

Tractable Robust Expected Utility and Risk Models for Portfolio Optimization

Karthik Natarajan* Melvyn Sim[†] Joline Uichanco[‡]

Submitted: March 13, 2008

Abstract

Expected utility models in portfolio optimization is based on the assumption of complete knowledge of the distribution of random returns. In this paper, we relax this assumption to the knowledge of only the mean, covariance and support information. No additional assumption on the type of distribution such as normality is made. The investor's utility is modeled as a piecewise-linear concave function. We derive exact and approximate optimal trading strategies for a robust or maximin expected utility model, where the investor maximizes his worst case expected utility over a set of ambiguous distributions. The optimal portfolios are identified using a tractable conic programming approach. Using the optimized certainty equivalent (OCE) framework of Ben-Tal and Teboulle [6], we provide connections of our results with robust or ambiguous convex risk measures, in which the investor minimizes his worst case risk under distributional ambiguity. New closed form expressions for the OCE risk measures and optimal portfolios are provided for two and three piece utility functions. Computational experiments indicate that such robust approaches can provide good trading strategies in financial markets.

1 Introduction

Consider an investor deciding to allocate wealth among a set of risky assets. The uncertain payoffs of the risky assets are assumed to be sampled from a particular distribution. Expected utility theory provides a foundation of how the investor would make his choice in this setting. Von Neumann and Morgenstern [40] argue that if an investor accepts certain assumptions concerning "rational choice",

*Department of Mathematics, National University of Singapore, Singapore 117543. Email: matkbn@nus.edu.sg. The research of the author was partially supported by Singapore-MIT Alliance, NUS Risk Management Institute and NUS startup grants R-146-050-070-133 & R146-050-070-101.

[†]NUS Business School, National University of Singapore, Singapore 117592. Email: dscsimm@nus.edu.sg. The research of the author was partially supported by Singapore-MIT Alliance, NUS Risk Management Institute and NUS academic research grant R-314-000-068-122.

[‡]Computational Engineering Programme, Singapore-MIT Alliance, National University of Singapore, Singapore 117576. Email: joline.uichanco@gmail.com

then he should use expected utility as a criteria for comparing investment alternatives. In other words, a rational investor will choose a trading strategy that maximizes his expected utility.

Under the expected utility maximization framework, a natural question that arises is how to model an investor's utility function. For mathematical tractability, one typically assumes a linear risk tolerance utility, exponential utility, isoelastic utility or logarithmic utility function. However, investors' utility functions may often prove to be highly complicated or of irregular form, leading to complex problems. Markowitz [24] addressed this issue by using a quadratic utility function, arguing that it approximates well enough many other functional forms. Under Markowitz's model, the portfolio optimization problem can be solved as a quadratic programming problem. Portfolio selection in this model is only based on mean and covariance information of the uncertain returns. While the resulting model may be simple, the quadratic utility function is subject to serious limitations in modeling an investor's behavior. Firstly, the quadratic function only exhibits positive marginal utility in a bounded range, beyond which it may not represent rational behavior (see Wipern [41]). This makes it necessary to bound the range of possible outcomes. Secondly, although empirical observations suggest otherwise, the degree of risk-aversion of the quadratic utility function is everywhere increasing (see Pratt [33]; Arrow [2]). Nevertheless, Markowitz's model has been sustained in the literature due to mathematical convenience, but with most authors recognizing its limitations.

A popular method to solve the portfolio optimization problem under more general nonlinear utility models is by a Taylor series approach (see Ziemba and Vickson [42]). The nonlinear utility is then replaced by its truncated Taylor series approximation. A second order Taylor series approximation results in a quadratic programming problem similar to Markowitz's model. Loist [23] argues that the truncated Taylor series approach can possibly lead to large errors when the infinite Taylor series does not converge. Empirical studies by Pulley [34], Kroll et al. [22], Hlawitschka [19] and Ederington [13] have found that mean-variance Taylor series expansions may serve as useful approximations to expected utility, even for a non convergent series. Another approach to solving nonlinear expected utility models is through a sample-based approach. The Sample Average Approximation (SAA) approach [21] uses the average of the utility values based on random samples from the past periods. The resulting problem can then be solved using deterministic nonlinear optimization methods. The SAA approach has been considered for two-stage stochastic optimization models by Kleywegt, Shapiro and Homem-De-Mello [21]. As the sample size goes to infinity, they show that the optimal value of the SAA approach converges to that of the original problem with probability one. However, sample-based methods may sometimes give highly inconsistent approximations due to insufficient data. Natarajan, Pachamanova and Sim [27] have investigated the quality of the solution to sample-based approaches for value-at-risk (VaR) optimization. Their computational experiments show that the behavior of sample-based approaches could be erratic. In fact, the optimal weights for portfolios sometimes vary widely from sample to sample. Therefore, practitioners must use caution in using such sample based methods.

A related issue in portfolio optimization is the quantification of risk. In addition to variance, Value-at-risk (VaR) is a risk metric which is widely used by banks, security firms and other organizations. More recently, Artzner et al. [3] have introduced a class of risk measures (referred to as *coherent risk measures*)

that satisfy certain desirable properties, some of which variance and VaR do not share. These properties are positive homogeneity, subadditivity, translation invariance and monotonicity. Many variations and extensions of coherent risk measures have been proposed and studied in literature (see [35], [30], [1], [15], [31]). An important extension is based on the relaxation of the positive homogeneity and subadditivity property into the weaker convexity property (see Föllmer and Schied [15]; Frittelli and Gianin [17]). This class is referred to as *convex risk measures*. Properties of these risk measures have intuitive interpretations in the context of the risk of a portfolio. For instance, the convexity property implies that the risk of a diversified portfolio is less than individual risks. Risk measures naturally impose preference orders to random outcomes. A related concept that also imposes a preference order is the *certainty equivalent*, or the sure amount for which an investor remains indifferent to an outcome. Examples of certainty equivalents based on utility functions can be found in Buhlmann [9] and Ben-Tal and Teboulle [5], [6]. For instance, the negative of the *optimized certainty equivalent* (OCE) introduced in [6] has been shown to define a convex risk measure, under specific assumptions on the utility function. Risk measures, such as conditional value-at-risk (Rockafellar and Uryasev [35]) and bounded shortfall risk (Föllmer and Schied [15]), can in fact be derived as special cases of the OCE.

Even if we can address the issue of accurately modeling the investor's risk profile, there is still a critical issue faced by practitioners. This deals with imperfect knowledge of the exact distribution of the random returns. In such cases, it is natural to assume only partial information about the distribution, such as the mean returns and the covariance matrix. This information could be derived by using historical data or from a reliable model of future returns. Ambiguity in distributions can be then modeled by a maximin utility or minimax risk approaches. Such approaches have been used in an inventory control model under an uncertain demand distribution by Scarf [37]. Instead of fitting the data to a unique distribution, the true distribution is allowed to lie in a family of distributions. The maximin approach, consists of maximizing the worst-case profit over this set of distributions. Other robust models enforce uncertainty in parameters in a deterministic manner using convex constraints (see Ben-Tal and Nemirovski [4]; Bertsimas and Sim [8]). In portfolio optimization, Gilboa and Schmeidler [18] have proposed a maximin model to find optimal trading strategies that maximize the worst-case expected utility of an investor over a set of allowed distributions. In a related recent work by Popescu [32], the problem of deriving robust solutions to the portfolio optimization problem based on mean and covariance information is addressed. Therein, for a certain class of utility functions, the maximin problem is reduced to solving a parametric quadratic program (PQP). Our work in this paper is closely related to this work.

In this paper, we assume that the investor's utility is represented by a piecewise-linear concave function. Our contributions can be summarized as follows:

1. In Section 2, we find the tightest possible lower bounds for the worst case expected utility under distributional families of known: (a) mean and covariance information, and (b) mean and support information. We also present a lower bound (not necessarily tight) under distributions of known mean, covariance and support. This bound is based on a convolution of the bounds in (a) and (b), thus providing a computationally tractable approximation to a NP-hard problem. We also

develop a partitioned statistics based lower bound that captures asymmetry information in the distribution. All bounds and the corresponding optimal portfolios are found by solving compact conic programs. Such conic optimization problems can be solved very efficiently in both theory and practise using interior point methods (see Nesterov and Nemirovski [29]).

2. In Section 3, we provide a connection of our utility bounds with ambiguous risk measures by defining a worst case OCE risk measure. For two and three piece utility functions, we provide simple closed form expressions for this new risk measure. In the two asset case with one risk-free and one risky asset, the optimal portfolio for the two piece function results in no diversification while the three piece function results in diversification. Connections with convex risk measures is provided based on our results in Section 2.
3. In Section 4, we use computational experiments on financial market data to verify the robustness of both the worst case OCE and sample based methods. Our results indicate that an optimal rebalancing portfolio derived using the worst case OCE approach can perform better than sample-based approaches, especially in out-of-sample data. The method appears to be robust across different periods of the investment horizon.

While finishing this paper, we became aware of recent unpublished work by Delage and Ye [12] that also solves the maximin expected utility model under moment information. We outline some of the key similarities and differences between our work and theirs.

- (a) *Ambiguity Modeling*: In Delage and Ye [12], the ambiguity in distribution is represented using support information with the true mean assumed to lie in an ellipsoidal set around an estimated mean and the difference of the estimated covariance matrix and the true covariance matrix assumed to lie in a positive semi-definite cone. In our paper, the ambiguity in distribution is represented using support information with the true mean and covariance matrix known exactly.
- (b) *Algorithms and Complexity*: Delage and Ye [12] prove polynomial time solvability for bounds under their ambiguity model using an ellipsoidal type algorithm for a fairly large class of utility functions (not just piecewise linear functions). On the other hand, we prove polynomial time solvability using compact conic programming formulations for piecewise linear functions under our ambiguity model. For the case with support equal to \mathbb{R}^n , they obtain the tight bound by solving a semidefinite program while we obtain the tight bound by solving a second order conic program. For general convex supports, they propose the use of an ellipsoidal type algorithm to find a lower bound (not tight). On the other hand, for conic representable supports, we obtain lower bounds (not tight) using compact conic programming formulations. An interesting future avenue of research is to compare the quality of bounds from the two approaches and the running times.
- (c) *Focus*: A key focus in Delage and Ye [12] is to use sample data to help derive confidence regions for the mean and covariance matrix to model their ambiguous distributions. On the other hand, our focus is on relating the results with worst case risk measures in finance using the OCE framework and deriving insights into the structure of the optimal portfolio for simple utility functions.

2 Expected Utility under Ambiguous Distributions

Let $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_n)$ be a vector of n random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We define the space of linear combinations of the random variables including constants c as follows:

$$\mathcal{X} = \{\tilde{x} : \exists(c, \mathbf{y}) \in \mathfrak{R} \times \mathfrak{R}^n : \tilde{x} = c + \mathbf{y}'\tilde{\mathbf{z}}\}.$$

We denote by \mathcal{W} the support of the random variable $\tilde{\mathbf{z}}$. Throughout the paper, we will use the notation $\tilde{x}_1 \geq \tilde{x}_2$ for $\tilde{x}_1, \tilde{x}_2 \in \mathcal{X}$ to represent state-wise dominance. Hence, if $\tilde{x}_1 = c_1 + \mathbf{y}_1'\tilde{\mathbf{z}}$ and $\tilde{x}_2 = c_2 + \mathbf{y}_2'\tilde{\mathbf{z}}$, then $\tilde{x}_1 \geq \tilde{x}_2$ is equivalent to

$$c_1 + \mathbf{y}_1'\mathbf{z} \geq c_2 + \mathbf{y}_2'\mathbf{z} \quad \forall \mathbf{z} \in \mathcal{W}. \quad (2.1)$$

Note that it is easy to see that the feasibility of (2.1) remains unchanged if we replace \mathcal{W} by its convex hull (see for instance Ben-Tal and Nemirovski [4]). Henceforth, without loss of the generality, we assume that the support of $\tilde{\mathbf{z}}$ is convex. Consider a piecewise linear concave utility function $u(\cdot) : \mathfrak{R} \rightarrow \mathfrak{R}$ defined as follows:

$$u(x) = \min_{k \in \{1, \dots, K\}} \{a_k x + b_k\}. \quad (2.2)$$

Such a piecewise linear function can serve as an approximation to general concave utility functions. The utility function is assumed to be non-degenerate with at least two distinct pieces ($K \geq 2$) but not too large (K is polynomial in the dimension of the problem n). No other assumption such as a non-decreasing function is needed for the results in this section. The expected utility is then given as

$$\mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right].$$

An implicit assumption in computing this expected value is the exact knowledge of the distribution \mathbb{P} . In practice, one seldom has full information about the multivariate distribution of $\tilde{\mathbf{z}}$. Furthermore, evaluating this expected utility can be a numerically challenging task. Suppose, instead that the true distribution \mathbb{P} is known to lie in a family of distributions \mathbb{F} . All distributions in \mathbb{F} are assumed to satisfy certain known properties, such as a fixed set of moments. The worst case expected utility is then defined as

$$\hat{u}(c + \mathbf{y}'\tilde{\mathbf{z}}) = \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right]. \quad (2.3)$$

In the portfolio optimization context, $\tilde{\mathbf{z}}$ represents the random risky payoffs of n assets and \mathbf{y} denotes the allocation vector. In addition, let $r \geq 0$ denotes the return of the risk-free asset and y_0 be the allocation in it. The random payoff of the portfolio is then given as

$$y_0 r + \mathbf{y}'\tilde{\mathbf{z}} \in \mathcal{X}.$$

We denote by Y the convex feasible region of the portfolio allocation vector (y_0, \mathbf{y}) . For example, if we normalize the sum of allocations to be one and prohibit short selling, we have

$$Y = \{(y_0, \mathbf{y}) : y_0 + \mathbf{e}'\mathbf{y} = 1, (y_0, \mathbf{y}) \geq \mathbf{0}\},$$

where \mathbf{e} is a vector of ones. The general portfolio expected utility maximization problem is then formulated as

$$\sup_{(y_0, \mathbf{y}) \in Y} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(y_0 r + \mathbf{y}' \tilde{\mathbf{z}}) + b_k\} \right], \quad (2.4)$$

for a given set of feasible portfolio allocations Y . To achieve robustness over a set of ambiguous distributions, the investor could choose to maximize his worst-case expected utility. The portfolio maximin expected utility problem is then formulated as:

$$\sup_{(y_0, \mathbf{y}) \in Y} \hat{u}(y_0 r + \mathbf{y}' \tilde{\mathbf{z}}) = \sup_{(y_0, \mathbf{y}) \in Y} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(y_0 r + \mathbf{y}' \tilde{\mathbf{z}}) + b_k\} \right]. \quad (2.5)$$

Our focus is on solving the maximin expected utility model where the ambiguity in the distribution is specified in terms of moments of the distribution.

