Tractable Robust Expected Utility and Risk Models for Portfolio Optimization

Karthik Natarajan* Melvyn Sim† Joline Uichanco‡§

Submitted: March 13, 2008. Revised: September 25, 2008, December 5, 2008

Abstract

Expected utility models in portfolio optimization are 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 restrictions 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 (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. Extensions of the model to capture asymmetry using partitioned statistics information and box-type uncertainty in the mean and covariance matrix are provided. Using the optimized certainty equivalent framework, 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 results for the worst-case OCE risk measures and optimal portfolios are provided for two and three-piece utility functions. For more complicated 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 expected utility maximization model (see Von-Neumann and Morgenstern [36]) provides a natural framework for a

*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 and NUS Risk Management Institute.

†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.

‡Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139. Email: joline.uichanco@gmail.com. The research was done while the author was a student in the Computational Engineering Programme, Singapore-MIT Alliance, National University of Singapore, Singapore 117576.

§We would like to thank a former student Ong Bi-Hui at the National University of Singapore for obtaining preliminary results on this problem. We are thankful to the co-editor, Jerome Detemple and two anonymous referees for valuable suggestions on how to improve the article.

rational investor to choose a portfolio allocation. Under standard assumptions on the utility function such as monotonicity and concavity, the optimal portfolio is well characterized in complete markets in both the single and multi-period settings (see Arrow and Debreu [1], Merton et. al. [21], Ocone and Karatzas [27], Detemple et. al. [11]). Using a second order-order Taylor series approximation for the utility results in a quadratic portfolio optimization problem in the spirit of Markowitz's [20] model. Portfolio selection under the Markowitz model is based only on the mean and covariance information of the uncertain returns. Despite it's popularity, the quadratic utility model suffers from behavioral limitations such as non-monotonicity (see Wipern [37]) or strong distributional assumptions such as normality.

A related issue in portfolio optimization models is the quantification of risk. Value-at-risk (VaR) is one such risk metric that 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 Rockafellar and Uryasev [31], Pflug [28], Acerbi and Tasche [2], Föllmer and Schied [13], Pflug and Ruszczynski [29]). An important extension is based on the relaxation of the positive homogeneity and subadditivity property into the weaker convexity property leading to the class of *convex risk measures* (see Föllmer and Schied [13], Fritelli and Gianin [15]). 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 [7] and Ben-Tal and Teboulle [5], [6]. For instance in [6], the negative of the *optimized certainty equivalent* has been shown to define a convex risk measure for a class of utility functions. Risk measures, such as conditional value-at-risk (Rockafellar and Uryasev [31]) and bounded shortfall risk (Föllmer and Schied [13]), can in fact be derived as special cases of the OCE.

Even if we can address the issue of accurately modeling the investor's utility function and risk preferences, there is an implicit assumption that is often made in these models. The investor is assumed to be in possession of a market model that accurately describes the distribution of the future random returns. Practitioners are however typically faced with ambiguity in the knowledge of the distribution. Aversion to ambiguity is well documented in the famous Ellsberg paradox with most people preferring to bet on an urn with 50 red and 50 blue balls, than in an urn with 100 balls containing an unknown number of red or blue balls. Under ambiguity-aversion, the investor can choose to maximize the minimum expected payoff over a set of possible distributions (see Gilboa and Schmeidler [17]). The uncertainty in the distribution is typically captured through information on partial moments such as the mean and covariance matrix. Popescu [30] studies the problem of deriving solutions to the single period robust (maximin) expected utility problem based on mean and covariance information. For a class of utility functions, she shows that the robust problem reduces to solving a parametric quadratic program. Our results in this paper closely relates to her work with stronger results for a subclass of the utility functions. Garlappi et. al. [16] derive closed form expressions for the optimal portfolios under an ambiguity-averse mean-variance model where the expected returns are assumed to lie within a confidence interval of it's estimated values. Tests on real market data therein indicate that the ambiguity-averse portfolios provide better out-of-sample performance and are more stable over time as compared to classical portfolios. In this paper, we do not restrict our attention to quadratic utilities. Interestingly, our computational results provide similar qualitative insights providing further justification for the use of such robust methods in practise. In a similar spirit, Calafiore [8] solves the mean-variance and mean-absolute deviation model with the true distribution assumed to lie within a distance from the nominal distribution with distance measured in the Kullback-Leibler divergence measure. A related model proposed in Korn and Menkens [19] assumes that the stock price follows a Black-Scholes type diffusion with a crash that can happen at an unknown time with an unknown magnitude. In their model, the investor determines the portfolio that maximizes the worst-case expected utility of the terminal wealth. In contrast, our approach to modeling ambiguity in distribution does not make any assumption on the form of the distribution. We however focus exclusively on single period models.

In this paper, we assume that the investor's utility is represented by a piecewise-linear concave function. Our contributions can then 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: (a) known mean and covariance information, and (b) known mean and support information. We present a lower bound (not necessarily tight) under distributions of: (c) known mean, covariance and support information. 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 lower bound using: (d) partitioned statistics information which captures asymmetry information in the distribution. Lastly, we find the tightest possible lower bound for the worst-case expected utility with: (e) box-type uncertainty in the mean and covariance matrix themselves. All bounds and the corresponding optimal portfolios are found by solving conic programs, specifically linear, second order cone and semidefinite programs. Such conic optimization problems can be solved very efficiently in both theory and practice using interior point methods (see Nesterov and Nemirovski [26]).

