expectations in high dimensions. The concept of approximating expectation functions using general moment problems is presented in Section 2, where the case of using only first moments is also considered. In Section 3, approximations using both first and second moments (including covariances) are developed using the underling moment problem. The portfolio optimization model used for demonstrating the approximations is in Section 4. Section 5 reports the results from portfolio analysis while conclusions are in Section 6.

2. Approximating the Expectation

When ξ has a large number K of random components that are possibly stochastically dependent, solution of the model (1) is computationally tedious, and thus, approximating the expectation of the risks is a critical component in the numerical solution. The basic approach is to determine a lower bound on $\psi(x)$, say $\psi_L(x)$, that is easily computable, for example, without the need for numerical integration w.r.t. the probability measure P^{tr}. Then, (1) is approximated with ψ_L in place of ψ, i.e.,

$$Z_L := \max_{x \in X} \{c(x) - \lambda \; \psi_L(x)\}, \tag{4}$$

the solution of which yields an optimal solution x_L^* that can serve as a near-optimal decision. For this reason, ψ_L is required to be a high-quality approximation for ψ. In fact, the lower and upper bounds resulting from the use of $\psi_L(x)$, as given by,

$$Z_L \geq Z^* \geq c(x_L^*) - \lambda\,\psi(x_L^*) \tag{5}$$

may be used to verify the quality of the computed allocation x_L^*. The lower bound in (5) requires computing the expected risk function once; however, if even that is complicated, one may use an upper bounding function $\psi_U(x)$ on the expected risk function, so that the quality of the computed allocation may be measured with respect to the relative gap, given by $[\psi_U(x_L^*) - \psi_L(x_L^*)]/|Z_L|$. In this sense, the lower approximation is quite important and it is the central focus here. We will utilize first and second moment information of P^{tr} for this purpose, and then demonstrate the quality of these approximations in the financial portfolio application later.

2.1. Generalized Moment Problems

While there exist varying approaches for constructing approximations on $\psi(x)$, one that has received most attention is the bounds using a given set of moment information of the underlying probability distribution, termed the generalized moment problem (GMP). Since the GMP is based on an optimizing criterion, often the GMP is used as the yardstick of bound *tightness*. We shall review this important class of problems first.

For ξ being a random vector mapping a measurable space $(\Xi, \mathcal{B})$ to $\Re^K$, with $\mathcal{B}$ the Borel sigma field of events in $\Xi(\subset \Re^K)$, let $f_i : \Xi \to \Re$, for $i = 1, \ldots, N$, be finite measurable functions. Suppose the knowledge of the true probability distribution P^{tr} is available through the moments $\mu_i := E[f_i(\xi)]$ for $i = 1, \ldots, N$. Given the vector $\mu \in \Re^N$, a "tight" lower bound on the expectation $\psi(x) = E[\phi(x, \xi)]$ is determined by solving the general moment problem (GMP):

$$\varphi(x) := \inf_{P \in \mathcal{P}} \left\{ \int_\Xi \phi(x,\xi) P(d\xi) \;:\; \int_\Xi f_i(\xi) P(d\xi) = \mu_i,\; i = 1, \ldots, N \right\}, \tag{6}$$

where $\mathcal{P}$ denotes the set of all probability measures on $(\Xi, \mathcal{B})$. Since the true measure P^{tr} satisfies the conditions in (6), it follows that $\psi(x) \geq \varphi(x)$.

Conditions for the existence of a probability measure that solves the GMP is derived in Kemperman [14]. The set of probability measures feasible in GMP has been studied extensively. For instance, extreme points of the set of admissible probability measures of GMP are discrete measures involving no more than $N+1$ atoms, see Karr [13]. Hence, the optimizing measure in (6) is a discrete measure with cardinality at most $N+1$. However, determination of this discrete measure is difficult, and it may generally involve nonconvex optimization, see Birge and Wets [2]. Instead, one may attempt to solve the semi-infinite dual problem of the GMP, given by

$$\varphi^*(x) := \sup_{\pi \in \Re^{N+1}} \left\{ \pi_0 + \sum_{i=1}^N \mu_i \pi_i \;:\; \pi_0 + \sum_{i=1}^N f_i(\xi) \pi_i \leq \phi(x,\xi),\; \xi \in \Xi \right\}. \tag{7}$$

Kall [12] investigates this duality relationship in the context of general moment problems. For a more general treatment of duality theory for semi-infinite linear programs, see Glashoff and Gustafson [9]. Weak duality holds for (6) and (7), i.e.,

Proposition 1. $\varphi(x) \geq \varphi^*(x)$.

Proof. When (6) is infeasible, we set $\varphi(x) = +\infty$ and if (7) is infeasible, we set $\varphi^*(x) = -\infty$, and then the proposition is held trivially. Otherwise, for some probability measure P feasible in (6) and for any vector $(\pi_0, \pi_1, \ldots, \pi_N)$ feasible in (7),

$$\begin{aligned} \pi_0 + \sum_{i=1}^{N} \mu_i \pi_i &= \int_\Xi P(d\xi)\pi_0 + \sum_{i=1}^{N}\Big(\int_\Xi f_i(\xi)P(d\xi)\Big)\pi_i \\ &= \int_\Xi [\pi_0 + \sum_{i=1}^{N} f_i(\xi)\pi_i] P(d\xi) \\ &\leq \int_\Xi \phi(x,\xi)P(d\xi). \quad \blacksquare \end{aligned} \tag{8}$$

Hence, any feasible solution of (7) generates a lower bound on $\psi(x)$; however, the intent is to generate the best lower bound by solving (7). Moreover, if $\varphi(x) = \varphi^*(x)$, then the computed lower bound is declared *tight* w.r.t to the GMP. The latter strong duality can be assured under mild conditions, as given below.

Proposition 2. *$\varphi(x) = \varphi^*(x)$ holds if*
(i) f_i, $i = 1, \ldots, N$, are continuous, and
(ii) Ξ is compact.

Proof. See Glashoff and Gustafson [9, p.79], Kall [12, Theorem 4]. ■

However, when Ξ is unbounded, a certain interior-type condition is required to ensure strong duality. Define $\mathcal{M}$ as the convex hull of the moment conditions $f_i(\xi)$, $i = 1, \ldots, N$ for $\xi \in \Xi$, i.e.,

$$\mathcal{M} := co\left\{(f_1(\xi), \ldots, f_N(\xi)) \;:\; \xi \in \Xi\right\}. \tag{9}$$

Proposition 3. *The semi-infinite dual (7) is solvable and $\varphi(x) = \varphi^*(x)$ holds if*
(i) (7) is feasible, and
(ii) the N-dimensional point $\mu \in int(\mathcal{M})$, the interior of the set $\mathcal{M}$.

Proof. See Glashoff and Gustafson [9, p.79], Kall [12, Theorem 4]. ■

While existence of solutions and strong duality are claimed as above, an explicit expression of the solution is not easy for arbitrary moment functions f_i. Also, note that the above results do not require any functional properties of ϕ. However, in our case, using the convexity of ϕ in ξ, the semi-infinite dual optimal solution can be expressed in closed-form under first moment conditions.