2.1 Mean and Covariance Information

Assume that for the random returns $\tilde{\mathbf{z}}$, only the mean vector $\boldsymbol{\mu} = \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{z}}]$ and covariance matrix $\mathbf{Q} = \mathbb{E}_{\mathbb{P}}[(\tilde{\mathbf{z}} - \boldsymbol{\mu})(\tilde{\mathbf{z}} - \boldsymbol{\mu})']$ are explicitly known. For this class of distributions, Popescu [32] shows that the problem of evaluating the worst case expected utility by optimizing over a n -variate distribution can be reduced to an optimization over a univariate distribution with the appropriate mean and variance. The result is stated next for completeness.

Proposition 1 (Popescu [32]). *Let \mathbb{F}_1 be the family of all distributions \mathbb{P} for $\tilde{\mathbf{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{Q} . For any utility function u , we have that*

$$\inf_{\mathbb{P} \in \mathbb{F}_1} \mathbb{E}_{\mathbb{P}} [u(c + \mathbf{y}' \tilde{\mathbf{z}})] = \inf_{\mathbb{P}_x \in \mathbb{F}_x} \mathbb{E}_{\mathbb{P}_x} [u(\tilde{x})],$$

where \mathbb{F}_x is the family of all univariate distributions of \tilde{x} with mean $\mu_x = c + \mathbf{y}' \boldsymbol{\mu}$ and variance $\sigma_x^2 = \mathbf{y}' \mathbf{Q} \mathbf{y}$.

Proposition 1 is based on the following projection property: for any random variable \tilde{x} with mean $c + \mathbf{y}' \boldsymbol{\mu}$ and variance $\mathbf{y}' \mathbf{Q} \mathbf{y}$, there exists a random vector $\tilde{\mathbf{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{Q} . The problem can be further reduced to optimizing over univariate distributions with at most three support points. This follows from the classical result in the problem of moments (see Rogosinsky [36]): for a moment problem with q known moments, there exists an extremal distribution with at most $q + 1$ support points. For a fixed portfolio, the worst case expected utility thus reduces to solving a deterministic optimization problem in at most three variables. For a special class of utility functions, Popescu [32] proposes the use of a parametric quadratic program to find the portfolio that maximizes the minimum expected utility. We state the result next.

Proposition 2 (Popescu [32]). Let \mathbb{F}_1 be the family of all distributions \mathbb{P} for $\tilde{\mathbf{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{Q} and \mathbb{F}_x be the family of all univariate distributions for \tilde{x} with mean $\mu_x = y_0 r + \mathbf{y}'\boldsymbol{\mu}$ and variance $\sigma_x^2 = \mathbf{y}'\mathbf{Q}\mathbf{y}$. Suppose the objective function $\inf_{\mathbb{P}_x \in \mathbb{F}_x} \mathbb{E}_{\mathbb{P}_x} [u(\tilde{x})]$, is continuous, nondecreasing in μ_x , nonincreasing in σ_x and quasi-concave in (μ_x, σ_x) . Then the maximin expected utility model is equivalent to solving a parametric quadratic program (PQP):

$$\operatorname{argsup}_{(y_0, \mathbf{y}) \in Y} \inf_{\mathbb{P} \in \mathbb{F}_1} \mathbb{E}_{\mathbb{P}} [u(y_0 r + \mathbf{y}'\tilde{\mathbf{z}})] = \operatorname{argsup}_{(y_0, \mathbf{y}) \in Y} \lambda(y_0 r + \mathbf{y}'\boldsymbol{\mu}) - (1 - \lambda)\mathbf{y}'\mathbf{Q}\mathbf{y},$$

for a suitable value of $\lambda \in [0, 1]$.

Utility functions for which Proposition 2 can be used include increasing concave utility functions with convex or concave-convex derivative. It also includes the class of increasing concave utility functions with at most one or two point support structure. We now show that for the class of piecewise linear concave utility functions (not necessarily with the one or two point support structure), the maximin expected utility problem can in fact be solved as a single compact second order cone program (SOCP). This provides a significant computational advantage over using a PQP to solve the maximin problem where one needs to solve multiple instances (possibly an exponential number) of a convex quadratic program for different λ values.

Theorem 2.1. Let \mathbb{F}_1 be the family of all distributions \mathbb{P} for $\tilde{\mathbf{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{Q} . The worst case expected utility:

$$\hat{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}}) = \inf_{\mathbb{P} \in \mathbb{F}_1} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right],$$

is given as the optimal objective value to the problem:

$$\hat{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}}) = \sup_{z \geq 0, t} \left\{ \min_{k \in \{1, \dots, K\}} (a_k(c + \mathbf{y}'\boldsymbol{\mu}) + b_k - a_k^2 z + a_k t) - \left(\frac{\mathbf{y}'\mathbf{Q}\mathbf{y} + t^2}{4z} \right) \right\}, \quad (2.6)$$

or equivalently:

$$\begin{aligned} \hat{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}}) = \inf_{\lambda_k} & \sum_{k=1}^K (a_k(c + \mathbf{y}'\boldsymbol{\mu}) + b_k)\lambda_k - \sqrt{\mathbf{y}'\mathbf{Q}\mathbf{y}} \sqrt{\sum_{k=1}^K a_k^2 \lambda_k - \left(\sum_{k=1}^K a_k \lambda_k \right)^2} \\ \text{s.t.} & \sum_{k=1}^K \lambda_k = 1, \\ & \lambda_k \geq 0, \quad \forall k = 1, \dots, K. \end{aligned} \quad (2.7)$$

Proof. See Appendix A.

A natural implication of Theorem 2.1 is that the maximin expected utility model in (2.5) under given mean and covariance matrix can be solved efficiently as a compact second order cone program. when the feasible region Y is representable using the SOCP constraints:

$$\begin{aligned}
& \sup_{z,t,w,s,y_0,\mathbf{y}} && w - s \\
& \text{s.t.} && w \leq a_k(y_0r + \mathbf{y}'\boldsymbol{\mu}) + b_k - a_k^2z + a_k t, \quad \forall k = 1, \dots, K, \\
& && 4zs \geq \mathbf{y}'\mathbf{Q}\mathbf{y} + t^2, \\
& && z \geq 0, \\
& && (y_0, \mathbf{y}) \in Y.
\end{aligned} \tag{2.8}$$

Note that the constraint $4zs \geq \mathbf{y}'\mathbf{Q}\mathbf{y} + t^2$, known as the rotated SOCP constraint, can be transformed to a standard SOCP constraint as follows

$$z + s \geq \sqrt{\mathbf{y}'\mathbf{Q}\mathbf{y} + t^2 + (z - s)^2}.$$

The variables λ_k in (2.7) can be interpreted as the probability that the k th piece of the utility function is chosen. Suppose, $\lambda_k = 1$, then we simply get the expected utility for the k th piece, $a_k\mu_x + b_k$. The first term in the bound is thus simply a convex combination of these linear approximations for each of the K pieces based on the mean while the second term is a penalty based on the variance. A similar formulation has been used to determine choice probabilities in a discrete choice model by Natarajan et.al [28]. The result here extends the approach to decision making in robust portfolio optimization problems. In some special cases, Theorem 2.1 can be solved in closed form. Two such results for the two and three-piece utility functions (see Figure 1) are provided next.

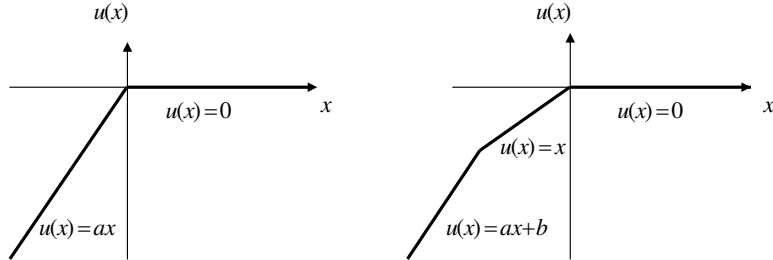


Figure 1: Two and three-piece utility functions

Proposition 3 (Cauchy-Schwarz inequality, Natarajan and Zhou [26]). *For any random variable \tilde{x} with mean μ_x and variance $\sigma_x^2 > 0$,*

- (a) *The worst case expected utility for the two-piece utility function $u(\tilde{x}) = \min\{a\tilde{x}, 0\}$ with $a > 1$ is given as*

$$\hat{u}_1(\tilde{x}) = \frac{a}{2} \left(\mu_x - \sqrt{\mu_x^2 + \sigma_x^2} \right).$$

(b) The worst case expected utility for the three-piece utility function $u(\tilde{x}) = \min\{a_k\tilde{x} + b_k, \tilde{x}, 0\}$ with $a > 1$ and $b > 0$ is given as

$$\hat{u}_1(\tilde{x}) = \begin{cases} \frac{1}{2} \left(\mu_x - \sqrt{\mu_x^2 + \sigma_x^2} \right), & \text{if } \sigma_x^2 \leq \left(\frac{b}{a(a-1)} + \mu_x \right) \left(\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left((a+1)\mu_x + b - \sqrt{((a-1)\mu_x + b)^2 + (a-1)^2\sigma_x^2} \right), & \text{if } \sigma_x^2 \leq \left(\frac{(2a-1)b}{a(a-1)} + \mu_x \right) \left(-\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left(a\mu_x + b - \sqrt{(a\mu_x + b)^2 + a^2\sigma_x^2} \right), & \text{if } \sigma_x^2 \geq \left(\frac{(2a-1)b}{a(a-1)} + \mu_x \right) \left(\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left(\mu_x - \frac{a(a-1)(\mu_x^2 + \sigma_x^2)}{2b} - \frac{b}{2a(a-1)} \right), & \text{otherwise.} \end{cases}$$

The bound in Proposition 3(a) is obtained by solving

$$\inf_{0 \leq \lambda \leq 1} a\lambda\mu_x - a\sigma_x\sqrt{\lambda(1-\lambda)},$$

and reduces to a version of the standard Cauchy-Schwarz inequality. The bound in Proposition 3(b) is obtained by solving

$$\begin{aligned} \inf_{\lambda_1, \lambda_2} & (a\mu_x + b)\lambda_1 + \mu_x\lambda_2 - \sigma_x\sqrt{a^2\lambda_1 + \lambda_2 - (a\lambda_1 + \lambda_2)^2} \\ \text{s.t.} & \lambda_1 + \lambda_2 \leq 1, \\ & \lambda_1, \lambda_2 \geq 0. \end{aligned}$$

Solving this convex programming problem in closed form, while more involved, can be done explicitly using the Karush-Kuhn-Tucker conditions. The reader is referred to Theorem 1, pg. 613 in Natarajan and Zhou [26] for a proof of this result. Our formulation in Theorem 2.1 generalizes these results to K pieces.

2.2 Mean and Support Information

Assume that for the random returns \tilde{z} , only the mean vector $\boldsymbol{\mu} = E_{\mathbb{P}}[\tilde{z}]$ and a support \mathcal{W} are explicitly known. The set \mathcal{W} is assumed to be a conic representable set of the form:

$$\mathcal{W} = \{z : \exists \mathbf{u} : \mathbf{D}z + \mathbf{F}\mathbf{u} - \mathbf{g} \in \mathbf{K}\}, \quad (2.9)$$

where \mathbf{K} is a regular cone, i.e., it is closed, convex, pointed, and has a non-empty interior. This includes the nonnegative orthant, the second order cone, the cone of positive semidefinite matrices and their cartesian product as special cases. The corresponding polar cone

$$\mathbf{K}^* = \{\mathbf{w} : \mathbf{w}'\mathbf{s} \geq 0 \forall \mathbf{s} \in \mathbf{K}\},$$

is also a regular cone. We assume that the mean vector lying strictly in the interior of this set, namely $\boldsymbol{\mu} \in \text{int}(\mathcal{W})$.

Theorem 2.2. Let \mathbb{F}_2 be the family of all distributions \mathbb{P} for $\tilde{\mathbf{z}}$ with mean $\boldsymbol{\mu}$ and support \mathcal{W} . Suppose $\boldsymbol{\mu}$ lies in the interior of the set \mathcal{W} . The worst case expected utility:

$$\hat{u}_2(c + \mathbf{y}'\tilde{\mathbf{z}}) = \inf_{\mathbb{P} \in \mathbb{F}_2} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right]$$

is given as the optimal objective value to the problem:

$$\hat{u}_2(c + \mathbf{y}'\tilde{\mathbf{z}}) = \sup_{\mathbf{s}} \left(\mathbf{s}'\boldsymbol{\mu} + \min_{k \in \{1, \dots, K\}} \left\{ \inf_{\mathbf{z} \in \mathcal{W}} (a_k \mathbf{y} - \mathbf{s})' \mathbf{z} + a_k c + b_k \right\} \right). \quad (2.10)$$

Proof. See Appendix A.