2. In Section 3, we provide a connection of our 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. For a single risk-free and a single risky asset, the optimal portfolios are very different with no diversification for two-piece utility and diversification for the three-piece utility. Connections with convex risk measures is provided based on our results in Section 2.

3. In Section 4, we perform computational experiments on real financial market data to compare the robustness of both the worst-case OCE and sample based OCE methods. Our results indicate

that an optimal rebalancing portfolio derived using the worst-case approach performs better than sample-based approaches in out-of-sample data. The method appears to be robust across different periods of the investment horizon.

2 Expected Utility under Ambiguous Distributions

Let $\tilde{\boldsymbol{z}} = (\tilde{z}_1, \ldots, \tilde{z}_n)$ 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:

$$\mathcal{X} = \{\tilde{x} \ : \ \exists (c, \boldsymbol{y}) \in \Re \times \Re^n \text{ such that } \tilde{x} = c + \boldsymbol{y}'\tilde{\boldsymbol{z}}\}.$$

We denote by $\mathcal{W}$ the support of the random variable $\tilde{\boldsymbol{z}}$. 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 $\tilde{x}_1 = c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}}$ and $\tilde{x}_2 = c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}}$, then $\tilde{x}_1 \geq \tilde{x}_2$ is equivalent to

$$c_1 + \boldsymbol{y_1}'\boldsymbol{z} \geq c_2 + \boldsymbol{y_2}'\boldsymbol{z} \qquad \forall \boldsymbol{z} \in \mathcal{W}. \tag{2.1}$$

It is easy to see that the feasibility of (2.1) remains unchanged if we replace $\mathcal{W}$ by its convex hull (see Ben-Tal and Nemirovski [4]). Hence, without loss of the generality, we assume that the support of $\tilde{\boldsymbol{z}}$ is convex. We define the class of utility functions that we study in this paper next.

Definition 2.1. *The utility function $u(\cdot) : \Re \rightarrow \Re$ is a piecewise-linear concave function defined as:*

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

and satisfies the following two properties:

1. *The number of linear pieces $K \geq 2$,*
2. *Each piece $k \in \{1, \ldots, K\}$ defines the utility function uniquely for at least one value of x.*

The expected utility is then given as

$$\mathrm{E}_{\mathbb{P}}\left(\min_{k \in \{1,\ldots,K\}} \left\{a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k\right\}\right).$$

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 $\tilde{\boldsymbol{z}}$. Even when completely specified, evaluating the 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 known set of partial moments. The worst-case expected utility is then defined as

$$\hat{u}(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \inf_{\mathbb{P} \in \mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\min_{k \in \{1,\ldots,K\}} \left\{a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k\right\}\right). \tag{2.2}$$

In the portfolio optimization context, $\tilde{\boldsymbol{z}}$ represents the random risky payoffs of n assets and $\boldsymbol{y}$ denotes the allocation vector. In addition, let $r \geq 0$ denote 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

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

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

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

where $\boldsymbol{e}$ is a vector of ones. For a fixed distribution $\mathbb{P}$, the expected utility maximization problem is formulated as

$$\sup_{(y_0,\boldsymbol{y})\in Y} \mathrm{E}_{\mathbb{P}}\left(\min_{k\in\{1,\ldots,K\}} \left\{a_k(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k\right\}\right)$$

To achieve robustness over a set of ambiguous distributions, the investor chooses to maximize his worst-case expected utility. This strategy is consistent with that of an ambiguity-averse investor, since it guarantees a minimum threshold on the expected utility. The robust expected utility problem is then formulated as:

$$\sup_{(y_0,\boldsymbol{y})\in Y} \hat{u}(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \sup_{(y_0,\boldsymbol{y})\in Y} \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\min_{k\in\{1,\ldots,K\}} \left\{a_k(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k\right\}\right). \tag{2.3}$$

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

Solving the portfolio problem (2.3) can be also used as a heuristic method for finding an investment strategy for the robust problem under more general nonlinear utilities. Intuitively, more the number of linear pieces, better is the piecewise-linear approximation and closer the heuristic solution would be to the true optimal solution. In Section 4, we provide numerical results to indicate that (2.3) can be solved in a matter of seconds to provide highly accurate solutions for nonlinear utilities using standard conic programming solvers. The next proposition relates the quality of the heuristic solution with the quality of the approximation of the utility function.

Proposition 1. *Suppose $f(\cdot) : \Re \mapsto \Re$ is a general nonlinear concave utility function. If $u(\cdot) : \Re \mapsto \Re$ is a piecewise-linear function such that*

$$f(x) - \epsilon \leq u(x) \leq f(x), \quad \forall x \in \Re,$$

with

$$\begin{aligned} Z_{opt} &= \sup_{(y_0,\boldsymbol{y})\in Y} \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(f(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}})\right), \\ (\hat{y}_0, \hat{\boldsymbol{y}}) &= \operatorname{argsup}_{(y_0,\boldsymbol{y})\in Y} \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(u(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}})\right), \\ Z_h &= \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(f(\hat{y}_0 r + \hat{\boldsymbol{y}}'\tilde{\boldsymbol{z}})\right). \end{aligned}$$

Then, we have

$$Z_{opt} - \epsilon \leq Z_h \leq Z_{opt}.$$

Proof. See Appendix A.