2.2. First moment approximation

Consider the GMP in (6) using only the first moments $\bar{\xi} = E[\xi]$, i.e.,

$$\varphi_1(x) = \inf_{P \in \mathcal{P}} \left\{ \int_\Xi \phi(x,\xi) P(d\xi) \;\; : \;\; \int_\Xi \xi P(d\xi) = \bar{\xi} \right\}. \tag{10}$$

Defining the graph

$$\mathcal{G}(x) := \{(\xi, \phi(x,\xi)) \; : \; \xi \in \Xi\}, \tag{11}$$

let $\mathcal{Z}(x) = co\,\mathcal{G}(x)$, the convex hull of $\mathcal{G}$, which is closed since ϕ is continuous in ξ. Then (10) is equivalent to

$$\varphi_1(x) = \inf_z \left\{ z_{K+1} \in \Re \;\; : \;\; (z_1, \ldots, z_{K+1}) \in \mathcal{Z}(x), \; z_k = \bar{\xi}_k, \; k = 1, \ldots, K \right\}. \tag{12}$$

Define the *epigraph* of $\phi(x,\xi)$ by the set $\mathcal{E}\phi(x)$, i.e.,

$$\mathcal{E}\phi(x) := \{(\xi_1, \ldots, \xi_K, z_{K+1}) \; : \; z_{K+1} \geq \phi(x,\xi), \; \xi \in \Xi\}. \tag{13}$$

Proposition 4. *$\mathcal{Z}(x) \subseteq \mathcal{E}\phi(x)$.*

Proof. Consider an arbitrary point $\hat{z} \in \mathcal{Z}(x)$. Since $\mathcal{Z}(x) = co\,\mathcal{G}(x)$, there exist a set of I points $(\xi^i, \phi(x,\xi^i))$, where $\xi^i \in \Xi$, and nonnegative multipliers λ_i, $i = 1, \ldots, I$, such that

$$\hat{z}_k = \sum_{i=1}^I \lambda_i \xi_k^i, \; k = 1, \ldots, K, \; \hat{z}_{K+1} = \sum_{i=1}^I \lambda_i \phi(x,\xi^i), \; \sum_{i=1}^I \lambda_i = 1. \tag{14}$$

Since ϕ is convex in ξ, (14) implies $\sum_{i=1}^I \lambda_i \phi(x,\xi^i) \geq \phi(x, \hat{z}_1, \ldots, \hat{z}_K)$, and thus, it follows that $\hat{z}_{K+1} \geq \phi(x, \hat{z}_1, \ldots, \hat{z}_K)$. Moreover, $\xi^i \in \Xi$ implies that $(\hat{z}_1, \ldots, \hat{z}_K) \in \Xi$ due to the convexity of Ξ. Thus, $\hat{z} \in \mathcal{E}\phi(x)$, completing the proof. ■

Proposition 5. *$\varphi_1(x) = \phi(x, \bar{\xi})$.*

Proof. Since $\mathcal{Z}(x) \subseteq \mathcal{E}\phi(x)$ holds due to Proposition 4, from (12),

$$\begin{aligned}
\varphi_1(x) \;\; &\geq \;\; \inf_z \left\{ z_{K+1} \in \Re \;\; : \;\; (z_1, \ldots, z_{K+1}) \in \mathcal{E}\phi(x), \; z_k = \bar{\xi}_k, \; k = 1, \ldots, K \right\} \\
&= \;\; \inf_z \left\{ z_{K+1} \in \Re \;\; : \;\; z_{K+1} \geq \phi(x, \bar{\xi}) \right\} \\
&= \;\; \phi(x, \bar{\xi}).
\end{aligned}$$

On the other hand, $(\bar{\xi}, \phi(x,\bar{\xi}) \in \mathcal{G}(x)$, and since $\mathcal{G}(x) \subseteq \mathcal{Z}(x)$, $(\bar{\xi}_1, \ldots, \xi_K, \phi(x,\bar{\xi}))$ is feasible in (12). Thus, $\varphi_1(x) \leq \phi(x,\bar{\xi})$ holds, which completes the proof. ■

Observe that the value of the first-moment GMP is thus obtained without any assumption on conditions required for strong duality. Note that $\phi(x,\bar{\xi})$ is indeed the well-known Jensen [10] lower bound on the expectation of a convex function. On the other hand,

$\bar{\xi} \in int(\Xi)$ holds since P^{tr} is nondegenerate, and due to finiteness of (10), Proposition 3 ensures that the semi-infinite dual satisfies

$$\phi(x,\bar{\xi}) = \sup_{\pi \in \Re^{N+1}} \left\{ \pi_0 + \sum_{k=1}^{K} \bar{\xi}_k \pi_k \ : \ \pi_0 + \sum_{k=1}^{K} \xi_k \pi_k \leq \phi(x,\xi), \ \xi \in \Xi \right\}. \tag{15}$$

The solution of the dual is then determined by the 'supporting hyperplane' of the convex function $\phi(x,\xi)$ at $\xi = \bar{\xi}$. Can the first moment lower bound be improved under additional information on the random vector ξ? For instance, if ξ has uncorrelated components, would the convexity information on ϕ allows one to develop a stronger bound? To answer this question, first, note that the components of ξ are mutually uncorrelated (under P^{tr}) if and only if

$$E\left[\prod_{j \in \Lambda} \xi_j\right] = \prod_{j \in \Lambda} \bar{\xi}_j, \ \forall \Lambda \in \mathbb{B}, \tag{16}$$

where $\mathbb{B}$ is the set of all subsets of $\{1,\ldots,K\}$ with cardinality 2. Then, the (tight) GMP lower bound under first moments and the uncorrelated information is formulated as

$$\varphi_{1u}(x) = \inf_{P \in \mathcal{P}} \left\{ \int_{\Xi} \phi(x,\xi)P(d\xi) \ : \ \int_{\Xi} \xi P(d\xi) = \bar{\xi}, \ \int_{\Xi} (\prod_{j \in \Lambda} \xi_j) P(d\xi) = (\prod_{j \in \Lambda} \bar{\xi}_j), \forall \Lambda \in B \right\}. \tag{17}$$

Proposition 6. *Jensen's lower bound remains tight even under uncorrelated information, i.e.,* $\varphi_{1u}(x) = \phi(x,\bar{\xi})$.