An implication of Theorem 2.2 is that the maximin expected utility problem in (2.5) under given mean and bounded support can be solved as a compact conic program when the feasible region Y is also conic representable. This follows directly by taking the dual formulation for the inner minimization problem over $\mathbf{z} \in \mathcal{W}$. The maximin expected utility model is then reformulated as the conic program:

$$\begin{aligned} \sup_{\mathbf{s}, t, \mathbf{w}^{(k)}, y_0, \mathbf{y}} \quad & \mathbf{s}'\boldsymbol{\mu} + t \\ \text{s.t.} \quad & t \leq \mathbf{w}^{(k)'} \mathbf{g} + a_k y_0 r + b_k, \quad \forall k = 1, \dots, K, \\ & \mathbf{D}' \mathbf{w}^{(k)} = a_k \mathbf{y} - \mathbf{s}, \quad \forall k = 1, \dots, K, \\ & \mathbf{F}' \mathbf{w}^{(k)} = \mathbf{0}, \quad \forall k = 1, \dots, K, \\ & \mathbf{w}^{(k)} \in \mathbf{K}^*, \quad \forall k = 1, \dots, K, \\ & (y_0, \mathbf{y}) \in Y. \end{aligned} \quad (2.11)$$

2.3 Mean, Covariance and Support Information

A natural question of interest arises when we consider distribution families of known mean, covariance and support information. Even though each of the previously established bounds remains valid under the intersection of the two families, there is no longer a guarantee of tightness. In fact, the complexity of finding the worst case expected utility under the intersection of these two sets is in general a NP-hard problem (see Murty and Kabadi [25]; Bertsimas and Popescu [7]). For example, characterizing distributions over $\mathcal{W} = \mathfrak{R}_+^n$ with a given mean-covariance matrix is equivalent to characterizing the cone of completely positive matrices. This is known to be a NP-hard problem. In this section, we sacrifice tightness in order to produce mathematically tractable bounds.

To develop the bound, we define $\pi_l(\mathbf{y}, \mathbf{d}) : \mathfrak{R}^n \times \mathfrak{R}^K \mapsto \mathfrak{R}$ to denote the worst case expected utility over a set of distributions \mathbb{F}_l :

$$\pi_l(\mathbf{y}, \mathbf{d}) = \inf_{\mathbb{P} \in \mathbb{F}_l} \mathbb{E} \left[\min_{k \in \{1, \dots, K\}} \{a_k \mathbf{y}' \tilde{\mathbf{z}} + d_k\} \right].$$

Observe that $\pi_l(\mathbf{y}, \mathbf{a}c + \mathbf{b}) = \hat{u}_l(c + \mathbf{y}'\tilde{\mathbf{z}})$ reduces to the utility bounds in the previous sections for the appropriate set of distributions \mathbb{F}_l . The reason for introducing $\pi_l(\mathbf{y}, \mathbf{d})$ is that this function is a positive homogeneous and concave function in its arguments. Using this, we improve on the existing bounds by taking a convolution.

Theorem 2.3. Let $\mathcal{L} = \{1, \dots, L\}$ and $\{\pi_l(\mathbf{y}, \mathbf{d}), l \in \mathcal{L}\}$ be jointly concave, positive homogenous functions that each denote the worst case expected utility over the family of distributions \mathbb{F}_l . Then

$$\mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right] \geq \hat{u}_{\mathcal{L}}(c + \mathbf{y}'\tilde{\mathbf{z}}) \geq \max_{l \in \mathcal{L}} \hat{u}_l(c + \mathbf{y}'\tilde{\mathbf{z}}), \quad \forall \mathbb{P} \in \bigcap_{l \in \mathcal{L}} \mathbb{F}_l,$$

where

$$\hat{u}_{\mathcal{L}}(c + \mathbf{y}'\tilde{\mathbf{z}}) = \pi_{\mathcal{L}}(\mathbf{y}, \mathbf{a}c + \mathbf{b})$$

and

$$\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d}) = \sup_{\mathbf{y}_l, \mathbf{d}_l} \left\{ \sum_{l \in \mathcal{L}} \pi_l(\mathbf{y}_l, \mathbf{d}_l) : \sum_{l \in \mathcal{L}} (\mathbf{y}_l, \mathbf{d}_l) = (\mathbf{y}, \mathbf{d}) \right\},$$

is itself a jointly concave and positive homogeneous bound.

Proof. See Appendix A.

A direct corollary of the Theorem 2.3 is that we can obtain a polynomial time solvable conic programming approximation for the maximin expected utility model under mean, covariance and support information by combining the results from Sections 2.1-2.2:

$$\begin{aligned} & \sup_{z_1, t_1, w_1, s_1, \mathbf{y}_1, \mathbf{s}_2, t_2, \mathbf{w}_2^{(k)}, \mathbf{d}_1, \mathbf{d}_2, \mathbf{y}_2, y_0, \mathbf{y}} && w_1 - s_1 + \mathbf{s}_2' \boldsymbol{\mu} + t_2 \\ & \text{s.t.} && w_1 \leq a_k \boldsymbol{\mu}' \mathbf{y}_1 + d_{1,k} - a_k^2 z_1 + a_k t_1, \quad \forall k = 1, \dots, K, \\ & && 4z_1 s_1 \geq \mathbf{y}_1' \mathbf{Q} \mathbf{y}_1 + t_1^2, \\ & && z_1 \geq 0, \\ & && t_2 \leq \mathbf{w}_2^{(k)'} \mathbf{g} + d_{2,k}, \quad \forall k = 1, \dots, K, \\ & && \mathbf{D}' \mathbf{w}_2^{(k)} = a_k \mathbf{y}_2 - \mathbf{s}_2, \quad \forall k = 1, \dots, K, \\ & && \mathbf{F}' \mathbf{w}_2^{(k)} = \mathbf{0}, \quad \forall k = 1, \dots, K, \\ & && \mathbf{w}_2^{(k)} \in \mathbf{K}^*, \quad \forall k = 1, \dots, K, \\ & && \mathbf{y}_1 + \mathbf{y}_2 = \mathbf{y}, \\ & && \mathbf{d}_1 + \mathbf{d}_2 = \mathbf{a}y_0 r + \mathbf{b}, \\ & && (y_0, \mathbf{y}) \in Y. \end{aligned} \tag{2.12}$$

The variables with subscripts of 1 correspond to mean-covariance bound in Section 2.1 and the variables with subscripts of 2 correspond to the mean-support bound in Section 2.2.

2.4 Partitioned Statistics

One of the main criticisms of using only the first and second moments information is the inability of capturing distributional skewness. To capture distributional asymmetry in a computationally tractable way, we partition the random variables into its positive and negative parts and calculate the mean and covariance matrices of each of these two vectors. We outline the steps next.

The random vector \tilde{z} can be expressed as

$$\tilde{z} = \tilde{z}^+ - \tilde{z}^-,$$

where \tilde{z}^+ and \tilde{z}^- are given entry-wise as $\tilde{z}_i^+ = \max(0, \tilde{z}_i)$, $\tilde{z}_i^- = \max(0, -\tilde{z}_i)$. We then have, $\mathbf{y}'\tilde{z} = \mathbf{y}'\tilde{z}^+ - \mathbf{y}'\tilde{z}^-$. In general, for any vectors $\mathbf{y}^p, \mathbf{y}^m \in \mathfrak{R}^n$, we can consider $\mathbf{y}^{p'}\tilde{z}^+ + \mathbf{y}^{m'}\tilde{z}^-$ and express the expected utility as

$$\mathbb{E}_{\mathbb{P}} \left(\min_{k \in \{1, \dots, K\}} \{a_k (c + \mathbf{y}^{p'}\tilde{z}^+ + \mathbf{y}^{m'}\tilde{z}^-) + b_k\} \right). \quad (2.13)$$

These partitioned random variables \tilde{z}^+, \tilde{z}^- naturally imply a positive support (i.e, \mathfrak{R}_+^{2n}). Moreover, partitioning into \tilde{z}^+ and \tilde{z}^- isolates the statistical information whenever returns are positive and negative respectively. Suppose that we have information about the first two moments of the partitioned random variables:

$$\begin{aligned} \boldsymbol{\mu}^p &= \mathbb{E}_{\mathbb{P}}(\tilde{z}^+), \\ \boldsymbol{\mu}^m &= \mathbb{E}_{\mathbb{P}}(\tilde{z}^-), \\ \bar{\mathbf{Q}} &= \mathbb{E}_{\mathbb{P}} \left(\begin{pmatrix} \tilde{z}^+ - \boldsymbol{\mu}^p \\ \tilde{z}^- - \boldsymbol{\mu}^m \end{pmatrix} \begin{pmatrix} \tilde{z}^+ - \boldsymbol{\mu}^p \\ \tilde{z}^- - \boldsymbol{\mu}^m \end{pmatrix}' \right). \end{aligned}$$

Clearly, $\boldsymbol{\mu} = \boldsymbol{\mu}^p - \boldsymbol{\mu}^m$. Note that the covariance matrix $\bar{\mathbf{Q}}$ is a $2n$ by $2n$ positive semidefinite matrix given by

$$\bar{\mathbf{Q}} = \text{var} \left(\begin{pmatrix} \tilde{z}^+ \\ \tilde{z}^- \end{pmatrix} \right) = \begin{bmatrix} \underbrace{\text{var}(\tilde{z}^+)}_{\bar{\mathbf{Q}}_{11}} & \underbrace{\text{cov}(\tilde{z}^+, \tilde{z}^-)}_{\bar{\mathbf{Q}}_{12}} \\ \underbrace{\text{cov}(\tilde{z}^+, \tilde{z}^-)}_{\bar{\mathbf{Q}}_{12}} & \underbrace{\text{var}(\tilde{z}^-)}_{\bar{\mathbf{Q}}_{22}} \end{bmatrix},$$

where

$$\text{cov}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \triangleq \mathbb{E}((\tilde{\mathbf{x}} - \mathbb{E}(\tilde{\mathbf{x}}))(\tilde{\mathbf{y}} - \mathbb{E}(\tilde{\mathbf{y}}))')$$

and

$$\text{var}(\tilde{\mathbf{x}}) \triangleq \text{cov}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}).$$

Moreover,

$$\text{var}(\mathbf{y}'\tilde{z}) = \mathbf{y}'\mathbf{Q}\mathbf{y} = \text{var}(\mathbf{y}'\tilde{z}^+ - \mathbf{y}'\tilde{z}^-) = (\mathbf{y}' - \mathbf{y}')\bar{\mathbf{Q}} \begin{pmatrix} \mathbf{y} \\ -\mathbf{y} \end{pmatrix}. \quad (2.14)$$

Hence,

$$\mathbf{Q} = \text{var}(\tilde{z}^+) - 2\text{cov}(\tilde{z}^-, \tilde{z}^+) + \text{var}(\tilde{z}^-) = \bar{\mathbf{Q}}_{11} - 2\bar{\mathbf{Q}}_{12} + \bar{\mathbf{Q}}_{22}.$$

Thus, clearly the mean and covariance of the asset returns can be derived from the partitioned statistics of the random variable.

We can find lower bounds to the expected utility in (2.13) using Theorems 2.1 and 2.2. From the mean and covariance information of the partitioned returns, we can establish the bound

$$\bar{u}_1(c + \mathbf{y}^{p'}\tilde{z}^+ + \mathbf{y}^{m'}\tilde{z}^-) = \bar{\pi}_1(\mathbf{y}^p, \mathbf{y}^m, \mathbf{a}c + \mathbf{b})$$

in which

$$\begin{aligned} \bar{\pi}_1(\mathbf{y}^p, \mathbf{y}^m, \mathbf{d}) &= \sup_{z, t, w, s} w - s \\ \text{s.t. } & w \leq a_k(\mathbf{y}^{p'} \boldsymbol{\mu}^p + \mathbf{y}^{m'} \boldsymbol{\mu}^m) + d_k - a_k^2 z + a_k t, \quad \forall k = 1, \dots, K, \\ & z \geq 0, \\ & 4zs \geq \begin{pmatrix} \mathbf{y}^p \\ \mathbf{y}^m \end{pmatrix}' \bar{\mathbf{Q}} \begin{pmatrix} \mathbf{y}^p \\ \mathbf{y}^m \end{pmatrix} + t^2. \end{aligned}$$

Likewise, under the mean and support information ($\mathcal{W} = \mathfrak{R}_+^{2n}$), we have

$$\bar{u}_2(c + \mathbf{y}^{p'} \tilde{\mathbf{z}}^+ + \mathbf{y}^{m'} \tilde{\mathbf{z}}^-) = \bar{\pi}_2(\mathbf{y}^p, \mathbf{y}^m, \mathbf{a}c + \mathbf{b})$$

where

$$\begin{aligned} \bar{\pi}_2(\mathbf{y}^p, \mathbf{y}^m, \mathbf{d}) &= \sup_{\mathbf{s}^p, \mathbf{s}^m, t} \mathbf{s}^{p'} \boldsymbol{\mu}^p + \mathbf{s}^{m'} \boldsymbol{\mu}^m + t \\ \text{s.t. } & t \leq d_k, \quad \forall k = 1, \dots, K, \\ & a_k \mathbf{y}^p - \mathbf{s}^p \geq \mathbf{0}, \quad \forall k = 1, \dots, K, \\ & a_k \mathbf{y}^m - \mathbf{s}^m \geq \mathbf{0}, \quad \forall k = 1, \dots, K. \end{aligned}$$

Then by convolution, we can derive the partitioned statistics lower bound for the worst-case expected utility in (2.3), which is tighter than the worst-case mean-covariance bound.