2.1 Mean and Covariance Information

Assume that for the random returns $\tilde{\boldsymbol{z}}$, only the mean vector $\boldsymbol{\mu} = \mathrm{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}})$ and covariance matrix $\boldsymbol{Q} = \mathrm{E}_{\mathbb{P}}((\tilde{\boldsymbol{z}} - \boldsymbol{\mu})(\tilde{\boldsymbol{z}} - \boldsymbol{\mu})')$ are explicitly known. For this class of distributions, Popescu [30] 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 2 (Popescu [30])**.** *Let $\mathbb{F}_1$ be the family of all distributions $\mathbb{P}$ for $\tilde{\boldsymbol{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$. For any utility function u, we have*

$$\inf_{\mathbb{P} \in \mathbb{F}_1} \mathrm{E}_{\mathbb{P}} \left(u(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) \right) = \inf_{\mathbb{P}_x \in \mathbb{F}_x} \mathrm{E}_{\mathbb{P}_x} \left(u(\tilde{x}) \right),$$

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

Proposition 2 is based on the following projection property: for any random variable $\tilde{x}$ with mean $c + \boldsymbol{y}'\boldsymbol{\mu}$ and variance $\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}$, there exists a random vector $\tilde{\boldsymbol{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$. 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 [32]): 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 class of utility functions, Popescu [30] proposes the use of a parametric quadratic program to find the portfolio that maximizes the worst-case expected utility. We state the result next.

Proposition 3 (Popescu [30])**.** *Let $\mathbb{F}_1$ be the family of all distributions $\mathbb{P}$ for $\tilde{\boldsymbol{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$ and $\mathbb{F}_x$ be the family of all univariate distributions for $\tilde{x}$ with mean $\mu_x = y_0 r + \boldsymbol{y}'\boldsymbol{\mu}$ and variance $\sigma_x^2 = \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}$. Suppose the objective function $\inf_{\mathbb{P}_x \in \mathbb{F}_x} \mathrm{E}_{\mathbb{P}_x} \left(u(\tilde{x}) \right)$, is continuous, nondecreasing in μ_x, nonincreasing in σ_x and quasi-concave in (μ_x, σ_x). Then the robust expected utility model is equivalent to solving a parametric quadratic program (PQP):*

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

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

Utility functions for which Proposition 3 holds 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 provide our first key result that shows for the class of piecewise-linear concave utility functions (not necessarily increasing or with the one or two point support structure), the robust expected utility problem can in fact be solved as a single second order cone program (SOCP). This provides a significant computational advantage over using PQP wherein one would need to solve multiple instances 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 $\tilde{\boldsymbol{z}}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$. The worst-case expected utility:*

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

is given as the optimal objective value to the problem:

$$\hat{u}_1(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \sup_{z\geq 0,t} \left(\min_{k\in\{1,\dots,K\}} \left(a_k(c + \boldsymbol{y}'\boldsymbol{\mu}) + b_k - a_k^2 z + a_k t\right) - \left(\frac{\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y} + t^2}{4z} \right) \right), \tag{2.4}$$

or equivalently:

$$\begin{aligned} \hat{u}_1(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \inf_{\lambda_k} \quad & \sum_{k=1}^{K} (a_k(c + \boldsymbol{y}'\boldsymbol{\mu}) + b_k)\lambda_k - \sqrt{\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}} \sqrt{\sum_{k=1}^{K} a_k^2 \lambda_k - \left(\sum_{k=1}^{K} a_k \lambda_k \right)^2} \\ s.t. \quad & \sum_{k=1}^{K} \lambda_k = 1, \\ & \lambda_k \geq 0, \qquad \forall k = 1, \dots, K. \end{aligned} \tag{2.5}$$

Proof. See Appendix A.

A natural implication of Theorem 2.1 is that the robust expected utility model in (2.3) under known 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. This problem is formulated as:

$$\begin{aligned} \sup_{z,t,w,s,y_0,\boldsymbol{y}} \quad & w - s \\ \text{s.t.} \quad & w \leq a_k(y_0 r + \boldsymbol{y}'\boldsymbol{\mu}) + b_k - a_k^2 z + a_k t, \qquad \forall k = 1, \dots, K, \\ & 4zs \geq \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y} + t^2, \\ & z \geq 0, \\ & (y_0, \boldsymbol{y}) \in Y. \end{aligned} \tag{2.6}$$

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

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

The variables λ_k in (2.5) can be interpreted as the probability that the kth piece of the utility function is chosen. Suppose, $\lambda_k = 1$, then we simply get the expected utility for the kth 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. [25]. The result here extends the approach to robust portfolio optimization problems. In special cases for two and three-piece utility functions, Theorem 2.1 can be solved in closed form (see Figure 1).

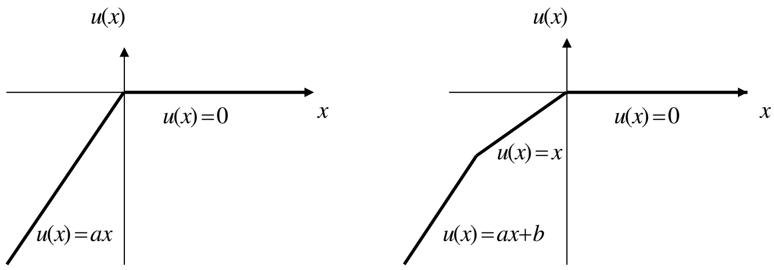


Figure 1: Two and three-piece utility functions.