Proof. Since the degenerate distribution with probability mass at $\bar{\xi}$ is feasible in (17), $\varphi_{1u}(x) \leq \phi(x,\bar{\xi})$. On the other hand, (17) is obtained by adding more constraints to (10), and thus, $\varphi_{1u}(x) \geq \varphi_1(x) = \phi(x,\bar{\xi})$, which completes the proof. ∎

3. Second Moment Approximation

While the 'uncorrelated' knowledge does not improve the first moment lower bound, is it possible to derive an improved lower bound when all variance-covariance information of ξ is available? Under the mean vector $\bar{\xi}$ and the covariance σ_{kl} between ξ_k and ξ_l, for $k,l = 1,\ldots,K$, a tight lower bound is determined by solving the GMP given by

$$\varphi_2(x) = \inf_{P} \left\{ \int_{\Xi} \phi(x,\xi)P(d\xi) \ : \ P \in \tilde{\mathcal{P}} \right\}, \tag{18}$$

where the set of probability measures $\tilde{\mathcal{P}}$ is characterized by

$$\tilde{\mathcal{P}} := \left\{ P \ : \ \int_{\Xi} \xi P(d\xi) = \bar{\xi}, \ \int_{\Xi} \xi_k \xi_l P(d\xi) = m_{kl}, \ k,l = 1,\ldots,K, \ k \geq l \right\}, \tag{19}$$

where $m_{kl} = \sigma_{kl} + \bar{\xi}_k\bar{\xi}_l$. Clearly, $\varphi_2(x) \geq \phi(x, \bar{\xi})$ with the latter equality surely being held when $\sigma_{kl} = 0$ for $k \neq l$, see Proposition 6. The semi-infinite dual of (18) is given by

$$\varphi_2(x) = \sup_{\pi \in \Pi} \left\{ \pi_0 + \pi_1 \bar{\xi} + \sum_{k,l=1, k \geq l}^{K} m_{kl}\pi_{kl} \right\} \tag{20}$$

where the dual feasible set

$$\Pi := \left\{ \pi \ : \ \pi_0 + \pi_1 \xi + \sum_{k,l=1, k \geq l}^{K} \xi_k \xi_l \pi_{kl} \ \leq \ \phi(x, \xi), \ \xi \in \Xi \right\}, \tag{21}$$

provided either Ξ is compact or

$$\begin{aligned} &(\bar{\xi}_1, \ldots, \bar{\xi}_K, m_{11}, \ldots, m_{1K}, m_{22}, \ldots, m_{2K}, \ldots, m_{K-1,K}, m_{KK}) \\ &\in \ \ int\, co \left\{ \xi_1, \ldots, \xi_K, (\xi_1)^2, \ldots, \xi_{1K}, (\xi_2)^2, \ldots, \xi_2\xi_K, \ldots, \xi_{K-1}\xi_K, (\xi_K)^2 \ : \ \xi \in \Xi \right\}. \end{aligned} \tag{22}$$

Observe that when $K = 1$ and $\Xi = (-\infty, +\infty)$, the interior condition in (22) is certainly satisfied when $E[(\xi_1)^2] > (\bar{\xi})^2$, i.e., P^{tr} is a nondegenerate distribution. However, we shall assume for (18) that Ξ is compact, as is followed in the remainder of the paper.

Solution of (18) or (20) remains an *open* research problem. The difficulty lies in that a lower bounding quadratic function on $\phi(x, .)$ over Ξ must be determined toward solving (20), which is an onerous task for general convex functions ϕ and arbitrary convex sets Ξ.

Under compact domains, let Ξ be a K-dimensional simplex (if not, the domain can be embedded in a simplex). The focus here is to develop bounds on $\varphi_2(x)$ using a lower bounding polyhedral function on $\phi(x, \xi)$. Toward this, let the vertices of the simplex Ξ be denoted by $u^i \in \Re^K$, $i = 1, \ldots, K+1$, and define the inverse of the vertex matrix by

$$V := \begin{bmatrix} u_1^1 & \cdots & \cdots & u_1^{K+1} \\ \vdots & \vdots & \vdots & \vdots \\ u_K^1 & \cdots & \cdots & u_K^{K+1} \\ 1 & \cdots & \cdots & 1 \end{bmatrix}^{-1} \tag{23}$$

and its j^{th} row by $v^j \equiv (v_1^j, \ldots, v_K^j, v_{K+1}^j)$. It is straightforward to show that:

Proposition 7. *Let the linear (measurable) function $\lambda_j(\xi)$ be defined by*

$$\lambda_j(\xi) := v_1^j \xi_1 + \cdots + v_K^j \xi_K + v_{K+1}^j. \tag{24}$$

Then, $\xi \in \Xi$ if and only if

$$\sum_{j=1}^{K+1} \lambda_j(\xi) = 1 \ \ and \ \ \lambda_j(\xi) \geq 0, \ j = 1, \ldots, K+1. \tag{25}$$

Consider the following construction. Multiply the constraints of (21) by nonnegative $\lambda_i(\xi)$, for each $i = 1, \ldots, K+1$, which yields

$$\lambda_i(\xi)\pi_0 + \pi_1[\lambda_i(\xi)\xi] + \sum_{k,l=1,k\geq l}^{K} [\lambda_i(\xi)\xi_k\xi_l]\pi_{kl} \;\leq\; \lambda_i(\xi)\phi(x,\xi),\; \forall \xi \in \Xi,\; i = 1, \ldots, K+1. \tag{26}$$

Upon taking the expectation of (26) with respect to any probability measure $P \in \tilde{\mathcal{P}}$,

$$E_P[\lambda_i(\xi)]\pi_0 + \pi_1 E_P[\lambda_i(\xi)\xi] + \sum_{k,l=1,k\geq l}^{K} E_P[\lambda_i(\xi)\xi_k\xi_l]\pi_{kl} \;\leq\; E_P[\lambda_i(\xi)\phi(x,\xi)] \tag{27}$$

holds for $i = 1, \ldots, K+1$. Summing the latter inequalities over all i, thus, every feasible solution $\pi \in \Pi$ must satisfy

$$\pi_0 + \pi_1 \sum_{i=1}^{K+1} t^i + \sum_{k,l=1,k\geq l}^{K} \left(\sum_{i=1}^{K+1} r^i_{kl} \right) \pi_{kl} \;\leq\; \sum_{i=1}^{K+1} E_P[\lambda_i(\xi)\phi(x,\xi)], \tag{28}$$

since $\sum_{i=1}^{K+1} E_P[\lambda_i(\xi)] = 1$ and by defining

$$t^i_k := E_P[\lambda_i(\xi)\xi_k],\; k = 1, \ldots, K \tag{29}$$

$$r^i_{kl} := E_P[\lambda_i(\xi)\xi_k\xi_l],\; k = 1, \ldots, K,\; l = k, \ldots, K. \tag{30}$$

Note that for any $P \in \tilde{\mathcal{P}}$, t^i_k and r^i_{kl} are (unique) constants. Then, referring to (25), it follows that $\sum_i t^i_k = \bar{\xi}_k$ and $\sum_i r^i_{kl} = m_{kl}$. Thus, every feasible solution $\pi \in \Pi$ must satisfy

$$\pi_0 + \pi_1\bar{\xi} + \sum_{k,l=1,k\geq l}^{K} m_{kl}\pi_{kl} \;\leq\; \sum_{i=1}^{K+1} E_P[\lambda_i(\xi)\phi(x,\xi)]. \tag{31}$$

Therefore, adding (31) to (20) is not a restriction; however, a lower bound on the right-hand side of (31) can be used to develop a restriction on (21).