Theorem 2.4. *Let \mathbb{F}_3 be the family of all distributions \mathbb{P} for $\tilde{\mathbf{z}}$ with known mean $(\boldsymbol{\mu}^p, \boldsymbol{\mu}^m)$ and covariance $(\bar{\mathbf{Q}})$ of the partitioned random variables. Let \mathbb{F}_1 be a family of distributions, with mean $\boldsymbol{\mu} = \boldsymbol{\mu}^p - \boldsymbol{\mu}^m$ and covariance $\mathbf{Q} = \bar{\mathbf{Q}}_{11} - 2\bar{\mathbf{Q}}_{12} + \bar{\mathbf{Q}}_{22}$. We define a new lower bound:*

$$\hat{u}_3(c + \mathbf{y}' \tilde{\mathbf{z}}) = \pi_3(\mathbf{y}, \mathbf{a}c + \mathbf{b}) \tag{2.15}$$

where

$$\begin{aligned} \pi_3(\mathbf{y}, \mathbf{d}) &= \sup_{\mathbf{y}_1^p, \mathbf{y}_1^m, \mathbf{y}_2^p, \mathbf{y}_2^m, \mathbf{d}_1, \mathbf{d}_2} \bar{\pi}_1(\mathbf{y}_1^p, \mathbf{y}_1^m, \mathbf{d}_1) + \bar{\pi}_2(\mathbf{y}_2^p, \mathbf{y}_2^m, \mathbf{d}_2) \\ \text{s.t. } & \mathbf{y}_1^p + \mathbf{y}_2^p = \mathbf{y}, \\ & \mathbf{y}_1^m + \mathbf{y}_2^m = -\mathbf{y}, \\ & \mathbf{d}_1 + \mathbf{d}_2 = \mathbf{d}, \end{aligned}$$

which satisfies

$$\inf_{\mathbb{P} \in \mathbb{F}_3} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}' \tilde{\mathbf{z}}) + b_k\} \right] \geq \hat{u}_3(c + \mathbf{y}' \tilde{\mathbf{z}}) \geq \underbrace{\inf_{\mathbb{P} \in \mathbb{F}_1} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}' \tilde{\mathbf{z}}) + b_k\} \right]}_{=\hat{u}_1(c + \mathbf{y}' \tilde{\mathbf{z}})}.$$

Proof. See Appendix A.

As in Section 2.3, we can optimize over this partitioned statistics bound using conic programming. Detailed empirical comparison of these bounds are provided in the computational results in Section 4.

3 Connection to Risk Measures

In this section, we will use the bounds to define new ambiguous risk measures. Based on the previous bounds, we can obtain approximations to the *worst case risk* (see Erdoğan and Iyengar [14]; Calafiore [10]) that the investor may face given the uncertain model of returns. Connections with convex (and under special cases, coherent) risk measures are also explored.

An important issue in portfolio optimization is the measurement of the risk of an investment. An axiomatic approach to defining acceptable properties of risk measures was introduced in Artzner et al. [3]. The class of risk measures introduced therein is called coherent risk measures. In financial risk management, a coherent risk can be viewed as a maximum expected loss under a set of probability measures. A relaxation of these properties gives rise to a larger class of risk measures known as convex risk measures (Föllmer and Schied [15]; Frittelli and Gianin [17]). Consider a random outcome $\tilde{x} = c + \mathbf{y}'\tilde{\mathbf{z}} \in \mathcal{X}$ which represents the uncertain payoff. The axiomatic characterization of convex and coherent risk measures is then given as:

Definition 3.1. *A function $\rho : \mathcal{X} \mapsto \mathfrak{R}$ is a convex risk measure if it satisfies, for all $\tilde{x}, \tilde{y} \in \mathcal{X}$:*

1. Monotonicity: *If $\tilde{x} \geq \tilde{y}$, then $\rho(\tilde{x}) \leq \rho(\tilde{y})$.*
2. Translation invariance: *If $c \in \mathfrak{R}$, then $\rho(\tilde{x} + c) = \rho(\tilde{x}) - c$.*
3. Convexity: *If $\lambda \in [0, 1]$, then $\rho(\lambda\tilde{x} + (1 - \lambda)\tilde{y}) \leq \lambda\rho(\tilde{x}) + (1 - \lambda)\rho(\tilde{y})$.*

If, in addition, we have

4. Positive homogeneity: *If $\lambda \geq 0$, then $\rho(\lambda\tilde{x}) = \lambda\rho(\tilde{x})$,*

we say that ρ is a coherent risk measure.

3.1 Optimized Certainty Equivalent (OCE)

The approach we use to define risk measures is based on the optimized certainty equivalent (OCE) notion introduced by Ben-Tal and Teboulle ([5], [6]). For a random variable \tilde{x} with probability distribution \mathbb{P}_x and a normalized concave utility function u , the OCE is defined as:

$$S_u(\tilde{x}) = \sup_{v \in \mathfrak{R}} \{v + \mathbb{E}_{\mathbb{P}_x} [u(\tilde{x} - v)]\}.$$

The OCE can be interpreted as the sure present value of a future uncertain income \tilde{x} . Suppose an investor expects an uncertain future income of \tilde{x} and can consume part of it at present. If he chooses to consume v , the resulting present value is then $v + \mathbb{E}_{\mathbb{P}_x} [u(\tilde{x} - v)]$. The optimized certainty equivalent is then a result of an optimal allocation of the payoffs between present and future consumption. The OCE risk measure is defined as:

$$\rho_u(\tilde{x}) = -S_u(\tilde{x}).$$

Consider the general class of functions $u(x) : \mathfrak{R} \mapsto [-\infty, \infty)$ that are proper, closed, concave, and nondecreasing utility functions with effective domain $\text{dom}(u) = \{t \in \mathfrak{R} : u(t) > -\infty\} \neq \emptyset$. Assume that the utility function satisfies the properties

$$u(0) = 0 \quad \text{and} \quad 1 \in \partial u(0),$$

where $\partial u(\cdot)$ denotes the subdifferential map of u . It is shown in [5], that for this class of utility functions, $\rho_u(\tilde{x})$ satisfies the properties in Definition 3.1 and defines a convex risk measure. Moreover, for piecewise linear utility functions with two pieces of the form

$$u(x) = \begin{cases} \gamma_2 x, & \text{if } x \leq 0, \\ \gamma_1 x, & \text{if } x > 0, \end{cases}$$

for some $\gamma_2 > 1 > \gamma_1 \geq 0$, $\rho_u(\tilde{x})$ defines a coherent risk measure.

We now specialize this definition of OCE risk measures for the class of piecewise linear utility functions.

Definition 3.2. Let $u(x) = \min_{k \in \{1, \dots, K\}} \{a_k x + b_k\}$ be a piecewise linear concave utility function satisfying the following three properties:

1. $a_1 > a_2 > \dots > a_l \geq 1 \geq a_{l+1} > \dots > a_K \geq 0$,
2. $b_k \geq 0$ for $k \in \{1, \dots, K\}$ and $b_l = 0$,
3. $a_l > 1$ implies that $b_{l+1} = 0$.

Property 1 implies that the utility is a piecewise linear increasing function. Property 2 implies that the utility is zero at $x = 0$ and this is attained at the l th piece. Property 1, 2 and 3 ensures that the value 1 is a subgradient for the utility function at $x = 0$. For the piecewise linear utility function, the OCE measure is defined as

$$S_u(\tilde{x}) = \sup_{v \in \mathfrak{R}} \left\{ v + \mathbb{E}_{\mathbb{P}_x} \left[\min_{k \in \{1, \dots, K\}} \{a_k(\tilde{x} - v) + b_k\} \right] \right\}, \quad (3.1)$$

and the corresponding risk measure is defined as

$$\rho_u(\tilde{x}) = \inf_{v \in \mathfrak{R}} \left\{ v - \mathbb{E}_{\mathbb{P}_x} \left[\min_{k \in \{1, \dots, K\}} \{a_k(\tilde{x} + v) + b_k\} \right] \right\}. \quad (3.2)$$

For the class of utility functions satisfying Definition 3.2, this risk measure is then a convex risk measure.

3.2 Risk Measures under Ambiguous Distributions

Consider an investor who wants to evaluate the OCE risk for a random payoff $\tilde{x} = c + \mathbf{y}'\tilde{\mathbf{z}} \in \mathcal{X}$. To evaluate this for a fixed vector \mathbf{y} , the complete knowledge of the multivariate distribution of $\tilde{\mathbf{z}}$ must be

known. Suppose that the actual distribution \mathbb{P} lies in the set of distributions \mathbb{F} . The worst case OCE is defined as:

$$\hat{S}_u(c + \mathbf{y}'\tilde{\mathbf{z}}) = \sup_{v \in \mathfrak{R}} \left\{ v + \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}} - v) + b_k\} \right] \right\}, \quad (3.3)$$

while the worst case OCE risk measure is defined as:

$$\hat{\rho}_u(c + \mathbf{y}'\tilde{\mathbf{z}}) = \inf_{v \in \mathfrak{R}} \left\{ v - \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}} + v) + b_k\} \right] \right\}, \quad (3.4)$$

where the expected utility is evaluated with respect to the worst case distribution. Using the bounds developed in Section 2, the problem of finding a trading strategy that minimizes the worst case OCE risk measure in a portfolio optimization problem can be formulated as

$$\inf_{(y_0, \mathbf{y}) \in Y} \hat{\rho}_u(y_0 r + \mathbf{y}'\tilde{\mathbf{z}}).$$

By introducing the variable v in the previous formulations, the optimal portfolio can be found exactly or approximately as a compact conic program. We now provide two special cases for which the worst case OCE risk measure can be computed explicitly under mean and covariance information.

Proposition 4. *For any random portfolio variable \tilde{x} with mean μ_x and variance $\sigma_x^2 > 0$,*

(a) *The worst case OCE risk measure for the two-piece utility function $u(\tilde{x}) = \min\{a\tilde{x}, 0\}$ with $a > 1$ is given as*

$$\hat{\rho}_u(\tilde{x}) = -\mu_x + \sqrt{a-1}\sigma_x.$$

(b) *The worst case OCE risk measure for the three-piece utility function $u(\tilde{x}) = \min\{a\tilde{x} + b, \tilde{x}, 0\}$ with $a > 1$ and $b > 0$ is given as*

$$\hat{\rho}_u(\tilde{x}) = \begin{cases} -\mu_x - \frac{b}{a} + \sqrt{a-1}\sigma_x, & \text{if } \sigma_x \geq \frac{2b}{a\sqrt{a-1}}, \\ -\mu_x + \frac{a(a-1)}{4b}\sigma_x^2, & \text{otherwise.} \end{cases}$$

Proof. See Appendix A.

The worst case OCE risk measure in Proposition 4(a) is simply the worst case conditional value at risk at confidence level $1/a$ while 4(b) is a generalization to three pieces. We now show that the optimal portfolios obtained for these two cases can in fact be significantly different.

Proposition 5. *Consider a two asset model with one risk-free asset with return $r \geq 0$ and one risky asset with return \tilde{z} with mean $\mu \geq r$ and variance σ^2 . Let $y \in [0, 1]$ denote the fractional allocation in the risky asset. The optimal allocation from the worst case OCE risk minimization problem is given as*

(a) *For the two-piece utility function $u(\tilde{x}) = \min\{a\tilde{x}, 0\}$ with $a > 1$:*

$$y = \begin{cases} 1, & \text{if } \mu - \sqrt{a-1}\sigma \geq r, \\ 0, & \text{otherwise.} \end{cases}$$

(b) For the three-piece utility function $u(\tilde{x}) = \min\{a\tilde{x} + b, \tilde{x}, 0\}$ with $a > 1$ and $b > 0$:

$$y = \begin{cases} 1, & \text{if } \mu - \min\left(\sqrt{a-1}\sigma, \frac{a(a-1)\sigma^2}{2b}\right) \geq r, \\ \frac{2b(\mu-r)}{a(a-1)\sigma^2}, & \text{otherwise.} \end{cases}$$

Proof. See Appendix A.

The optimal portfolio allocations are plotted in Figures 2 and 3 respectively. One of the standard criticisms raised against worst case approaches is that the optimal portfolios could be too pessimistic. For the two piece utility function, we observe that this is indeed the case with the optimal portfolio involving no diversification. Based on the larger of the two values r and $\mu - \sqrt{a-1}\sigma$, we invest completely in the risk-free asset or the risky asset. On the other hand, it is interesting to note with a three piece utility function, we model a less risk-averse investor and obtain non-trivial diversification for the optimal portfolio under the worst case OCE risk measure. This suggests that there is merit in studying the more complicated piecewise linear utility functions and the ambiguous versions of it.

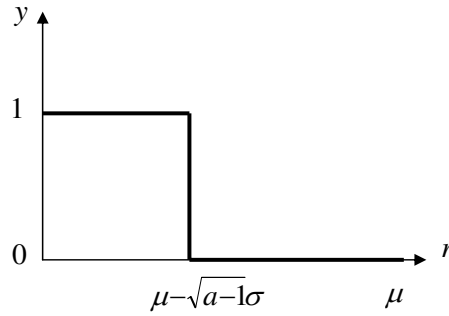


Figure 2: Optimal portfolio for the two-piece utility function.

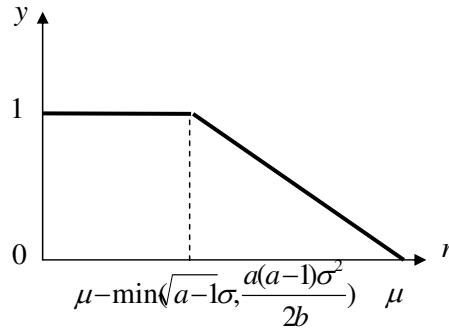


Figure 3: Optimal portfolio for the three-piece utility function.

While the OCE risk measure in (3.2) is a convex risk measure for utility functions satisfying Definition 3.2, the ambiguous risk measure violates the axiom of monotonicity. We show that with the support information, we can characterize the convexity of the ambiguous risk measures next.

Theorem 3.1. Consider the class of utility functions $u(\cdot)$ satisfying Definition 3.2. Let \mathcal{W} be a support for the random returns $\tilde{\mathbf{z}}$ (possibly bounded or unbounded) and $\pi : \mathfrak{R}^n \times \mathfrak{R}^K \mapsto \mathfrak{R}$ be a jointly concave, positive homogeneous function satisfying

$$\pi(\mathbf{y}, \mathbf{d}) \geq \min_{k \in \{1, \dots, K\}} \left\{ \inf_{\mathbf{z} \in \mathcal{W}} a_k \mathbf{y}' \mathbf{z} + d_k \right\}, \quad \forall \mathbf{y} \in \mathfrak{R}^n, \mathbf{d} \in \mathfrak{R}^K,$$

and $\pi(\mathbf{0}, \mathbf{d}) = \min_{k \in \{1, \dots, K\}} \{d_k\}$. Then the functional $\hat{\rho}_u : \mathcal{X} \mapsto \mathfrak{R}$ defined as

$$\hat{\rho}_u(c + \mathbf{y}' \tilde{\mathbf{z}}) = \inf_{v \in \mathfrak{R}} \{v - \pi(\mathbf{y}, \mathbf{a}(v + c) + \mathbf{b})\}, \quad (3.5)$$

is a convex risk measure over the space of random variables \mathcal{X} . If in addition, $\mathbf{b} = \mathbf{0}$, then the risk measure is coherent over \mathcal{X} .