Proposition 4 (Cauchy-Schwarz inequality, Natarajan and Zhou [23])**.** *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 $a > 0$ 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 $u(\tilde{x}) = \min\{a\tilde{x} + b, \tilde{x}, 0\}$ with $a > 1$ and $b > 0$ is given as

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

The bound in Proposition 4(a) is obtained by solving

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

and reduces to a version of the Cauchy-Schwarz inequality. The bound in Proposition 4(b) is obtained by solving

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

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 [23] for a proof of this result. Theorem 2.1 generalizes these results to arbitrary piecewise-linear concave functions.

2.2 Mean and Support Information

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

$$\mathcal{W} = \{\boldsymbol{z} \ : \ \boldsymbol{D}\boldsymbol{z} + \boldsymbol{F}\boldsymbol{u} - \boldsymbol{g} \in \boldsymbol{K} \text{ for some } \boldsymbol{u}\},$$

where $\boldsymbol{K}$ 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 defined as

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

is also a regular cone.

Theorem 2.2. *Let $\mathbb{F}_2$ be the family of all distributions $\mathbb{P}$ for $\tilde{\boldsymbol{z}}$ 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:*

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

is given as the optimal objective value to the problem:

$$\hat{u}_2(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \sup_{\boldsymbol{s}} \left(\boldsymbol{s}'\boldsymbol{\mu} + \min_{k\in\{1,\ldots,K\}} \left\{ \inf_{\boldsymbol{z}\in\mathcal{W}} (a_k\boldsymbol{y} - \boldsymbol{s})'\boldsymbol{z} + a_k c + b_k \right\} \right). \tag{2.7}$$

Proof. See Appendix A.

An implication of Theorem 2.2 is that the robust expected utility problem in (2.3) under given mean and bounded support can be solved as a 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 $\boldsymbol{z} \in \mathcal{W}$. The robust expected utility model is then reformulated as the conic program:

$$\begin{array}{lll} \sup\limits_{\boldsymbol{s},t,\boldsymbol{w}^{(\boldsymbol{k})},y_0,\boldsymbol{y}} & \boldsymbol{s}'\boldsymbol{\mu} + t & \\ \text{s.t.} & t \leq \boldsymbol{w}^{(\boldsymbol{k})\prime}\boldsymbol{g} + a_k y_0 r + b_k, & \forall k = 1,\ldots,K, \\ & \boldsymbol{D}'\boldsymbol{w}^{(\boldsymbol{k})} = a_k\boldsymbol{y} - \boldsymbol{s}, & \forall k = 1,\ldots,K, \\ & \boldsymbol{F}'\boldsymbol{w}^{(\boldsymbol{k})} = \boldsymbol{0}, & \forall k = 1,\ldots,K, \\ & \boldsymbol{w}^{(\boldsymbol{k})} \in \boldsymbol{K}^*, & \forall k = 1,\ldots,K, \\ & (y_0, \boldsymbol{y}) \in Y. & \end{array} \tag{2.8}$$

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 typically a NP-hard problem (see Murty and Kabadi [22]). For example, characterizing distributions over $\mathcal{W} = \Re^n_+$ 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 $\pi_l(\boldsymbol{y}, \boldsymbol{d}) : \Re^n \times \Re^K \mapsto \Re$ to denote the worst-case expected utility over a set of distributions $\mathbb{F}_l$:

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

Observe that $\pi_l(\boldsymbol{y}, \boldsymbol{a}c + \boldsymbol{b}) = \hat{u}_l(c + \boldsymbol{y}'\tilde{\boldsymbol{z}})$ reduces to the utility bounds derived in the previous sections for the appropriate set of distributions $\mathbb{F}_l$. The reason for introducing $\pi_l(\boldsymbol{y}, \boldsymbol{d})$ 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 $\{\pi_l(\boldsymbol{y}, \boldsymbol{d}),\ l \in \mathcal{L}\}$ be jointly concave, positive homogenous functions that each denote the worst-case expected utility over the family of distributions $\mathbb{F}_l$. Then:*

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

where

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

and

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

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 robust expected utility model under mean, covariance and support information by combining the results from Sections 2.1-2.2:

$$\begin{array}{rll}
\sup\limits_{z_1,t_1,w_1,s_1,\boldsymbol{y_1},\boldsymbol{s_2},t_2,\boldsymbol{w_2^{(k)}},\boldsymbol{d_1},\boldsymbol{d_2},\boldsymbol{y_2},y_0,\boldsymbol{y}} & w_1 - s_1 + \boldsymbol{s_2}'\boldsymbol{\mu} + t_2 & \\
\text{s.t.} & w_1 \leq a_k\boldsymbol{\mu}'\boldsymbol{y_1} + d_{1,k} - a_k^2 z_1 + a_k t_1, & \forall k = 1,\ldots,K, \\
& 4z_1 s_1 \geq \boldsymbol{y_1}'\boldsymbol{Q}\boldsymbol{y_1} + t_1^2, & \\
& z_1 \geq 0, & \\
& t_2 \leq \boldsymbol{w_2^{(k)}}'\boldsymbol{g} + d_{2,k}, & \forall k = 1,\ldots,K, \\
& \boldsymbol{D}'\boldsymbol{w_2^{(k)}} = a_k\boldsymbol{y_2} - \boldsymbol{s_2}, & \forall k = 1,\ldots,K, \\
& \boldsymbol{F}'\boldsymbol{w_2^{(k)}} = \boldsymbol{0}, & \forall k = 1,\ldots,K, \\
& \boldsymbol{w_2^{(k)}} \in \boldsymbol{K}^*, & \forall k = 1,\ldots,K, \\
& \boldsymbol{y_1} + \boldsymbol{y_2} = \boldsymbol{y}, & \\
& \boldsymbol{d_1} + \boldsymbol{d_2} = \boldsymbol{a}y_0 r + \boldsymbol{b}, & \\
& (y_0, \boldsymbol{y}) \in Y. &
\end{array} \tag{2.9}$$