Let $\ell^i(x,\xi)$, $i = 1, \ldots, K+1$, be a family of lower bounding linear functions on the (proper) convex function $\phi(x,\xi)$ over Ξ, i.e.,

$$\ell^i(x,\xi) \leq \phi(x,\xi),\; w.p.1\,, \quad \forall i = 1, \ldots, K+1. \tag{32}$$

Construct the polyhedral lower bounding function $g(x,\xi)$ on $\phi(x,\xi)$, where for $\xi \in \Xi$,

$$\phi(x,\xi) \geq g(x,\xi) := \max_{i=1,\ldots,K+1} \ell^i(x,\xi). \tag{33}$$

Next, define the points $\tilde{\xi}^i \in \Re^K$, $i = 1, \ldots, K+1$, as follows:

$$\tilde{\xi}^i_k := \frac{1}{\rho_i}\left[v^i_1 m_{k1} + \cdots + v^i_K m_{kK} + v^i_{K+1}\bar{\xi}_k \right],\;\; k = 1, \ldots, K, \tag{34}$$

$$\text{where}\;\; \rho_i := v^i_1\bar{\xi}_1 + \cdots + v^i_K\bar{\xi}_K + v^i_{K+1}\,,\;\; i = 1, \ldots, K+1. \tag{35}$$

Proposition 8. $\tilde{\xi}^i \in \Xi$ *for all* $i = 1, \ldots, K+1$.

Proof. For $\xi \in \Xi$, multiplying the expressions in (25) by nonnegative $\lambda_i(\xi)$, for some vertex index i,

$$\sum_{j=1}^{K+1} \lambda_j(\xi)\lambda_i(\xi) = \lambda_i(\xi) \text{ and } \lambda_j(\xi)\lambda_i(\xi) \geq 0.$$

Noting the linearity of $\lambda_j(\xi)$ in (24),

$$\sum_{j=1}^{K+1} [v_1^j \xi_1 + \cdots + v_K^j \xi_K + v_{K+1}^j]\lambda_i(\xi) = \lambda_i(\xi) \text{ and } [v_1^j \xi_1 + \cdots + v_K^j \xi_K + v_{K+1}^j]\lambda_i(\xi) \geq 0.$$

Upon taking the expectation of the above w.r.t. any $P \in \tilde{\mathcal{P}}$, and noting the definitions in (29) and (35),

$$\sum_{j=1}^{K+1} [v_1^j t_1^i + \cdots + v_K^j t_K^i + v_{K+1}^j \rho_i] = \rho_i \text{ and } v_1^j t_1^i + \cdots + v_K^j t_K^i + v_{K+1}^j \rho_i \geq 0.$$

Dividing by ρ_i and noting the definition in (34),

$$\sum_{j=1}^{K+1} [v_1^j \tilde{\xi}_1^i + \cdots + v_K^j \tilde{\xi}_K^i + v_{K+1}^j] = 1 \text{ and } v_1^j \tilde{\xi}_1^i + \cdots + v_K^j \tilde{\xi}_K^i + v_{K+1}^j \geq 0.$$

Since $\lambda_j(\tilde{\xi}^i) = v_1^j \tilde{\xi}_1^i + \cdots + v_K^j \tilde{\xi}_K^i + v_{K+1}^j$, it follows that

$$\sum_{j=1}^{K+1} \lambda_j(\tilde{\xi}^i) = 1 \text{ and } \lambda_j(\tilde{\xi}^i) \geq 0,$$

which implies due to Proposition 7 that $\tilde{\xi}^i \in \Xi$. ■

Proposition 9. *For any* $P \in \tilde{\mathcal{P}}$,

$$\sum_{i=1}^{K+1} E_P[\lambda_i(\xi)\phi(x,\xi)] \geq \sum_{i=1}^{K+1} E_P[\lambda_i(\xi)g(x,\xi)] \geq \sum_{i=1}^{K+1} \rho_i \, g(x,\tilde{\xi}^i). \tag{36}$$

Proof. The first inequality follows for any probability measure P on Ξ because $\phi \geq g$ and $\lambda_i \geq 0$, $\forall i$, on Ξ. For the second inequality, noting the definition of g in (33), and denoting the linear function $\ell^j(x,\xi)$ by $\alpha^j \xi + \theta_j$ for some row vector $\alpha^j \in \Re^K$ and scalar $\theta_j \in \Re$,

$$E_P\left[\lambda_i(\xi)g(x,\xi)\right] = E_P\left[\lambda_i(\xi) \max_{j=1,\ldots,K+1} \ell^j(x,\xi)\right]$$

$$
\begin{aligned}
&= E_P\left[\max_{j=1,\dots,K+1} \lambda_i(\xi)\ell^j(x,\xi)\right] \\
&\geq \max_{j=1,\dots,K+1} \left\{E_P\left[\lambda_i(\xi)\ell^j(x,\xi)\right]\right\} \\
&= \max_{j=1,\dots,K+1} \left\{\alpha^j\, E_P[\lambda_i(\xi)\xi] + \theta_j\, E_P[\lambda_i(\xi)]\right\} \\
&= \max_{j=1,\dots,K+1} \left\{\alpha^j\, t^i + \theta_j\, \rho_i\right\} \;=\; \rho_i\left[\max_{j=1,\dots,K+1}\left\{\alpha^j\left(\frac{t^i}{\rho_i}\right) + \theta_j\right\}\right] \\
&= \rho_i\left[\max_{j=1,\dots,K+1} \ell^j(x,\tilde{\xi}^i)\right] \\
&= \rho_i\, g(x,\tilde{\xi}^i). \quad \blacksquare
\end{aligned}
$$

Therefore, it is straightforward to see that

$$
\begin{aligned}
\varphi_2(x) \geq \varphi_2^R(x) := \quad & \sup_{\pi\in\Pi} \quad \pi_0 + \pi_1\bar{\xi} + \sum_{k,l=1,k\geq l}^{K} m_{kl}\pi_{kl} \\
& \text{s.t.} \quad \pi_0 + \pi_1\bar{\xi} + \sum_{k,l=1,k\geq l}^{K} m_{kl}\pi_{kl} \;\leq\; \sum_{i=1}^{K+1} \rho_i\, g(x,\tilde{\xi}^i).
\end{aligned} \tag{37}
$$

It is important to note that if an optimal solution $\pi^* \in \Pi$ of (20) is feasible in the optimization problem (37), then indeed $\varphi_2(x) = \varphi_2^R(x)$ follows for the chosen lower bounding linear functions ℓ^j. Generally, an arbitrary simplicial lower approximating function on ϕ cannot be expected to solve the moment problem in (18), implying $\varphi_2(x) > \varphi_2^R(x)$.