Proof. See Appendix A.

In practice, due to the limited historical data, it is difficult to determine the support of a multivariate random variable. Therefore, it may be reasonable to assume unbounded support, $\mathcal{W} = \mathfrak{R}^n$. In that case, the condition in Theorem 3.1 simply reduces to $\pi(\mathbf{y}, \mathbf{d}) \geq -\infty$ for all $\mathbf{y} \neq \mathbf{0}$, which is always true. As such, the mean-covariance bound from Section 3 is a convex risk measure over random variables with unbounded support. For random variables with bounded conic representable support \mathcal{W} , using only mean-covariance bounds but neglecting support, can lead to a loss in the monotonicity property and hence the convexity of the risk measure. In this case, we propose the use of the convolution bound from Section 2.3 to obtain convex risk measures.

4 Computational Experiments

In this section, we discuss and compare the performance of the different trading strategies on real market data. The data set analyzed consists of historical daily returns for a 49 industry portfolio obtained from the Fama & French data library [16]. The portfolio consists of NYSE, AMEX and NASDAQ stocks classified by industry. These include industries such as finance, health, textiles, food and machinery. Daily return data of 2772 observations is obtained spanning a total of 11 years, from September 1, 1996 to August 31, 2007.

Consider an investor planning to invest in a portfolio of $n = 49$ risky assets. He would like to minimize the risk of his investment, while guaranteeing a certain average level of percentage returns. To ensure that the returns do not deviate greatly from the required target return, the investor rebalances his portfolio at the beginning of each half-year (September or March). We assume that no short selling is allowed. The set of allowable trading strategies is given by

$$Y = \{\mathbf{y} \in \mathfrak{R}^n : \mathbf{y} \geq \mathbf{0}, \mathbf{y}' \mathbf{e} = 1, \mathbf{y}' \boldsymbol{\mu} = \mu_t\},$$

where \mathbf{e} is a vector of ones and μ_t is a target return. It has been observed in numerous empirical tests (see Schwartz and Whitcomb [38]; Simonds et al. [39]; Conrad and Kaul [11]) that the distribution

of stock returns exhibits nonstationarity. To better reflect a time-varying distribution and avoid any potential bias due to choice of period, we adopt in our experiments a semiannual rebalancing portfolio strategy over the ten year period. At the beginning of each rebalancing period, data from the past one year is used to determine a six-month trading strategy. This semiannual rebalancing strategy is adopted for a total of 10 years.

The underlying utility model is assumed to be a linear concave utility function approximating the normalized exponential utility function,

$$u(x) = (1 - \exp(-\alpha x))/\alpha$$

with risk aversion parameter of $\alpha = 200$. The function is normalized so that the resulting OCE risk is a convex risk measure. The ten-piece linear function parameters are given in Table 1. Note that the piecewise linear function satisfies the properties of Definition 3.2.

k	a_k	b_k	k	a_k	b_k
1	1.3521	0.0002	6	0.4179	0.0011
2	1.1070	0	7	0.3178	0.0016
3	0.8848	0	8	0.2355	0.0021
4	0.6891	0.0002	9	0.1626	0.0027
5	0.5367	0.0006	10	0.1037	0.0033

Table 1: Parameters of the piecewise-linear utility function.

In the beginning of each rebalancing period, the investor chooses among the following three methods to determine his six-month trading strategy:

1. A sample-based approach (SB): The samples of returns from the one-year training set are used to construct an empirical distribution of the asset returns. The portfolio optimization problem is then solved with respect to this empirical distribution. In particular, if the N_{tr} samples in the training data set are denoted as $\{\mathbf{z}_1, \dots, \mathbf{z}_{N_{tr}}\}$, then the sample-based method for the OCE risk minimization portfolio problem solves

$$\inf_{\mathbf{y} \in Y} \inf_{v \in \mathfrak{R}} \left\{ v - \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} \min_{k \in \{1, \dots, K\}} \{a_k(v + \mathbf{y}' \mathbf{z}_i) + b_k\} \right\}.$$

2. A robust approach using mean and covariance information (MC): Each one-year training data set is used to estimate the mean $\boldsymbol{\mu}$ and covariance \mathbf{Q} of returns of the risk assets. The maximin approach then solves

$$\inf_{\mathbf{y} \in Y} \inf_{v \in \mathfrak{R}} \{v - \hat{u}_1(v + \mathbf{y}' \tilde{\mathbf{z}})\},$$

where the worst case expected utility is computed from Theorem 2.1.

3. A robust approach using partitioned statistics (PS): The one-year training data set is partitioned into the positive and negative parts, and the corresponding mean and covariance of the partitioned distribution are estimated. The maximin approach then solves

$$\inf_{\mathbf{y} \in Y} \inf_{v \in \mathfrak{R}} \{v - \hat{u}_3(v + \mathbf{y}'\tilde{\mathbf{z}})\},$$

where the bound on expected utility is computed from (2.15).

Under each approach, a 10-year dynamic trading strategy is obtained. We compare the performance of each of these strategies by comparing the realized mean and realized OCE risk. Suppose N_{ts} is the number of data points in each six-month rebalancing period. Denote the realization of returns over the 10 year period as $\{\mathbf{z}_1, \dots, \mathbf{z}_{N_{ts}}, \mathbf{z}_{N_{ts}+1}, \dots, \mathbf{z}_{20N_{ts}}\}$. For any 10-year dynamic strategy, $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_{20})$, the realized mean and OCE risk is then given by

$$\begin{aligned} \bar{\mu}(\mathbf{y}) &= \frac{1}{20N_{ts}} \sum_{i=1}^{20} \sum_{j=1}^{N_{ts}} \mathbf{y}'_i \mathbf{z}^{(i-1)N_{ts}+j}, \\ \bar{\rho}(\mathbf{y}) &= \inf_{v \in \mathfrak{R}} \left\{ v - \frac{1}{20N_{ts}} \sum_{i=1}^{20} \sum_{j=1}^{N_{ts}} \min_{k \in \{1, \dots, 5\}} \{a_k(\mathbf{y}'_i \mathbf{z}^{(i-1)N_{ts}+j} + v) + b_k\} \right\}. \end{aligned}$$

The experiment is repeated for varying target mean levels. We find the efficient frontiers of each method by plotting the realized average returns against the realized OCE risk. On each frontier, a data point corresponds to the mean return and risk level of the optimal trading strategy under a specific target mean. The computational experiments are performed using the linear and second order cone programming solvers in ILOG CPLEX 10.1. Computing each semiannual optimal trading strategy for the robust methods took less than one second.

Table 2 shows the 10-year realized OCE risk and average daily returns of the three methods under each fixed target mean. Figure 4 shows this information in terms of each method's implied efficient frontier. Under every target mean, the robust methods provide the least risky 10-year trading strategies. Moreover, implementing the partitioned statistics strategies always results in the lowest OCE risk. This is more obvious in the efficient frontiers of the three methods. The partitioned statistics efficient frontier most closely approximates the true efficient frontier of a 10-year semiannual rebalancing portfolio. One possible reason for the poor performance of the sample-based method is possible over-fitting of the distribution. Approximate methods that use in-sample data to assume the complete distribution appear to result in large errors in the out-of-sample data. These strategies however perform better for in-sample data as one would expect (see Figure 5 for two in-sample periods). Yet this improvement is only slight and the OCE risk of the sample-based method is only marginally better than those of the robust methods.

Figure 6 plots the cumulative wealth of the dynamic trading strategies over the ten-year period for the target mean daily return 0.06%. We can observe that the portfolio derived by the sample-based method gives the lowest overall cumulative wealth over the period. Of the three portfolios, the one derived by the partitioned statistics approach provides the greatest cumulative wealth. It is also

Target Mean (%)	Average Daily Returns (%)			Realized OCE Risk (%)		
	Sample-Based	Mean-Covariance	Partitioned Statistics	Sample-Based	Mean-Covariance	Partitioned Statistics
0.0400	0.0339	0.0337	0.0363	0.1843	0.1838	0.1827
0.0425	0.0338	0.0337	0.0364	0.1848	0.1840	0.1828
0.0450	0.0337	0.0340	0.0366	0.1851	0.1843	0.1829
0.0475	0.0334	0.0343	0.0370	0.1862	0.1847	0.1832
0.0500	0.0337	0.0346	0.0375	0.1869	0.1851	0.1835
0.0525	0.0342	0.0351	0.0379	0.1873	0.1854	0.1841
0.0550	0.0344	0.0357	0.0383	0.1881	0.1858	0.1847
0.0575	0.0349	0.0366	0.0391	0.1892	0.1859	0.1853
0.0600	0.0358	0.0375	0.0399	0.1897	0.1864	0.1860
0.0625	0.0369	0.0384	0.0409	0.1903	0.1871	0.1866
0.0650	0.0376	0.0394	0.0419	0.1913	0.1881	0.1875
0.0675	0.0389	0.0403	0.0429	0.1925	0.1894	0.1888
0.0700	0.0401	0.0411	0.0439	0.1937	0.1910	0.1903
0.0725	0.0409	0.0418	0.0449	0.1959	0.1930	0.1921
0.0750	0.0417	0.0427	0.0459	0.1981	0.1952	0.1943

Table 2: Realized OCE risk and average daily returns under different target means when tested over the 10-year investment horizon.

interesting to note that cumulative wealth of the three portfolios appear to moving simultaneously over the ten-year period. Based on all these observations, we can conclude that the maximin approach based on the first two moments can provide good portfolio trading strategies.

5 Acknowledgement

The authors would like to thank a former student Bi-Hui Ong at the National University of Singapore for stimulating discussions on this problem.

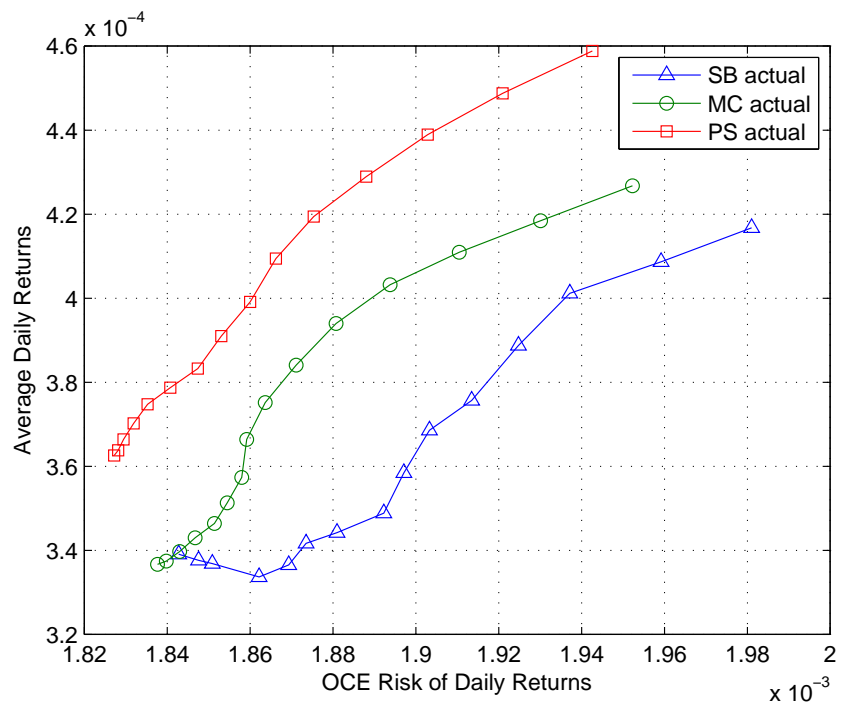
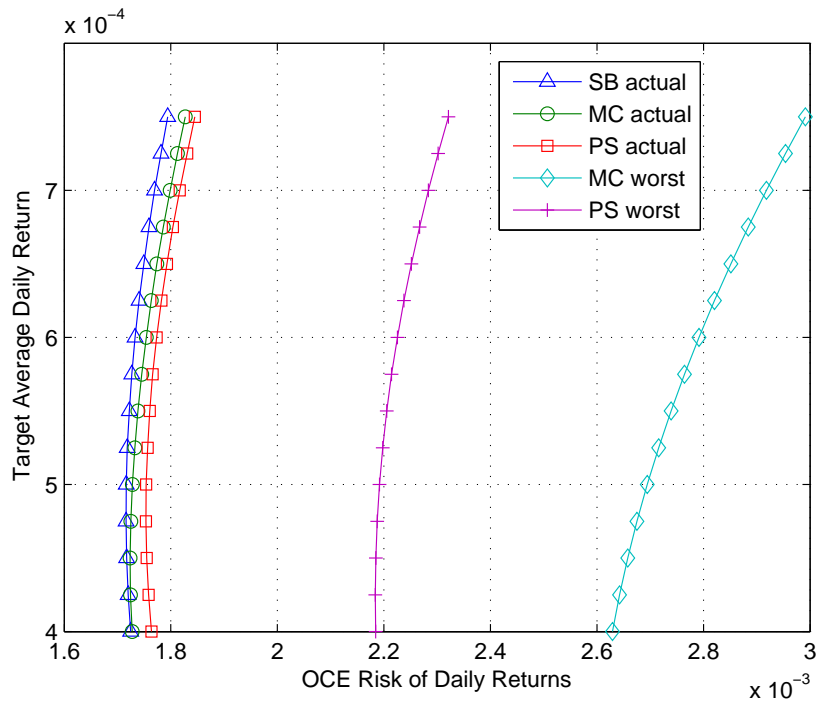
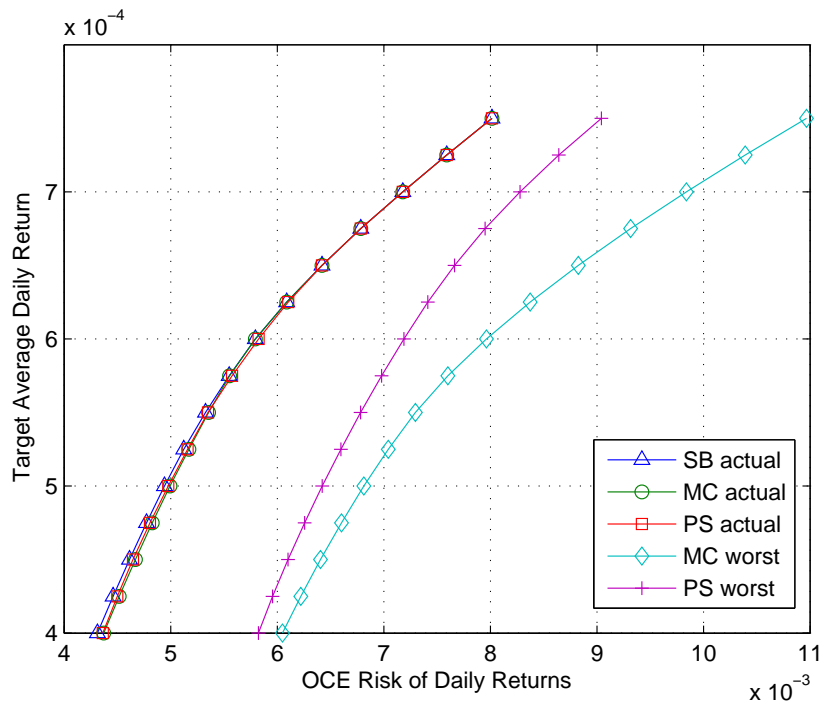


Figure 4: The efficient frontier of the rebalanced portfolio over September 1997 to August 2007.