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

2.4 Partitioned Statistics Information

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 $\tilde{\boldsymbol{z}}$ can be expressed as

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

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

$$\mathrm{E}_{\mathbb{P}}\left(\min_{k\in\{1,\ldots,K\}} \left\{a_k\left(c + \boldsymbol{y^p}'\tilde{\boldsymbol{z}}^+ + \boldsymbol{y^m}'\tilde{\boldsymbol{z}}^-\right) + b_k\right\}\right). \tag{2.10}$$

These partitioned random variables $\tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-$ naturally imply a positive support (i.e, $\Re_+^{2n}$). Moreover, partitioning into $\tilde{\boldsymbol{z}}^+$ and $\tilde{\boldsymbol{z}}^-$ 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:

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

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

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

where

$$\text{cov}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \stackrel{\Delta}{=} \text{E}\left((\tilde{\boldsymbol{x}} - \text{E}(\tilde{\boldsymbol{x}}))(\tilde{\boldsymbol{y}} - \text{E}(\tilde{\boldsymbol{y}}))'\right),$$

and

$$\text{var}(\tilde{\boldsymbol{x}}) \stackrel{\Delta}{=} \text{cov}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}).$$

Moreover,

$$\text{var}(\boldsymbol{y}'\tilde{\boldsymbol{z}}) = \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y} = \text{var}(\boldsymbol{y}'\tilde{\boldsymbol{z}}^+ - \boldsymbol{y}'\tilde{\boldsymbol{z}}^-) = (\boldsymbol{y}' \ \ -\boldsymbol{y}')\bar{\boldsymbol{Q}}\begin{pmatrix} \boldsymbol{y} \\ -\boldsymbol{y} \end{pmatrix}. \tag{2.11}$$

Hence,

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

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 (2.10) using Theorems 2.1 and 2.2. From the mean and covariance information of the partitioned returns, we can establish the bound

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

in which

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

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

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

where

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

Then by convolution, we can derive the partitioned statistics lower bound for the worst-case expected utility in (2.2), 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 $\tilde{\boldsymbol{z}}$ with known mean $(\boldsymbol{\mu^p}, \boldsymbol{\mu^m})$ and covariance $(\bar{\boldsymbol{Q}})$ 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 $\boldsymbol{Q} = \bar{\boldsymbol{Q}}_{11} - 2\bar{\boldsymbol{Q}}_{12} + \bar{\boldsymbol{Q}}_{22}$. We define a new lower bound:*

$$\hat{u}_3(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \pi_3(\boldsymbol{y}, \boldsymbol{a}c + \boldsymbol{b}) \tag{2.12}$$

where

$$\begin{aligned} \pi_3(\boldsymbol{y}, \boldsymbol{d}) = \sup_{\boldsymbol{y_1^p}, \boldsymbol{y_1^m}, \boldsymbol{y_2^p}, \boldsymbol{y_2^m}, \boldsymbol{d_1}, \boldsymbol{d_2}} \quad & \bar{\pi}_1(\boldsymbol{y_1^p}, \boldsymbol{y_1^m}, \boldsymbol{d_1}) + \bar{\pi}_2(\boldsymbol{y_2^p}, \boldsymbol{y_2^m}, \boldsymbol{d_2}) \\ s.t. \quad & \boldsymbol{y_1^p} + \boldsymbol{y_2^p} = \boldsymbol{y}, \\ & \boldsymbol{y_1^m} + \boldsymbol{y_2^m} = -\boldsymbol{y}, \\ & \boldsymbol{d_1} + \boldsymbol{d_2} = \boldsymbol{d}, \end{aligned}$$

which satisfies

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

Proof. See Appendix A.

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

2.5 Box-Type Uncertainty in Mean and Covariance Information

The models developed thus far in this paper have assumed that the moment information is known exactly. However, in some cases, it is desirable to specify uncertainty sets on the moments themselves. For instance, Delage and Ye [10] study a worst-case model assuming that the true mean vector lies in an ellipsoidal set centered around an estimated mean and the true covariance matrix lie in the intersection of two positive semidefinite cones. A related uncertainty set proposed by Tütüncü and Koenig [35] assumes that the moments lie in a box region specified by upper and lower bounds. Our next results shows that the worst-case expected utility under the box-type uncertainty model in the mean and covariance information can be found by solving a semidefinite program (SDP) instead of a second order cone program (SOCP).