Proposition 10. *For polyhedral (simplicial) function g, lower approximating ϕ, suppose $\varphi_2(x) > \varphi_2^R(x)$ holds. Then,*

$$
\varphi_2^R(x) = \sum_{i=1}^{K+1} \rho_i\, g(x,\tilde{\xi}^i). \tag{38}
$$

Proof. Considering the semi-infinite program in (37), strong duality must hold since Ξ is compact and ϕ is continuous, see for instance, Anderson and Nash [1]. Denoting $F_g(x) := \sum_{i=1}^{K+1} \rho_i\, g(x,\tilde{\xi}^i)$, therefore,

$$
\begin{aligned}
\varphi_2^R(x) = \quad & \inf_{\sigma,Q} \quad \sigma\, F_g(x) + \int_\Xi \phi(x,\xi)Q(d\xi) \\
& \text{s.t.} \quad \sigma + \int_\Xi Q(d\xi) = 1 \\
& \qquad \sigma\,\bar{\xi} + \int_\Xi \xi Q(d\xi) = \bar{\xi} \\
& \qquad \sigma\, m_{kl} + \int_\Xi \xi_k\xi_l Q(d\xi) = m_{kl},\ k,l = 1,\dots,K,\ k\geq l \\
& \qquad \sigma \geq 0,\ Q(.) \geq 0.
\end{aligned} \tag{39}
$$

Since, $\sigma \leq 1$ must hold for feasibility of (39), it follows that

$$\begin{aligned} \varphi_2^R(x) &= \inf_{\sigma,P} \quad \sigma\, F_g(x) + (1-\sigma)\int_\Xi \phi(x,\xi)P(d\xi) && (40)\\ & \quad \text{s.t.} \quad 0 \leq \sigma \leq 1, \quad P \in \tilde{\mathcal{P}}. \\ &= \inf_{\sigma\in[0,1]} \quad \sigma\, F_g(x) + (1-\sigma)\,\varphi_2(x) && (41)\\ &= \min\left\{F_g(x)\ ,\ \varphi_2(x)\right\}. \end{aligned}$$

Since $\varphi_2(x) > \varphi_2^R(x) = \min\{F_g(x)\ ,\ \varphi_2(x)\}$, we have $\varphi_2^R(x) = F_g(x)$. ■

So far, the lower approximating polyhedral (simplicial) function g is chosen quite arbitrarily, and accordingly, the above lower bound $F_g(x) = \varphi_2^R(x)$ may become arbitrarily weaker compared to $\varphi_2(x)$. Moreover, this lower bound is computationally tedious since g can be difficult to compute. Is it possible to *lift* g such that it is lower-approximating to ϕ over Ξ and it is relatively-easy to compute? Indeed, it is possible if the linear functions ℓ_i are chosen as *supporting* hyperplanes to the convex function ϕ. For this, for each $i = 1, \ldots, K+1$, define the supporting hyperplanes to ϕ precisely at $\tilde{\xi}^i \in \Xi$ by

$$\ell^i(x,\xi) := \phi(x,\tilde{\xi}^i) + \partial\phi(x,\tilde{\xi}^i)'(\xi - \tilde{\xi}^i),$$

where $\partial\phi$ is a subgradient to ϕ at $\tilde{\xi}^i$. Then, by convexity of ϕ in ξ, $\ell^i(x,\xi) \leq \phi(x,\xi)$ on Ξ and $\ell^i(x,\tilde{\xi}^i) = \phi(x,\tilde{\xi}^i)$. Hence,

$$\begin{aligned} g(x,\tilde{\xi}^j) &= \max_{i=1,\ldots,K+1} \left\{\phi(x,\tilde{\xi}^i) + \partial\phi(x,\tilde{\xi}^i)'(\tilde{\xi}^j - \tilde{\xi}^i)\right\} && (42)\\ &= \phi(x,\tilde{\xi}^j) \end{aligned}$$

since the right hand maximum in (42) is attained with $i = j$. This leads to the following main result:

Proposition 11.

$$\psi(x) \geq \varphi_2(x) \geq \varphi_2^R(x) = \sum_{i=1}^{K+1} \rho_i\, \phi(x,\tilde{\xi}^i) =: \psi_L(x). \tag{43}$$

Proof. Let an optimal probability measure solving the GMP in (18) be denoted by $P^* \in \tilde{\mathcal{P}}$, i.e., $\varphi_2(x) = E_{P^*}[\phi(x,\xi)]$. Then, applying Theorem 4 in Edirisinghe [4] on the convex function ϕ under measure P^*, $\varphi_2(x) \geq \psi_L(x)$ follows. Then, from the proof of Proposition 10, $\varphi_2^R(x) = \min\{\psi_L(x)\ ,\ \varphi_2(x)\} = \psi_L(x)$. ■

Therefore, the *lifted* simplicial lower 'supporting' approximation on ϕ at $\tilde{\xi}^i$, $i = 1, \ldots, K+1$, results in solving the restricted semi-infinite dual in (37). While the result in Edirisinghe [4, Theorem 4] was derived under a quite complicated procedure without any reference to a GMP, the preceding analysis reveals that it indeed is the solution to the restricted dual of the GMP. That is, there exists a feasible quadratic function lower approximating ϕ, see (21), with an expected value under P^{tr} exactly equal to the expected value of the supporting simplicial (polyhedral) function on ϕ under the derived probability measure $\{(\tilde{\xi}^i, \rho_i) \ : \ i = 1, \ldots, K+1\}$.

This is the first instance the latter relation to the underlying mean-covariance moment problem is investigated. Furthermore, the derivation in this paper underscores the fact that a simplicial polyhedral lower approximation is likely to fail in solving the moment problem (18) as evident from the inequality $\varphi_2(x) \geq \varphi_2^R(x)$. Hence, it remains an open question at this point whether $\varphi_2(x) = \varphi_2^R(x)$ can be attained with the supporting simplicial function constructed at $\tilde{\xi}^i$, $i = 1, \ldots, K+1$, for a general convex function ϕ.

4. Financial Optimization Model

The usefulness and quality of the mean and variance-covariance approximation of expectation functionals are investigated in the context of a financial optimization model. As presented in Section 1, the expectation functional is the risk component $\psi(x)$ of the investment portfolio problem (1), where specific risks are computed by (2) adapted to realizations ξ. We will specialize this model in this section to the following case.

Consider K risky assets for portfolio allocation of a total budget B^0, given the current (initial) \$ investment in each asset by x_j^0, $j = 1, \ldots, K$. An optimal allocation is desired for a future period (of length τ days) and it is denoted by x_j, $j = 1, \ldots, K$. Such a portfolio revision incurs transactions and slippage costs and this loss function is denoted by $L(.)$. As proposed in Edirisinghe [5], to account for possibly investing in stocks with relatively *light* trading volume, a loss function that is inversely proportional to the trade size, z_j in asset j, is desired. Adopting from the latter reference, we use the quadratic loss function, for (market calibrated) constants a_{1j} and a_{2j},

$$L_j(z_j) = a_{1j} z_j + a_{2j} \frac{(z_j)^2}{vol_j},$$

where vol_j is the (estimated) market total daily trading dollar volume and z_j is the dollar volume of shares purchased/sold in asset j, i.e., $z_j = |x_j - x_j^0|$. We will allow 'going long' or 'selling short' in each asset, implying $x_j \in \Re$, and thus, such portfolios are likely to encounter greater risk relative to the overall *market.* One way to circumvent the market dependent risk is to require that the portfolio's correlation with the market be controlled in the sense that portfolio *beta* is within acceptable levels, see Edirisinghe [5], referred to as portfolio's *degree of market neutrality* (DMN). Let the "beta" of asset j be given by β_j, which measures the correlated-dependence of asset return with market return. Denoting the market (τ-day) random return by ζ_M, and that of the asset by ξ_j, we have $\beta_j = Cov(\xi_j, \zeta_M)/[\sigma_{jj} Var(\zeta_M)]^{0.5}$, where σ_{jj} is the variance of asset return ξ_j. We use Standard and Poors 500 index as the proxy for the market. Specifying the 'portfolio beta' to be within $\pm 100\nu\%$, the DMN constraint is given by, $\sum_{j=1}^K \beta_j x_j \in [-\nu\, B^0 \ , \ +\nu\, B^0]$, and thus, when $\nu \approx 0$, the optimal portfolio is required to be nearly *beta-neutral.*