(a) March 1998 to February 1999



(b) March 2002 to February 2003

Figure 5: The efficient frontier of optimal trading strategy during in-sample periods. The “actual” OCE risk values are taken under the empirical distribution, whereas the “worst” OCE risk values are under the worst-case distribution.

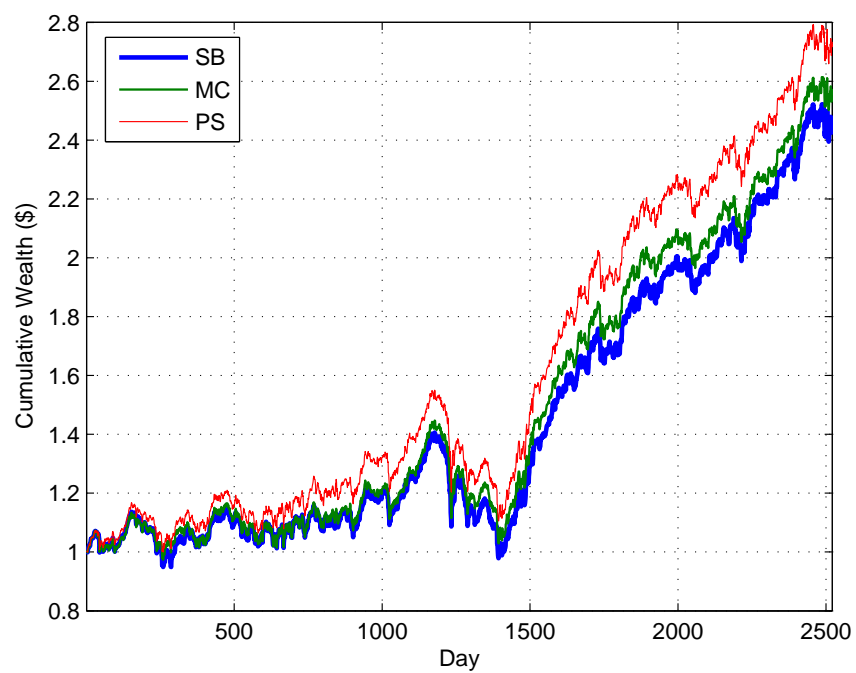


Figure 6: The cumulative wealth of the trading strategies over the period September 1997 to August 2007 for target mean return, $\mu_t = 0.06\%$.

A Proofs

Proof of Theorem 2.1: By Proposition 1, the moment problem over n random variables in \mathfrak{R}^n with given mean and covariance matrix

$$\inf_{\tilde{\mathbf{z}} \sim (\mathfrak{R}^n, \boldsymbol{\mu}, \mathbf{Q})} \mathbb{E}_{\mathbb{P}} \left[\min_k \{a_k(c + \mathbf{y}'\tilde{\mathbf{z}}) + b_k\} \right], \quad (\text{A.1})$$

is equivalent to the moment problem over a single random variable in \mathfrak{R} with given mean and variance

$$\inf_{\tilde{x} \sim (\mathfrak{R}, c + \mathbf{y}'\boldsymbol{\mu}, \mathbf{y}'\mathbf{Q}\mathbf{y})} \mathbb{E}_{\mathbb{P}_x} \left[\min_k \{a_k\tilde{x} + b_k\} \right].$$

Let $\mu_x = c + \mathbf{y}'\boldsymbol{\mu}$ and $\sigma_x^2 = \mathbf{y}'\mathbf{Q}\mathbf{y}$. The dual formulation (see Isii [20]) is given as

$$\begin{aligned} \sup_{x_0, x_1, x_2} \quad & x_0 + \mu_x x_1 + (\mu_x^2 + \sigma_x^2)x_2 \\ & x_0 - b_k + (x_1 - a_k)x + x_2 x^2 \leq 0, \quad \forall x \in \mathfrak{R}, \forall k = 1, \dots, K. \end{aligned} \quad (\text{A.2})$$

The decision variables x_0 , x_1 and x_2 are the dual variables for the probability-mass, the mean and second moment constraints respectively. From Isii [20], strong duality holds for this problem under the regularity condition that $\sigma_x^2 = \mathbf{y}'\mathbf{Q}\mathbf{y} > 0$. Without loss of generality, we assume that this condition is satisfied¹. The left hand side of the constraint in (A.2) is a quadratic function in x . For utility functions with two or more distinct linear pieces, the feasible region is thus nonempty when $x_2 < 0$. If the feasible set is nonempty, then for each of the K constraints, the maximum value in the left hand side is attained at $x^* = (a_k - x_1)/2x_2$. Therefore, (A.2) is equivalent to

$$\begin{aligned} \sup_{x_0, x_1, x_2} \quad & x_0 + \mu_x x_1 + (\mu_x^2 + \sigma_x^2)x_2 \\ \text{s.t.} \quad & x_0 - \frac{(x_1 - a_k)^2}{4x_2} - b_k \leq 0, \quad \forall k = 1, \dots, K, \\ & x_2 \leq 0. \end{aligned} \quad (\text{A.3})$$

Using the change of variables

$$\begin{aligned} x_0 &= w - \frac{(t + \mu_x)^2}{4z}, \\ x_1 &= \frac{t + \mu_x}{2z}, \\ x_2 &= \frac{-1}{4z}, \end{aligned}$$

we get the equivalent formulation for (A.3)

$$\begin{aligned} \sup_{z, t, w} \quad & w - \frac{\sigma_x^2 + t^2}{4z} \\ \text{s.t.} \quad & w \leq a_k \mu_x + b_k - a_k^2 z + a_k t, \quad \forall k = 1, \dots, K, \\ & z \geq 0. \end{aligned}$$

¹It is easy to see that the objective function reduces to $\min_k \{a_k \mu_x + b_k\}$ for $\sigma_x = 0$.

Linearizing the objective, this can be reformulated as

$$\begin{aligned}
& \sup_{z,t,w,s} w - s \\
& \text{s.t. } w \leq a_k \mu_x + b_k - a_k^2 z + a_k t, \quad \forall k = 1, \dots, K, \\
& \quad 4zs \geq \sigma_x^2 + t^2, \\
& \quad z \geq 0.
\end{aligned} \tag{A.4}$$

Formulation (A.4) can be rewritten as a standard SOCP:

$$\begin{aligned}
& \sup_{z,t,v,w,s} w - s \\
& \text{s.t. } w \leq a_k \mu_x + b_k - a_k^2 z + a_k t, \quad \forall k = 1, \dots, K, \\
& \quad z + s \geq \sqrt{\sigma_x^2 + t^2 + (z - s)^2}, \\
& \quad z \geq 0.
\end{aligned} \tag{A.5}$$

From strong conic program duality for SOCP (A.5) (see Nesterov and Nemirovski [29]), the equivalent primal formulation is:

$$\begin{aligned}
& \inf_{\lambda_k, v_0, v_1, v_2, v_3} \sum_{k=1}^K (a_k \mu_x + b_k) \lambda_k - \sigma_x v_1 \\
& \text{s.t. } \sum_{k=1}^K \lambda_k = 1, \\
& \quad \lambda_k \geq 0, \quad \forall k = 1, \dots, K, \\
& \quad v_2 - \sum_{k=1}^K a_k \lambda_k = 0, \\
& \quad v_0 + v_3 + \sum_{k=1}^K a_k^2 \lambda_k \geq 0, \\
& \quad v_0 - v_3 = -1, \\
& \quad v_0 \geq \sqrt{v_1^2 + v_2^2 + v_3^2},
\end{aligned} \tag{A.6}$$

where λ_k are the dual variables for the K inequality constraints while v_0, v_1, v_2, v_3 are the dual variables for the SOCP constraint. The optimal solution sets the value as:

$$\begin{aligned}
v_0^2 - v_3^2 &= \sum_{k=1}^K a_k^2 \lambda_k, \\
v_2 &= \sum_{k=1}^K a_k \lambda_k, \\
v_1 &= \sqrt{\sum_{k=1}^K a_k^2 \lambda_k - \left(\sum_{k=1}^K a_k \lambda_k \right)^2}.
\end{aligned}$$

Formulation (A.6) can then be solved as:

$$\begin{aligned} \inf_{\lambda_k} \quad & \sum_{k=1}^K (a_k \mu_x + b_k) \lambda_k - \sigma_x \sqrt{\sum_{k=1}^K a_k^2 \lambda_k - \left(\sum_{k=1}^K a_k \lambda_k \right)^2} \\ \text{s.t.} \quad & \sum_{k=1}^K \lambda_k = 1, \\ & \lambda_k \geq 0, \quad \forall k = 1, \dots, K. \end{aligned}$$

□

Proof of Theorem 2.2: From strong duality results of Isii [20], the problem

$$\inf_{\tilde{\mathbf{z}} \sim (\mathcal{W}, \boldsymbol{\mu})} \mathbb{E} \left[\min_{k \in \{1, \dots, K\}} \{a_k (c + \mathbf{y}' \tilde{\mathbf{z}}) + b_k\} \right],$$

is equivalent to the dual formulation

$$\begin{aligned} \sup_{s_0, \mathbf{s}} \quad & s_0 + \mathbf{s}' \boldsymbol{\mu} \\ \text{s.t.} \quad & s_0 + \mathbf{s}' \mathbf{z} \leq \min_{k \in \{1, \dots, K\}} \{a_k \mathbf{y}' \mathbf{z} + a_k c + b_k\} \quad \forall \mathbf{z} \in \mathcal{W}. \end{aligned}$$

The optimal value for s_0 is then

$$s_0 = \inf_{\mathbf{z} \in \mathcal{W}} \min_{k \in \{1, \dots, K\}} (a_k \mathbf{y} - \mathbf{s})' \mathbf{z} + a_k c + b_k$$

The dual formulation thus reduces to

$$\sup_{\mathbf{s}} \left(\mathbf{s}' \boldsymbol{\mu} + \min_{k \in \{1, \dots, K\}} \left\{ \inf_{\mathbf{z} \in \mathcal{W}} (a_k \mathbf{y} - \mathbf{s})' \mathbf{z} + a_k c + b_k \right\} \right).$$

□

Proof of Theorem 2.3: We first show that that $\hat{u}_{\mathcal{L}}(c + \mathbf{y}' \tilde{\mathbf{z}})$ is a tighter lower bound to the problem for all distributions in $\cap_{l \in \mathcal{L}} \mathbb{F}_l$. It suffices to show that for all (\mathbf{y}, \mathbf{d}) ,

$$\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d}) \geq \max_{l \in \mathcal{L}} \pi_l(\mathbf{y}, \mathbf{d}),$$

which is straightforward, since $\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d})$ is the convolution of L lower bounds and by positive homogeneity, $\pi_l(\mathbf{0}, \mathbf{0}) = 0$. For all feasible solutions to the maximization problem, the following inequality holds for all $\mathbb{P} \in \cap_{l \in \mathcal{L}} \mathbb{F}_l$ and all (\mathbf{y}, \mathbf{d})

$$\begin{aligned} \sum_{l \in \mathcal{L}} \pi_l(\mathbf{y}_l, \mathbf{d}_l) & \leq \sum_{l \in \mathcal{L}} \mathbb{E}_{\mathbb{P}} \left[\min_k \{a_k \mathbf{y}_l' \tilde{\mathbf{z}} + d_{l,k}\} \right], \\ & = \mathbb{E}_{\mathbb{P}} \left[\sum_{l \in \mathcal{L}} \min_k \{a_k \mathbf{y}_l' \tilde{\mathbf{z}} + d_{l,k}\} \right]. \end{aligned}$$

However, since

$$\sum_{l \in \mathcal{L}} \min_k \{a_k \mathbf{y}_l' \tilde{\mathbf{z}} + d_{l,k}\} \leq \min_k \left\{ \sum_{l \in \mathcal{L}} (a_k \mathbf{y}_l' \tilde{\mathbf{z}} + d_{l,k}) \right\},$$

and $\sum_{l \in \mathcal{L}} \mathbf{y}_l = \mathbf{y}$ and $\sum_{l \in \mathcal{L}} \mathbf{d}_l = \mathbf{d}$, then for all feasible solutions, we have

$$\sum_{l \in \mathcal{L}} \pi_l(\mathbf{y}_l, \mathbf{d}_l) \leq \pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d}) \leq \mathbb{E}_{\mathbb{P}} \left[\min_k \{a_k \mathbf{y}' \tilde{\mathbf{z}} + d_k\} \right].$$

This shows that $\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d})$ remains a lower bound of (2.3) for all $\mathbb{P} \in \cap_l \mathbb{F}_l$.