Theorem 2.5. *Let $\mathbb{F}_{1s}$ be the family of all distributions $\mathbb{P}$ for $\tilde{\boldsymbol{z}}$ with $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$ lying in the uncertainty set:*

$$\mathcal{U} = \{(\boldsymbol{\mu}, \boldsymbol{Q}) \ : \ \underline{\boldsymbol{\mu}} \leq \boldsymbol{\mu} \leq \overline{\boldsymbol{\mu}}, \ \underline{\boldsymbol{Q}} \leq \boldsymbol{Q} \leq \overline{\boldsymbol{Q}}, \ \boldsymbol{Q} \succeq \boldsymbol{0}\}.$$

Suppose there exists a positive definite matrix lying in this uncertainty set and $a_k \geq 0$ for all $k = 1, \ldots, K$, then the worst-case expected utility:

$$\hat{u}_{1s} = \inf_{\mathbb{P} \in \mathbb{F}_{1s}} \mathrm{E}_{\mathbb{P}} \left(\min_{k=1,\ldots,K} \{ a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k \} \right)$$

is given as the optimal objective value to the problem:

$$
\begin{array}{rl}
\sup\limits_{z,t,w,s,p,v,\underline{\boldsymbol{\omega}},\overline{\boldsymbol{\omega}},\underline{\boldsymbol{\Sigma}},\overline{\boldsymbol{\Sigma}}} & w - s \\
s.t. & w \le a_k(c+v) + b_k - a_k^2 z + a_k t, \quad \forall k = 1,\dots,K, \\
& v \le \underline{\boldsymbol{\mu}}'\underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\mu}}'\overline{\boldsymbol{\omega}}, \\
& z + s - p \ge \sqrt{(z-s+p)^2 + t^2}, \\
& p \ge \overline{\boldsymbol{Q}} \bullet \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{Q}} \bullet \underline{\boldsymbol{\Sigma}}, \\
& \underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\omega}} = \boldsymbol{y}, \\
& \begin{pmatrix} 4z & \boldsymbol{y}' \\ \boldsymbol{y} & \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{\Sigma}} \end{pmatrix} \succeq \mathbf{0}, \\
& z \ge 0, \ \underline{\boldsymbol{\omega}}, \overline{\boldsymbol{\omega}} \ge \mathbf{0}, \ \underline{\boldsymbol{\Sigma}}, \overline{\boldsymbol{\Sigma}} \ge \mathbf{0}.
\end{array}
\tag{2.13}
$$

Proof. See Appendix A.

3 Connection to Risk Measures

In this section, we will use the results in Section 2 to define new ambiguous risk measures. Based on the previous bounds, we can obtain approximations to the *worst-case risk* (see Erdoğan and Iyengar [12], Calafiore [8]) 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 a 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 (see Föllmer and Schied [13], Fritelli and Gianin [15]). Consider the random outcome $\tilde{x} = c + \boldsymbol{y}'\tilde{\boldsymbol{z}} \in \mathcal{X}$ which represents the uncertain payoff. The axiomatic characterization of convex and coherent risk measures is then given as:

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

1. Monotonicity*: If $\tilde{x} \ge \tilde{y}$, then $\rho(\tilde{x}) \le \rho(\tilde{y})$.*

2. Translation invariance*: If $c \in \Re$, then $\rho(\tilde{x} + c) = \rho(\tilde{x}) - c$.*

3. Convexity*: If $\lambda \in [0,1]$, then $\rho(\lambda\tilde{x} + (1-\lambda)\tilde{y}) \le \lambda\rho(\tilde{x}) + (1-\lambda)\rho(\tilde{y})$.*

If, in addition, we have

4. Positive homogeneity*: If $\lambda \ge 0$, 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 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:

$$S_u(\tilde{x}) = \sup_{v \in \Re} \left(v + \mathrm{E}_{\mathbb{P}_x} \left(u(\tilde{x} - v) \right) \right).$$

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 $v + \mathrm{E}_{\mathbb{P}_x}(u(\tilde{x} - v))$. 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 $u(x) : \Re \mapsto [-\infty, \infty)$ that are proper, closed, concave, and nondecreasing utility functions with effective domain $\mathrm{dom(u)} = \{t \in \Re \; : \; u(t) > -\infty\} \neq \emptyset$. 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 [6], 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* $u(x) = \min\limits_{k \in \{1,\ldots,K\}} \{a_k x + b_k\}$ *be a piecewise-linear concave utility function satisfying the following properties:*

1. *The number of linear pieces* $K \geq 2$*,*
2. *Each piece* $k \in \{1, \ldots, K\}$ *defines the utility function uniquely for at least one value of* x*,*
3. *Utility function is non-decreasing with* $u(0) = 0$*,* $1 \in \partial u(0)$*.*

For this class of piecewise-linear utility functions, the OCE is defined as

$$S_u(\tilde{x}) = \sup_{v \in \Re} \left\{ v + \mathrm{E}_{\mathbb{P}_x} \left(\min_{k \in \{1,\ldots,K\}} \{a_k(\tilde{x} - v) + b_k\} \right) \right\},$$

and the corresponding risk measure is defined as

$$\rho_u(\tilde{x}) = \inf_{v \in \Re} \left\{ v - \mathrm{E}_{\mathbb{P}_x} \left(\min_{k \in \{1,\ldots,K\}} \{a_k(\tilde{x} + v) + b_k\} \right) \right\}. \tag{3.1}$$

The corresponding 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 $\tilde{x} = c + \boldsymbol{y}'\tilde{\boldsymbol{z}} \in \mathcal{X}$. To evaluate this for a fixed vector $\boldsymbol{y}$, the complete knowledge of the multivariate distribution of $\tilde{\boldsymbol{z}}$ 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:

$$\hat{S}_u(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \sup_{v \in \Re} \left\{ v + \inf_{\mathbb{P} \in \mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\min_{k \in \{1,\ldots,K\}} \{a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}} - v) + b_k\} \right) \right\},$$

while the worst-case OCE risk measure is defined as:

$$\hat{\rho}_u(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \inf_{v \in \Re} \left\{ v - \inf_{\mathbb{P} \in \mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\min_{k \in \{1,\ldots,K\}} \{a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}} + v) + b_k\} \right) \right\}, \tag{3.2}$$

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

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

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 5. *For any random portfolio payoff $\tilde{x}$ 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

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

Proof. See Appendix A.