The investor desires to maximize portfolio expected return over the τ-day period subject to satisfying above concerns, as well as asset positions being controlled within given bounds, denoted by $x_j^{\min} \leq x_j \leq x_j^{\max}$. The number of shares purchased/sold must be an integer and the share price of asset j (at the time of forming the portfolio) is denoted by p_j. Furthermore, an appropriate risk function $\psi(x)$ must be incorporated to ensure that the inherent risk in asset returns (due to their correlation with each other) is controlled in an *efficient* manner. Consider the portfolio optimization model:

$$\begin{aligned} Z^* := \max_x \quad & \bar{\xi}'x - \textstyle\sum_{j=1}^K L_j(z_j) - \lambda\, \psi(x) \\ \text{s.t.} \quad & \mathbf{1}'x \leq B^0 \\ & -x + x^0 \leq z \leq x - x^0 \\ & -\nu\, B^0 \leq \beta' x \leq +\nu\, B^0 \\ & x^{\min} \leq x \leq x^{\max}, \quad |\tfrac{x_j}{p_j}| : \text{integer},\ j = 1, \dots, K. \end{aligned} \tag{44}$$

Portfolio efficiency is varied using the risk-aversion parameter $\lambda \geq 0$. Choice of $\psi(x)$ has been a topic of considerable debate and significant progresses have been made in this regard. Edirisinghe [5] considers $\psi(x)$ in the context of static or dynamic risk control. Artzner *et al.* [17] describe the principles of measuring risk attitude, where the authors propose the concept of coherent risk measures that allows numerical expression of risk attitude. Another interesting method of risk measurement is to use the conditional value-at-risk (CVaR), see Rockafellar and Uryasev [18] and Ogryczak and Ruszczynski [16]. In particular, risk measures based on mean and CVaR are coherent.

The focus here is not not engage in a detailed discussion of the pros and cons of various risk measures, but to pick a risk measure that is consistent with coherency and rationality. Setting $\psi(x) = x'Mx$, where M is the variance-covariance matrix of ξ, also see Section 1, i.e., Markowitz model, can lead to inconsistent portfolio choices that contradicts rationality of an investor. The concept of 'rational' risk measures proposed by Bychkov and Edirisinghe [3] circumvents such inconsistencies by requiring the risk functions to satisfy properties of coherence and first-order stochastic dominance (FSD). The latter reference shows that even the well-known mean/semi-variance trade off, where one sets $\psi(x) = E[(\max\{\bar{\xi}'x - \xi'x\,,\, 0\})^2]$, is not consistent with FSD. However, setting $\psi(x) = E[(\max\{C - \xi'x\,,\, 0\})^\gamma]$, for $\gamma \geq 1$, yields a rational risk measure that is FSD consistent and convex, where C is a constant return target for the portfolio. In the sequel, we employ this risk function with $\gamma = 2$. Consequently, $\psi(x) = E[\phi(x, \xi)]$, where

$$\begin{aligned} \phi(x, \xi) := \min_y \quad & (y)^2 \\ \text{s.t.} \quad & y \geq C - \xi'x \\ & y \geq 0. \end{aligned} \tag{45}$$

4.1. Data and parameters

Asset (stock) returns are quite well-known to have non-symmetric distributions. More importantly, stock return distributions are shown to have 'fatter' tails than the normal distributions would imply, see Ziemba [21]. Consequently, normal distribution assumption (or many other theoretical distributions) on ξ often leads to portfolios that perform poorly in practice. The main reason is that *rare* events do occur much more frequently than most theoretical distributions correspond to.

In this paper, we employ a historical daily return sample of T days for the K assets ($T >> K$), and only the first and second-order cross moments are estimated from the sample. Without making further distributional assumptions, these moment estimates are employed in the context of the second moment lower bound discussed in Section 3 to approximate the risk function $\psi(x)$. This approximation then constructs precisely $K+1$ stock return vectors that are located relative to both the mean and variance-covariance information, as well as based on the extremeness of returns observed during the historical T periods.

Let the return sample be denoted by $\xi^1, \ldots, \xi^T \in \Re^K$. A simplicial support Ξ is needed such that $co\{\xi^1, \ldots, \xi^T\} \subseteq \Xi$, see Figure 1. The problem of determining a compact multi-dimensional simplex covering multivariate points is addressed in Edirisinghe [6]; also see Edirisinghe and You [7]. For Ξ as determined above, denoting the inverse of vertex matrix by V, the second moment approximated scenario sample, given by $\tilde{\xi}^i$, $i = 1, \ldots, K+1$, is computed according to equations (34)-(35), where $\bar{\xi}_k = \frac{1}{T}\sum_{t=1}^{T} \xi_k^t$ and $m_{kl} = \frac{1}{T}\sum_{t=1}^{T} \xi_k^t \xi_l^t$, $\forall k, l$. For the illustration here, we use the 10-year historical period from January 2000 to

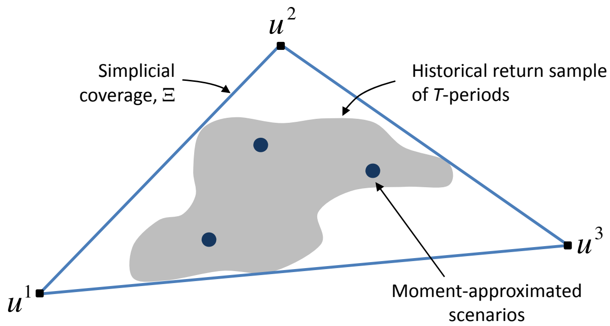

Figure 1: Second moment approximated historical return sample in $\Re^2$.

December 2009, and the underlying assets are S&P 100 stocks, see Table 1. Six stocks have been removed from the S&P 100 list as those do not have listings for the entire 10-year duration, and thus, $K = 95$. We use the SPDR Trust, which is an exchange-traded fund that holds all of the S&P 500 index stocks, as the market barometer for stock-beta calculations. SPDR trades under the ticker symbol SPY and it is listed as the last ticker in Table 1. Since $K = 95$, the second moment approximation yields 96 return scenarios, while the historical sample has over 2,500 return vectors.