To see that $\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d})$ is positive homogenous, note that for all $\lambda > 0$,

$$\begin{aligned} \pi_{\mathcal{L}}(\lambda \mathbf{y}, \lambda \mathbf{d}) &= \sup_{\mathbf{y}_l, \mathbf{d}_l} \sum_{l \in \mathcal{L}} \pi_l(\mathbf{y}_l, \mathbf{d}_l) \\ &\quad \text{s.t.} \quad \sum_{l \in \mathcal{L}} \mathbf{y}_l = \lambda \mathbf{y}, \\ &\quad \quad \quad \sum_{l \in \mathcal{L}} \mathbf{d}_l = \lambda \mathbf{d}, \\ &= \sup_{\mathbf{y}_l, \mathbf{d}_l} \sum_{l \in \mathcal{L}} \pi_l(\lambda \mathbf{y}_l, \lambda \mathbf{d}_l) \\ &\quad \text{s.t.} \quad \sum_{l \in \mathcal{L}} \mathbf{y}_l = \mathbf{y}, \\ &\quad \quad \quad \sum_{l \in \mathcal{L}} \mathbf{d}_l = \mathbf{d}. \end{aligned}$$

From positive homogeneity of π_l for all l , it follows that $\pi_{\mathcal{L}}(\lambda \mathbf{y}, \lambda \mathbf{d}) = \lambda \pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d})$ for all $\lambda > 0$. Finally, the concavity of $\pi_{\mathcal{L}}$ follows immediately from the convexity of the epigraph of $-\pi_{\mathcal{L}}$ which results due to the concavity of the functions π_l ,

$$\begin{aligned} &\{(\mathbf{y}, \mathbf{d}, t) : -\pi_{\mathcal{L}}(\mathbf{y}, \mathbf{d}) \leq t\}, \\ &= \left\{ (\mathbf{y}, \mathbf{d}, t) : \exists(\mathbf{y}_l, \mathbf{d}_l), l \in \mathcal{L} : \sum_{l \in \mathcal{L}} -\pi_l(\mathbf{y}_l, \mathbf{d}_l) \leq t, \sum_{l \in \mathcal{L}} \mathbf{y}_l = \mathbf{y}, \sum_{l \in \mathcal{L}} \mathbf{d}_l = \mathbf{d} \right\}. \end{aligned}$$

□

Proof of Theorem 2.4: We can show that $\hat{u}_3(c + \mathbf{y}' \tilde{\mathbf{z}})$ is a lower bound for the worst-case expected utility under partitioned statistics. The details of the proof are similar to the proof of Theorem 2.3. By convolution, we find that for all feasible solutions of the optimization problem (2.15) and for all $(\tilde{\mathbf{z}}^+, \tilde{\mathbf{z}}^-)$ satisfying the first two moments and support \mathfrak{R}_+^{2n} ,

$$\mathbb{E}_{\mathbb{P}} \left[\min_{k \in \{1, \dots, K\}} \{a_k(c + \mathbf{y}' \tilde{\mathbf{z}}^+ - \mathbf{y}' \tilde{\mathbf{z}}^-) + b_k\} \right] \geq \bar{\pi}_1(\mathbf{y}_1^p, \mathbf{y}_1^m, \mathbf{d}_1) + \bar{\pi}_2(\mathbf{y}_2^p, \mathbf{y}_2^m, \mathbf{d}_2).$$

Using the fact that $\tilde{\mathbf{z}} = \tilde{\mathbf{z}}^+ - \tilde{\mathbf{z}}^-$, and by taking the supremum of the righthand side over all feasible solutions, we find that $\hat{u}_3(\mathbf{y}, \mathbf{a}c + \mathbf{b})$ is a lower bound of the worst-case expected utility under partitioned

statistics. We can also find a feasible solution to problem (2.15) by setting $(\mathbf{y}_1^p, \mathbf{y}_1^m, \mathbf{d}_1) = (\mathbf{y}, -\mathbf{y}, \mathbf{ac} + \mathbf{b})$ and $(\mathbf{y}_2^p, \mathbf{y}_2^m, \mathbf{d}_2) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$. From the equivalence (2.14) and $\boldsymbol{\mu} = \boldsymbol{\mu}^p - \boldsymbol{\mu}^m$, it follows that

$$\begin{aligned}\bar{\pi}_1(\mathbf{y}, -\mathbf{y}, \mathbf{ac} + \mathbf{b}) &= \bar{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}}^+ - \mathbf{y}'\tilde{\mathbf{z}}^-) \\ &= \hat{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}}).\end{aligned}$$

Since $\bar{\pi}_2(\mathbf{0}, \mathbf{0}, \mathbf{0}) = 0$, then we have $\hat{u}_3(c + \mathbf{y}'\tilde{\mathbf{z}}) \geq \hat{u}_1(c + \mathbf{y}'\tilde{\mathbf{z}})$. \square

Proof of Proposition 4: Consider the random variable \tilde{x} with mean μ_x and variance $\sigma_x^2 > 0$.

(a) From Proposition 3, the worst case OCE for the two-piece utility function, is obtained by solving the single variable maximization problem

$$\hat{S}_u(\tilde{x}) = \sup_{v \in \mathfrak{R}} \left\{ v + \frac{a}{2} \left(\mu_x - v - \sqrt{(\mu_x - v)^2 + \sigma_x^2} \right) \right\}.$$

By setting the derivative of the objective to zero, the optimal value for v is obtained as

$$v = \mu_x - \sigma_x \frac{(a-2)}{2\sqrt{a-1}}.$$

The corresponding worst case OCE is then given as

$$\hat{S}_u(\tilde{x}) = \mu_x - \sigma_x \sqrt{a-1}.$$

and the worst case OCE risk measure is

$$\hat{\rho}_u(\tilde{x}) = -\hat{S}_u(\tilde{x}) = -\mu_x + \sigma_x \sqrt{a-1}.$$

(b) From Proposition 3, the worst case OCE for the three-piece utility function is obtained by solving the single variable maximization problem

$$\hat{S}_u(\tilde{x}) = \sup_{v \in \mathfrak{R}} \{v + \hat{u}(\tilde{x} - v)\},$$

where

$$\hat{u}(\tilde{x}) = \begin{cases} \frac{1}{2} \left(\mu_x - \sqrt{\mu_x^2 + \sigma_x^2} \right), & \text{(i): if } \sigma_x^2 \leq \left(\frac{b}{a(a-1)} + \mu_x \right) \left(\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left((a+1)\mu_x + b - \sqrt{((a-1)\mu_x + b)^2 + (a-1)^2\sigma_x^2} \right), & \text{(ii): if } \sigma_x^2 \leq \left(\frac{(2a-1)b}{a(a-1)} + \mu_x \right) \left(-\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left(a\mu_x + b - \sqrt{(a\mu_x + b)^2 + a^2\sigma_x^2} \right), & \text{(iii): if } \sigma_x^2 \geq \left(\frac{(2a-1)b}{a(a-1)} + \mu_x \right) \left(\frac{b}{a(a-1)} - \mu_x \right), \\ \frac{1}{2} \left(\mu_x - \frac{a(a-1)(\mu_x^2 + \sigma_x^2)}{2b} - \frac{b}{2a(a-1)} \right), & \text{(iv): otherwise.} \end{cases}$$

We partition the four different regions (i)-(iv) in the (v, σ_x) space as indicated in Figure 7. In fact, we show next that only regions (iii) or (iv) can occur in the optimal solution.

Case (a): $\sigma_x \geq b/(a-1)$:

For $\sigma_x \geq b/(a-1)$, the worst case utility function \hat{u} takes the value in region (iii). The maximization problem is then given as

$$\sup_{v \in \mathfrak{R}} \left\{ v + \frac{1}{2} \left(a(\mu_x - v) + b - \sqrt{(a(\mu_x - v) + b)^2 + a^2\sigma_x^2} \right) \right\}.$$

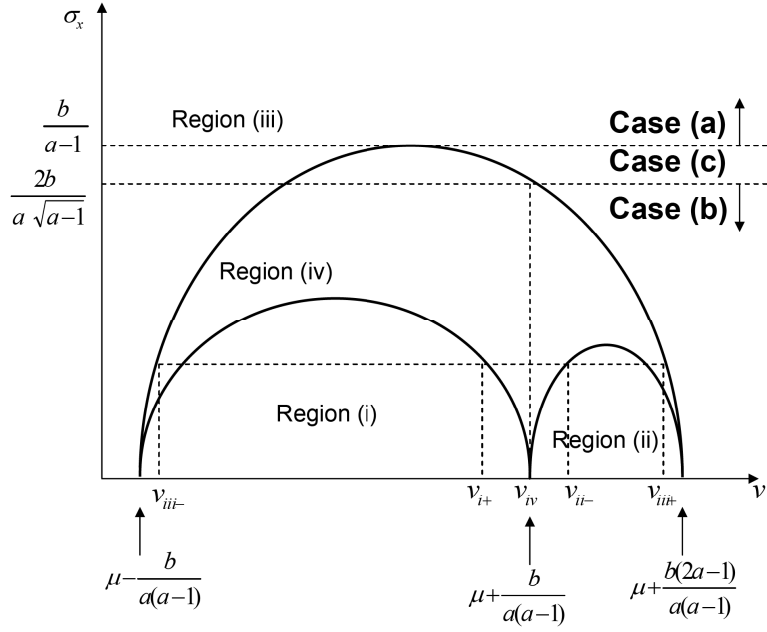


Figure 7: Characterization of the regions for three-piece utility function.

By setting the derivative to zero, the unconstrained maximum is obtained at

$$v_{iii} = \mu_x + \frac{b}{a} - \frac{(a-2)\sigma_x}{2\sqrt{a-1}},$$

with an optimal objective value of

$$\hat{S}_u(\tilde{x}) = \mu_x + \frac{b}{a} - \sqrt{a-1}\sigma_x.$$

Case (b): $\sigma_x \leq 2b/(a\sqrt{a-1})$:

For $\sigma_x \leq 2b/(a\sqrt{a-1})$, the solution obtained in case (a) is no longer feasible. The value v_{iii} in fact lies inside the outer envelope of region (iv). To check this, we first evaluate the two roots of the equation

$$\sigma_x^2 = \left(\frac{(2a-1)b}{a(a-1)} + \mu_x - v \right) \left(\frac{b}{a(a-1)} - \mu_x + v \right),$$

which is given as

$$v_{iii\pm} = \mu_x + \frac{b}{a} \pm \sqrt{\frac{b^2}{(a-1)^2} - \sigma_x^2}.$$

Thus we obtain

$$\begin{aligned} (v_{iii+} - v_{iii})(v_{iii} - v_{iii-}) &= \frac{b^2}{(a-1)^2} - \sigma_x^2 - \frac{(a-2)^2\sigma_x^2}{4(a-1)}, \\ &= \frac{a^2}{4(a-1)} \left(\frac{4b^2}{a^2(a-1)} - \sigma_x^2 \right), \\ &\geq 0. \end{aligned}$$

where the last inequality follows from the assumption that $\sigma_x \leq 2b/(a\sqrt{a-1})$. Hence in case (b), the optimal value of v over the region (iii) is attained at either v_{iii+} or v_{iii-} . This follows from the concavity

of the objective function (see Figure 7).

Over the region (i), the objective function for the worst case OCE is given as

$$\left\{ v + \frac{1}{2} \left(\mu_x - v - \sqrt{(\mu_x - v)^2 + \sigma_x^2} \right) \right\},$$

which is an increasing function in v . Hence the optimal v over region (i) is given by the larger root of the equation

$$\sigma_x^2 = \left(\frac{b}{a(a-1)} + \mu_x - v \right) \left(\frac{b}{a(a-1)} - \mu_x + v \right).$$

The corresponding v value is given as (see Figure 7)

$$v_{i+} = \mu_x + \sqrt{\frac{b^2}{a^2(a-1)^2} - \sigma_x^2},$$

with an objective value of

$$\mu_x - \frac{b - \sqrt{b^2 - \sigma_x^2 a^2 (a-1)^2}}{2a(a-1)}.$$

A similar argument shows that the worst case OCE over the region (ii) is decreasing in v . Hence the optimal v over region (ii) is given by v_{ii-} (see Figure 7). Lastly over the region (iv), the objective function for the worst case OCE is given as

$$\left\{ v + \frac{1}{2} \left(\mu_x - v - \frac{a(a-1)((\mu_x - v)^2 + \sigma_x^2)}{2b} - \frac{b}{2a(a-1)} \right) \right\}.$$

By setting the derivative to zero, the unconstrained maximum for this function is attained at

$$v_{iv} = \mu_x + \frac{b}{a(a-1)},$$

with the optimal objective value of

$$\mu_x - \frac{a(a-1)\sigma_x^2}{4b}.$$

Since this unconstrained optimal v lies in the region (iv), it is also optimal over the region (iv). Lastly, from the continuity of the objective function at the boundaries of the four different regions, it follows that the optimal objective value at v_{iv} is greater than the values at v_{iii-} , v_{i+} , v_{ii-} , v_{iii+} , proving that the minimum under this case occurs in region (iv).

Case (c): $2b/(a\sqrt{a-1}) \leq \sigma_x \leq b/(a-1)$:

In this case, v_{iii} definitely lies in region (iii) and is feasible while v_{iv} always lies in region (iii) (see Figure 7(a)). Due to concavity of the objective, this implies that the optimal value v over region (iv) is attained at either v_{iii+} or v_{iii-} . We have also previously established that the optimal v over regions (i) and (ii) are attained at v_{i+} and v_{ii-} respectively. From continuity of the objective function at the boundaries of the regions, it follows that the optimal objective value v_{iii} is greater than the values at v_{iii-} , v_{i+} , v_{ii-} , v_{iii+} . Thus, the optimal objective lies in region (iii). By combining these three cases, we get our desired result. \square

Proof of Proposition 5: The portfolio allocation problem is given as

$$\min_{0 \leq y \leq 1} \hat{\rho}_u((1-y)r + y\tilde{z}).$$

where y is the allocation in the risky asset and $1-y$ is the allocation in the risk-free asset.