The worst-case OCE risk measure in Proposition 5(a) is simply the worst-case conditional value at risk at confidence level $1/a$ while 5(b) is a generalization to three pieces. We now show that the optimal portfolios obtained for these two cases can be significantly different.

Proposition 6. *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 $b > 0$:

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

Proof. See Appendix A.

The optimal portfolio allocations are plotted in Figure 2.

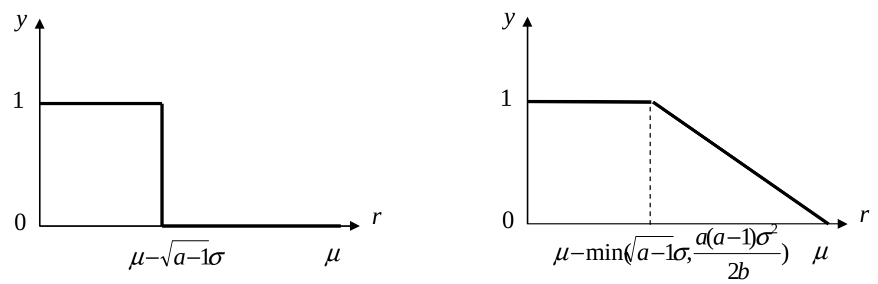


Figure 2: Optimal allocation in risky asset for two and three-piece utility functions.

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, with one risk-free and one risky asset the optimal worst-case portfolio in fact involves 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, with a three-piece utility function, we obtain non-trivial diversification for the optimal portfolio under the worst-case OCE risk measure. This is consistent with the notion of diversification and indicates the merit in studying the ambiguous versions of more complicated piecewise-linear utility functions.

While the OCE risk measure in (3.1) is a convex risk measure for utility functions satisfying Definition 3.2, the ambiguous risk measure in (3.2) could violate 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 $\tilde{\boldsymbol{z}}$ (possibly bounded or unbounded) and $\pi : \Re^n \times \Re^K \mapsto \Re$ be a jointly concave, positive homogeneous function satisfying*

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

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

$$\hat{\rho}_u(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \inf_{v \in \Re} \left\{ v - \pi\left(\boldsymbol{y}, \boldsymbol{a}(v + c) + \boldsymbol{b}\right) \right\}, \qquad (3.3)$$

is a convex risk measure over the space of random variables $\mathcal{X}$. If in addition, $\boldsymbol{b} = \mathbf{0}$, 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, $\mathcal{W} = \Re^n$. In that case, the condition in Theorem 3.1 simply reduces to $\pi(\boldsymbol{y}, \boldsymbol{d}) \geq -\infty$ for all $\boldsymbol{y}$, 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 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 [14]. 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

$$Y = \left\{ \boldsymbol{y} \in \Re^n : \boldsymbol{y} \geq \boldsymbol{0},\ \boldsymbol{y}'\boldsymbol{e} = 1,\ \boldsymbol{y}'\boldsymbol{\mu} = \mu_t \right\},$$

where $\boldsymbol{e}$ is a vector of ones and μ_t is a target return. It has been observed in numerous empirical tests (see Schwartz and Whitcomb [33], Simonds et al. [34], Conrad and Kaul [9]) that the distribution of stock returns exhibits nonstationarity. To avoid any potential bias due to the choice of the 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. For time-varying distributions, the resulting trading strategy would be myopically optimum. However this provides an easily implementable model that can validate the effect of capturing ambiguity in a single-period portfolio selection problem (see Garlappi et. al. [16] for a similar experiment).

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. An approximation of this function with a ten-piece linear function is given in Table 1. This 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 $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{N_{tr}}\}$, then the sample-based method for the OCE risk minimization portfolio problem solves

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

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 $\boldsymbol{Q}$ of returns of the risk assets. The robust approach then solves

$$\inf_{\boldsymbol{y}\in Y} \inf_{v\in\Re} \left(v - \hat{u}_1\left(v + \boldsymbol{y}'\tilde{\boldsymbol{z}}\right)\right),$$

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 robust approach then solves

$$\inf_{\boldsymbol{y}\in Y} \inf_{v\in\Re} \left(v - \hat{u}_3\left(v + \boldsymbol{y}'\tilde{\boldsymbol{z}}\right)\right),$$

where the bound on expected utility is computed from Theorem 2.4.

The computations were performed using the linear and second order cone programming solvers in ILOG CPLEX 10.1. In our experiments, we tested the quality of different approximations to the exponential utility function using 10, 100, 1000 and 10000 linear pieces. Table 2 displays the discretization (K), the approximation error (ϵ) and running time for the mean-covariance (MC) second order conic program. Even the large scale SOCPs run very quickly (in a matter of seconds) using standard solvers. For sake of brevity, we restrict our results to the case with $K = 10$ with an accuracy level of 6.799 $\times 10^{-5}$.