The objective of the financial decision problem is to choose assets to invest (long or short) at the beginning of the year 2010 for a duration of a half-year, i.e., until June 30. However, the portfolio formed at the beginning of 2010 will be rebalanced at the end of every month as market dynamics evolve. Therefore, when the portfolio is rebalanced at the beginning of February 2010, an additional data set for January 2010 will be available and it is used to append the historical data set of 2000-2009 when constructing the approximated scenario set for the period of February 2010. Under the monthly rebalancing

AA	AAPL	ABT	AEP	ALL	AMGN	AMZN	AVP	AXP	BA	BAC	BAX
BHI	BK	BMY	BRK-B	C	CAT	CL	CMCSA	COF	COP	COST	CPB
CSCO	CVS	CVX	DD	DELL	DIS	DOW	DVN	EMC	ETR	EXC	F
FCX	FDX	GD	GE	GILD	GS	HAL	HD	HNZ	HON	HPQ	IBM
INTC	JNJ	JPM	KFT	KO	LMT	LOW	MCD	MDT	MET	MMM	MO
MON	MRK	MS	MSFT	NKE	NOV	NSC	NWSA	ORCL	OXY	PEP	PFE
PG	QCOM	RF	RTN	S	SLB	SLE	SO	T	TGT	TWX	TXN
UNH	UPS	UTX	VZ	WAG	WFC	WMB	WMT	WY	XOM	XRX	*SPY*

Table 1: List of stock ticker symbols

strategy, such an approximation is needed at the beginning of each month in the horizon, conditional upon the data available prior to that point in time. This approach results in a dynamically evolving portfolio, and the resulting portfolios are (out-of-sample) simulated using the (actual) realized stock price series during the concerned monthly period. Then, the portfolio performance is compared against the market tracker, SPY, for January-June, 2010.

For the specific experiment here, we set B^0 =\$1 million at the end of 2009, $x_j^{\min} = -10\% B^0$, $x_j^{\max} = 10\% B^0$, and $\nu = 5\%$. Therefore, no stock receives more than 10% of wealth for long/short investment, and the portfolio's market dependency is controlled within 5%. All initial positions, at the beginning of 2010 is set to zero. As the portfolio is monthly-rebalanced, B^0 is automatically adjusted to the cash position carried forward in the portfolio and x^0 is set to the beginning stock positions at the rebalancing time. For the transactions and slippage cost model, $a_{1j} = 0.02$ and $a_{2j} = 1.0$ are set for all stocks. While these parameters depend on the asset and they need to be calibrated to the market dynamics, for simplicity here, stationary constants are assumed. Expected trading volume for each asset for computing the slippage costs is determined for the trading day by the average volume of the preceding (historical) month. The portfolio monthly target return is set at an aggressive $C = 3\%$ during Jan-Jun, 2010. Thus, it corresponds to a compounded half-year return target of 19.41%. The riskfree rate is assumed to be zero.

5. Portfolio Performance

The portfolio model with approximating scenarios is evaluated by computing performance metrics for the (out-of-sample simulated) wealth series of the managed portfolio. We consider annualized rate of return (ARoR), which is the portfolio daily average rate of return, net of trading costs, annualized over 250 days of trading, as well as the annualized standard deviation (AStD), which is the standard deviation of the daily portfolio net rate of return series, annualized over 250 days of trading.

In addition, we consider an important metric portfolio performance, typically used by fund mangers, termed the maximum draw down (maxDD). Investors do not wish to see the value of the portfolio decline considerably over time. Such drastic declines in portfolio value may lead to perceptions that the fund is too risky; it may even lead to losing important client accounts from the fund. Portfolio draw down is defined as the relative equity loss from the highest peak to the lowest valley of a portfolio value decline within a given window of observation. We set this time window to Jan 01-Jun 30, 2010 for the monthly rebalanced portfolio.

For the S&P 500 index (SPY), for the period of interest, ARoR=−13.49%, AStD=20.46%, and maxDD=14.86%, expressed as 'percent of the initial budget'. We consider two models for comparison, both of which use the historical returns from 2000-2009 for computing mean and var/cov information of the 95 stocks in Table 1:

- *Target Deviation* (TD) model: Under the hypothesis that future stock returns are possibly asymmetric and non-normal, compute the second moment-approximated scenarios for the risk function in (45).
- *Mean-variance* (MV) model: Under the assumption that stock returns are normally distributed, apply Markowitz's mean-variance trade off by setting the risk function $\phi(x) = x'Mx$.

The model-based (in-sample) ARoR vs AStD portfolio trade-off under monthly rebalancing of the model in (44) is plotted for TD and MV models, see Figure 2. As expected, MV model displays a better in-sample performance relative to TD model since the MV model is optimized for mean/variance trade off. The relative performance of TD model is weaker at low portfolio AStD, while at increased risk levels, TD performs as well as the MV. Portfolio strategies of the two models are quite different as depicted in Figure 3. At low values of λ, i.e., less risk-averse, both models indicate lower levels of diversification with increased short positions in the portfolios. But, as the investor becomes more risk-averse, TD model increases long positions at the expense of short positions, but with no significant change in diversification. In contrast, the MV model dictates increased diversification, both in the long and short positions.

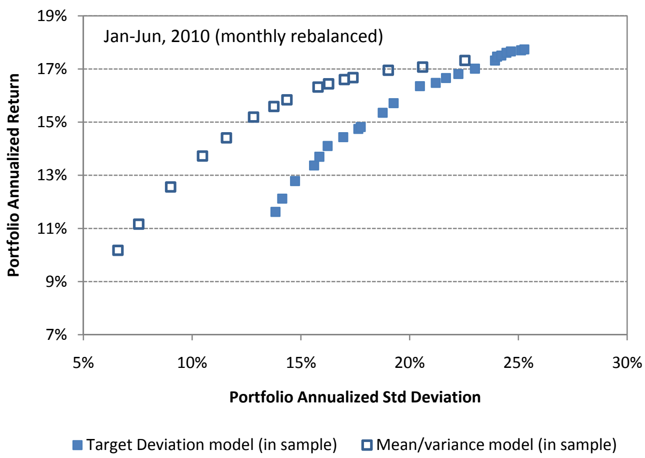

Figure 2: In-sample efficient frontiers for approximated versus normal returns.

How do the above optimal strategies perform in the out-of-sample period from January-June of 2010 when applied against the actual observed returns? Figure 4 presents the out-of-sample ARoR/AStD relationship. Although the MV model optimized the in-sample

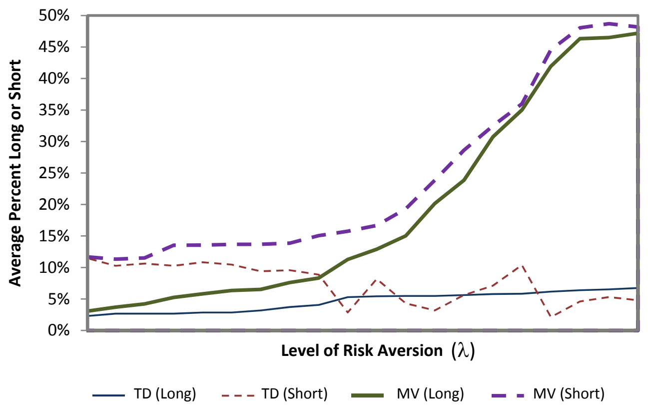

Figure 3: Portfolio long/short strategies for target-based and mean-var models.

mean-variance trade off, its actual performance during the 6 months in 2010 is significantly inferior to that of the TD model. The main insight here is that the assumption of normally distributed returns in the MV model yielded a diversification strategy that is not consistent with the actually observed returns. On the contrary, the second moment-based scenarios allow for extremal scenarios with higher probabilities (than normal distributions would allow), and accordingly, the out-of-sample performance is much improved. Observe that the market itself (S&P 500 index) performed quite poorly, indicating the extremal nature of the actual returns during the out-of-sample period.