(a) For the two-piece utility function, the problem to be solved is

$$\min_{0 \leq y \leq 1} -(1-y)r - y\mu + \sqrt{a-1}\sigma y.$$

The optimal portfolio is then given by the following simple rule of investing completely in a single instrument:

$$y = \begin{cases} 1, & \text{if } \mu - \sqrt{a-1}\sigma \geq r, \\ 0, & \text{otherwise.} \end{cases}$$

(b) For the three-piece utility function, the optimal worst case OCE risk minimizing portfolio is found by splitting into two cases. For $\sigma \leq \frac{2b}{a\sqrt{a-1}}$, we need to solve

$$\min_{0 \leq y \leq 1} -(1-y)r - y\mu + \frac{a(a-1)}{4b}\sigma^2 y^2.$$

The optimal solution to this problem is given as:

$$y = \begin{cases} 1, & \text{if } \mu - \frac{a(a-1)\sigma^2}{2b} \geq r, \\ \frac{2b(\mu-r)}{a(a-1)\sigma^2}, & \text{otherwise.} \end{cases}$$

For $\sigma \geq \frac{2b}{a\sqrt{a-1}}$, we need to solve

$$\min \left\{ \min_{0 \leq y \leq \frac{2b}{a\sqrt{a-1}\sigma}} -(1-y)r - y\mu + \frac{a(a-1)}{4b}\sigma^2 y^2, \min_{\frac{2b}{a\sqrt{a-1}\sigma} \leq y \leq 1} -(1-y)r - y\mu - \frac{b}{a} + \sqrt{a-1}\sigma y, \right\}$$

The optimal solution to this problem is given as:

$$y = \begin{cases} 1, & \text{if } \mu - \sqrt{a-1}\sigma \geq r, \\ \frac{2b(\mu-r)}{a(a-1)\sigma^2}, & \text{otherwise.} \end{cases}$$

Combining these two solutions, we get the optimal portfolio:

$$y = \begin{cases} 1, & \text{if } \mu - \min\left(\sqrt{a-1}\sigma, \frac{a(a-1)\sigma^2}{2b}\right) \geq r, \\ \frac{2b(\mu-r)}{a(a-1)\sigma^2}, & \text{otherwise.} \end{cases}$$

□

Proof of Theorem 3.1: To prove that $\hat{\rho}_u : \mathcal{X} \mapsto \mathfrak{R}$ is a convex risk measure, we must show that it satisfies monotonicity, translation invariance and convexity properties for all random returns in \mathcal{X} . To show monotonicity, we consider two random returns $c_1 + \mathbf{y}_1' \tilde{\mathbf{z}}, c_2 + \mathbf{y}_2' \tilde{\mathbf{z}} \in \mathcal{X}$ such that

$$c_1 + \mathbf{y}_1' \mathbf{z} \geq c_2 + \mathbf{y}_2' \mathbf{z}, \quad \forall \mathbf{z} \in \mathcal{W}.$$

Equivalently, we have

$$\inf_{\mathbf{z} \in \mathcal{W}} (c_1 - c_2 + (\mathbf{y}_1 - \mathbf{y}_2)' \mathbf{z}) \geq 0.$$

Then for all $v \in \mathfrak{R}$, we have

$$\begin{aligned} \pi(\mathbf{y}_1, \mathbf{a}(c_1 + v) + \mathbf{b}) &= \pi(\mathbf{y}_2 + (\mathbf{y}_1 - \mathbf{y}_2), \mathbf{a}(c_2 + v) + \mathbf{b} + \mathbf{a}(c_1 - c_2)), \\ &\geq \pi(\mathbf{y}_2, \mathbf{a}(c_2 + v) + \mathbf{b}) + \pi(\mathbf{y}_1 - \mathbf{y}_2, \mathbf{a}(c_1 - c_2)), \end{aligned}$$

where the inequality follows from the joint concavity and positive homogeneity of π . Observe that since $\mathbf{a} \geq \mathbf{0}$,

$$\begin{aligned} \pi(\mathbf{y}_1 - \mathbf{y}_2, \mathbf{a}(c_1 - c_2)) &\geq \min_{k \in \{1, \dots, K\}} \left\{ \inf_{\mathbf{z} \in \mathcal{W}} a_k (c_1 - c_2 + (\mathbf{y}_1 - \mathbf{y}_2)' \mathbf{z}) \right\} \\ &\geq 0 \end{aligned}$$

Hence

$$\inf_v \{v - \pi(\mathbf{y}_2, \mathbf{a}(c_2 + v) + \mathbf{b})\} \geq \inf_v \{v - \pi(\mathbf{y}_1, \mathbf{a}(c_1 + v) + \mathbf{b})\},$$

or equivalently, $\hat{\rho}_u(c_2 + \mathbf{y}_2' \tilde{\mathbf{z}}) \geq \hat{\rho}_u(c_1 + \mathbf{y}_1' \tilde{\mathbf{z}})$. This implies that the risk measure satisfies the monotonicity property. The rest of the proof is exactly as in Ben-Tal and Teboulle [6]. We add it here for completeness.

To show translation invariance, note that for any $d \in \mathfrak{R}$,

$$\begin{aligned} \hat{\rho}_u(c + \mathbf{y}' \tilde{\mathbf{z}} + d) &= \inf_v \{v - \pi(\mathbf{y}, \mathbf{a}(v + c + d) + \mathbf{b})\}, \\ &= \inf_v \{v - d - \pi(\mathbf{y}, \mathbf{a}(v + c) + \mathbf{b})\}, \\ &= \inf_v \{v - \pi(\mathbf{y}, \mathbf{a}(v + c) + \mathbf{b})\} - d, \\ &= \hat{\rho}_u(c + \mathbf{y}' \tilde{\mathbf{z}}) - d. \end{aligned}$$

To show convexity of the risk measure, note that for any $\lambda \in [0, 1]$

$$\hat{\rho}_u(\lambda(c_1 + \mathbf{y}_1' \tilde{\mathbf{z}}) + (1 - \lambda)(c_2 + \mathbf{y}_2' \tilde{\mathbf{z}})) = \inf_v \{v - \pi(\lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2, \mathbf{a}(\lambda c_1 + (1 - \lambda) c_2 + v) + \mathbf{b})\}.$$

We introduce variables v_1, v_2 such that $v = \lambda v_1 + (1 - \lambda) v_2$. From the concavity of π ,

$$\begin{aligned} &\pi(\lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2, \mathbf{a}(\lambda c_1 + (1 - \lambda) c_2 + v) + \mathbf{b}) \\ &= \pi(\lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2, \lambda(\mathbf{a}(c_1 + v_1) + \mathbf{b}) + (1 - \lambda)(\mathbf{a}(c_2 + v_2) + \mathbf{b})), \\ &\geq \lambda \pi(\mathbf{y}_1, \mathbf{a}(c_1 + v_1) + \mathbf{b}) + (1 - \lambda) \pi(\mathbf{y}_2, \mathbf{a}(c_2 + v_2) + \mathbf{b}). \end{aligned}$$

Thus,

$$\begin{aligned} &\hat{\rho}_u(\lambda(c_1 + \mathbf{y}_1' \tilde{\mathbf{z}}) + (1 - \lambda)(c_2 + \mathbf{y}_2' \tilde{\mathbf{z}})) \\ &\leq \inf_{v_1, v_2} \{\lambda(v_1 - \pi(\mathbf{y}_1, \mathbf{a}(c_1 + v_1) + \mathbf{b})) + (1 - \lambda)(v_2 - \pi(\mathbf{y}_2, \mathbf{a}(c_2 + v_2) + \mathbf{b}))\}, \\ &= \lambda \inf_{v_1} \{v_1 - \pi(\mathbf{y}_1, \mathbf{a}(c_1 + v_1) + \mathbf{b})\} + (1 - \lambda) \inf_{v_2} \{v_2 - \pi(\mathbf{y}_2, \mathbf{a}(c_2 + v_2) + \mathbf{b})\}, \\ &= \lambda \hat{\rho}_u(c_1 + \mathbf{y}_1' \tilde{\mathbf{z}}) + (1 - \lambda) \hat{\rho}_u(c_2 + \mathbf{y}_2' \tilde{\mathbf{z}}). \end{aligned}$$

Therefore, this defines a convex risk measure.

Finally, suppose $\mathbf{b} = \mathbf{0}$. For any $\lambda > 0$,

$$\begin{aligned}\hat{\rho}_u(\lambda(c + \mathbf{y}'\tilde{\mathbf{z}})) &= \inf_v \{v - \pi(\lambda\mathbf{y}, \mathbf{a}(\lambda c + v))\}, \\ &= \inf_v \{\lambda v - \pi(\lambda\mathbf{y}, \mathbf{a}\lambda(c + v))\}.\end{aligned}$$

Since π is positive homogeneous, then

$$\begin{aligned}\hat{\rho}_u(c + \lambda\mathbf{y}'\tilde{\mathbf{z}}) &= \lambda \inf_v \{v - \pi(\mathbf{y}, \mathbf{a}(c + v))\}, \\ &= \lambda \hat{\rho}_u(c + \mathbf{y}'\tilde{\mathbf{z}}).\end{aligned}$$

Therefore, if $\mathbf{b} = \mathbf{0}$, this defines a coherent risk measure. □

References

- [1] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26:1487–1503, 2002.
- [2] K. J. Arrow. Aspects of the theory of risk-bearing, 1965. Helsinki: Yrjö Hahnsson Foundation.
- [3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- [4] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
- [5] A. Ben-Tal and M. Teboulle. Expected utility, penalty functions and duality in stochastic nonlinear programming. *Management Science*, 32:1445–1466, 1986.
- [6] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.
- [7] D. Bertsimas and I. Popescu. On the relation between option and stock prices: A convex optimization approach. *Operations Research*, 50(2):358–374, 2002.
- [8] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2003.
- [9] H. Buhlmann. *Mathematical Methods in Risk Theory*. Springer, Berlin, 1970.
- [10] G. C. Calafiore. Ambiguous risk measures and optimal robust portfolios, 2007. To appear in SIAM Journal on Optimization.
- [11] J. Conrad and G. Kaul. Time-variation in expected returns. *The Journal of Business*, 61(4):409–425, 1988.
- [12] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems, 2008. Working Paper.
- [13] L. H. Ederington. Mean-variance as an approximation to expected utility maximization: Semi ex-ante results. In M. Hirschey and W. Marr, editors, *Advances in Financial Economics*, volume 1. 1995.
- [14] E. Erdoğan and G. Iyengar. Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107:37–61, 2006.
- [15] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6:429–447, 2002.
- [16] K. French and E. Fama. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
- [17] M. Frittelli and E. R. Gianin. Putting order in risk measures. *Journal of Banking and Finance*, 26:1473–1486, 2002.
- [18] I. Gilboa and D. Schmeidler. Maximin expected utility theory with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.
- [19] W. Hlawitschka. The empirical nature of Taylor-series approximations to expected utility. *The American Economic Review*, 84(3):713–719, 1994.
- [20] K. Isii. On the sharpness of Chebyshev-type inequalities. *Annals of the Institute of Statistical Mathematics*, 12:185–197, 1963.
- [21] A. J. Kelywegt, A. Shapiro, and T. Homem-De-Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12:479–502, 2001.
- [22] Y. Kroll, H. Levy, and H. Markowitz. Mean-variance versus direct utility maximization. *Journal of Finance*, 39(1):47–61, 1984.
- [23] O. Loistl. The erroneous approximation of expected utility by means of a Taylor’s series expansion: Analytic and computational results. *The American Economic Review*, 66(5):904–910, 1976.

- [24] H. Markowitz. Portfolio selection. *Journal of Finance*, 8:77–91, 1952.
- [25] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.
- [26] K. Natarajan and Z. Linyi. A mean-variance bound for a three piece linear function. *Probability in the Engineering and Informational Sciences*.
- [27] K. Natarajan, D. Pachamanova, and M. Sim. Incorporating asymmetric distributional information in robust value-at-risk optimization, 2007. Forthcoming in *Management Science*.
- [28] K. Natarajan, M. Song, and Chung Piau Teo. Persistency model and its applications in choice modeling. *Submitted for publication*.
- [29] Y. Nesterov and A. Nemirovski. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.
- [30] G. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In S. Uryasev, editor, *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Dordrecht, 2000.
- [31] G. Pflug and A. Ruszczyński. A risk measure for income processes. In G. Szegoe, editor, *Risk Measures for the 21st Century*. J. Wiley and Sons, New York, 2004.
- [32] I. Popescu. Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112, 2007.
- [33] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32:122–136, 1964.
- [34] L. B. Pulley. Mean-variance approximations to expected logarithmic utility. *Operations Research*, 31(4):685–696, 1983.
- [35] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.
- [36] W. W. Rogosinsky. Moments of non-negative mass. *Proc. Roy. Soc. London. Ser. A*, 245:1–27, 1958.
- [37] H. E. Scarf. A min-max solution to an inventory problem. In K. J. Arrow and S. Karlin, editors, *Studies in Mathematical Theory of Inventory and Production*, pages 201–209. Stanford University Press, Stanford, CA, 1958.
- [38] R. A. Schwartz and D. K. Whitcomb. Evidence on the presence and causes of serial correlation in market model residuals. *The Journal of Financial and Quantitative Analysis*, 12(2):291–313, 1977.
- [39] R. R. Simonds, L. R. LaMotte, and Jr. A. McWhorter. Testing for nonstationarity of market: An exact test and power considerations. *The Journal of Financial and Quantitative Analysis*, 21(2):209–220, 1986.
- [40] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.
- [41] R. F. Wipperfurth. Utility implications of portfolio selection and performance appraisal models. *The Journal of Financial and Quantitative Analysis*, 6(3):913–924, 1971.
- [42] W. T. Ziemba and R. G. Vickson editors. *Stochastic optimization models in finance*. New Jersey : World Scientific, New Jersey, 2006.