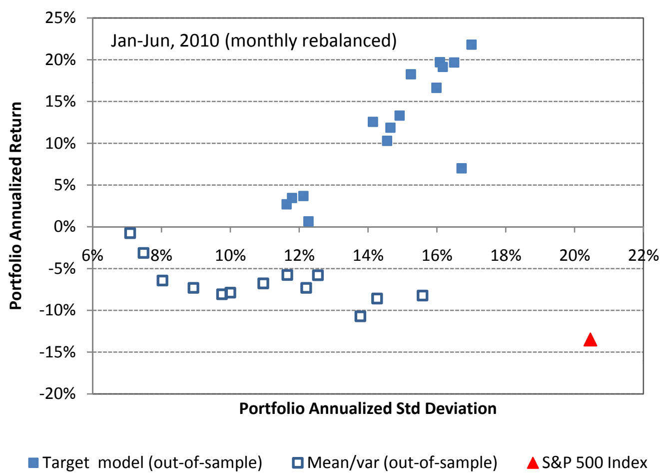

Figure 4: Out-of-sample efficient frontiers for approximated versus normal returns.

The maxDD performance metric too is quite inferior for the MV model relative to the TD model, see Figure 5. With increasing maxDD, both models yield diminished portfolio returns; however, for moderate risks levels (thus, moderate maxDD values), the performance of the TD-portfolio is outstanding. Also, observe from Figures 4 and 5 that increased standard deviation for the portfolio does not necessarily imply increased draw downs.

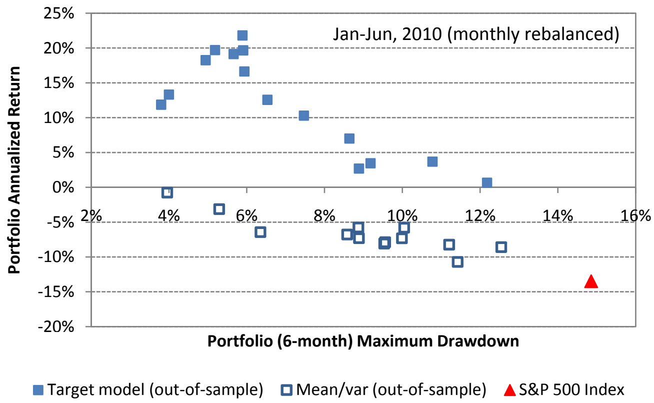

Figure 5: Portfolio Return to Maximum draw down performance.

6. Concluding Remarks

Rather than focusing on time series modeling, this paper aims at controlling risks directly via portfolio optimization, recognizing that scenarios of the future cannot be known with certainty. This is done through incorporating risk functions that involve computing expectations in high dimensions, embedded within an outer decision optimization model. The basic theory of approximating expectation functions using first and second moments utilizes the generalized moment problem, and they are applicable for general convex functions. Application of the approximation requires compact domains and simplicial coverage of such domains of random vectors. This can be easily done for financial optimization problems when return scenarios for the future must be based on historical (discrete) returns observed over some period of time.

Two competing models are compared, one in which normality is assumed for random returns, and thus, the usual mean-variance quadratic optimization (due to Markowitz) is applied for risk-return trade off. Second, under no such distributional assumption, historical returns are approximated with second moment-approximated discrete scenarios and applied within a risk function that is not symmetric, in this case, a target deficit risk model. While in-sample performance of the mean-variance trade off model is superior, as for the out-of-sample (actual) performance, the scenario-approximated target deficit model is far superior. These results corroborate well with the well-known fact that stock returns have heavier-tails and discounting extremal possibilities in returns can often lead to significant declines in fund performance.

References

[1] E.J. Anderson and P. Nash. *Linear Programming in Infinite-Dimensional Spaces.* John Wiley and Sons, New York, 1987.

[2] J.R. Birge and R.J-B. Wets. Designing Approximation Schemes for Stochastic Optimization Problems, in particular for Stochastic Programs with Recourse. *Mathematical Programming Study*, 27:54–102, 1986.

[3] M. Bychkov and N.C.P. Edirisinghe. Rational Risk Measures and Applications. Working paper, University of Tennessee, College of Business, Knoxville, TN, 2008.

[4] N.C.P. Edirisinghe. New second-order bounds on the expectation of saddle functions with applications to stochastic linear programming. *Operations Research*, 44:909–922, 1996.

[5] N.C.P. Edirisinghe. Integrated risk control using stochastic programming ALM models for money management. In S.A. Zenios and W.T. Ziemba, editors, *Handbook of Asset and Liability Management*, volume 2, chapter 16, pages 707–750. Elsevier Science BV, 2007.

[6] N.C.P. Edirisinghe. An efficient scheme for computing a compact simplex covering multivariate points. Working paper, University of Tennessee, College of Business, Knoxville, TN (under review), 2009.

[7] N.C.P. Edirisinghe and G-M. You. Second-order scenario approximation and refinement in optimization under uncertainty. *Annals of Operations Research*, 19:314–340, 1996.

[8] N.C.P. Edirisinghe and W.T. Ziemba. Tight bounds for stochastic convex programs. *Operations Research*, 40:660–677, 1992.

[9] K. Glashoff and S-A. Gustafson. *Linear Optimization and Approximation*. Springer-Verlag, New York, 1983.

[10] J.L. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30:173–177, 1906.

[11] P. Kall. *Stochastic Linear Programming*. Springer-Verlag, Berlin, 1976.

[12] P. Kall. Stochastic Programming with Recourse : Upper Bounds and Moment Problems. In K. Lommatzsch M. Vlach K. Zimmermann J. Guddat, P. Kall, editor, *Advances in Mathematical Optimization and Related Topics*. Akademie-Verlag, Berlin, 1988.

[13] A.F. Karr. Extreme Points of Certain Sets of Probability Measures, With Applications. *Mathematics of Operations Research*, 8:74–85, 1983.

[14] J. Kemperman. The General Moment Problem. A Geometric Approach. *Annals of Mathematical Statistics*, 39:93–112, 1968.

[15] H. Markowitz. *Portfolio Selection - Efficient Diversification of Investments*. John Wiley and Sons, New York, 1959.

[16] W. Ogryczak and A. Ruszczynski. Dual stochastic dominance and related mean-risk models. *SIAM Journal of Optimization*, 13(1):60–78, 2002.

[17] J.M. Eber P. Artzner, F. Delbaen and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–227, 1999.

[18] R.T. Rockafellar and S. Uryasev. Optimization of conditional value at risk. *Journal of Risk*, 2(3):21–41, 2000.

[19] R.J-B. Wets. Stochastic Programs with Fixed Recourse: The Equivalent Deterministic Problem. *SIAM Review*, 16:309–339, 1974.

[20] R.J-B. Wets. Stochastic Programming: Solution Techniques and Approximation Schemes. In B. Korte A. Bachem, M. Groetschel, editor, *Mathematical Programming: State-of-the-art*. Springer-Verlag, Berlin, 1982.

[21] W.T. Ziemba. *The Stochastic Programming Approach to Asset Liability and Wealth Management*. AIMR, 2003.