K	Approximation error (ϵ)	Running time (sec)
10	6.799 $\times 10^{-5}$	< 1
100	7.729 $\times 10^{-7}$	< 1
1000	7.830 $\times 10^{-9}$	4.9498
10000	7.840 $\times 10^{-11}$	14.3033

Table 2: Running times for different discretization.

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 $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{N_{ts}}, \boldsymbol{z}_{N_{ts}+1}, \ldots, \boldsymbol{z}_{20N_{ts}}\}$. For any 10-year dynamic strategy, $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{20})$, the realized mean and OCE risk is then given by

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

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. Computing the semiannual optimal trading strategy for each method took less than one second.

Table 3 provides the 10-year realized OCE risk and average daily returns of the three methods under each fixed target mean. Figure 3 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. Methods that use in-sample data to assume the complete distribution appear to result in large errors in the out-of-sample data. This is consistent with numerical results in Natarajan et. al. [24] for Value-at-risk (VaR) optimization. These strategies however perform better for in-sample data as one would expect (see Figure 4 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 5 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 interesting to note that cumulative wealth of the three portfolios appear to be moving simultaneously over the ten-year period. Based on all these observations, we can conclude that the robust approach based on the first two moments can provide good portfolio trading strategies.

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

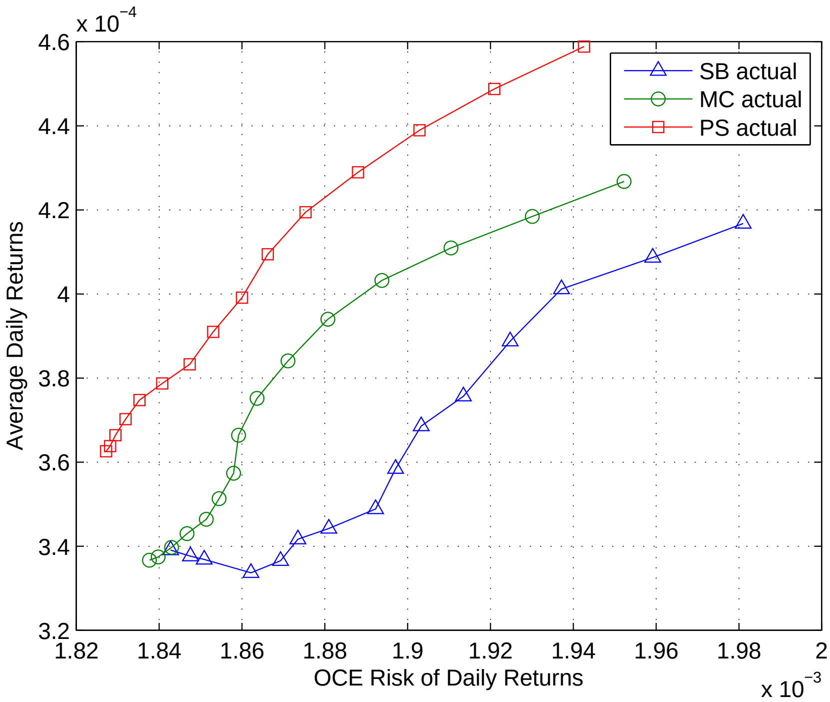
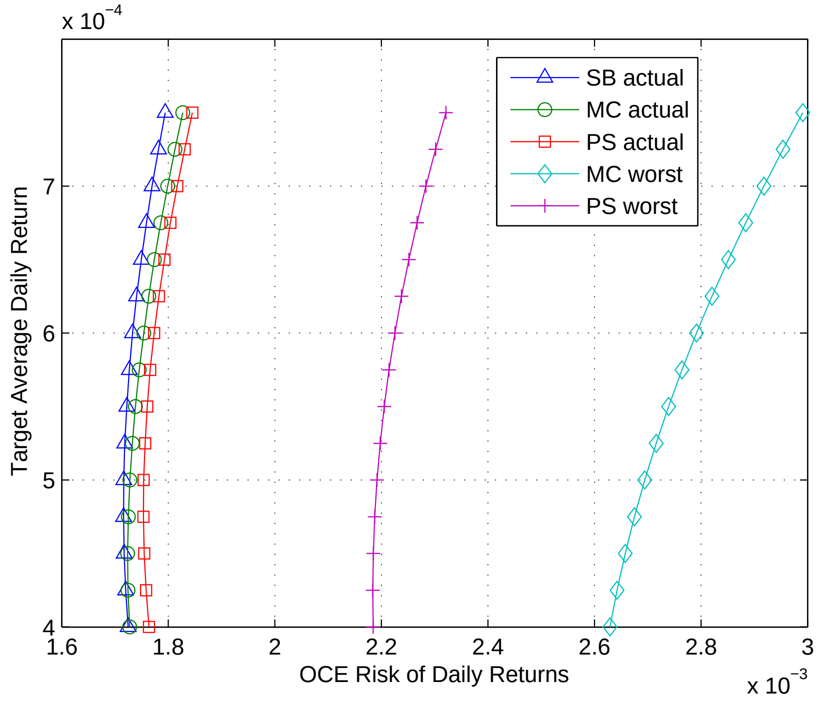
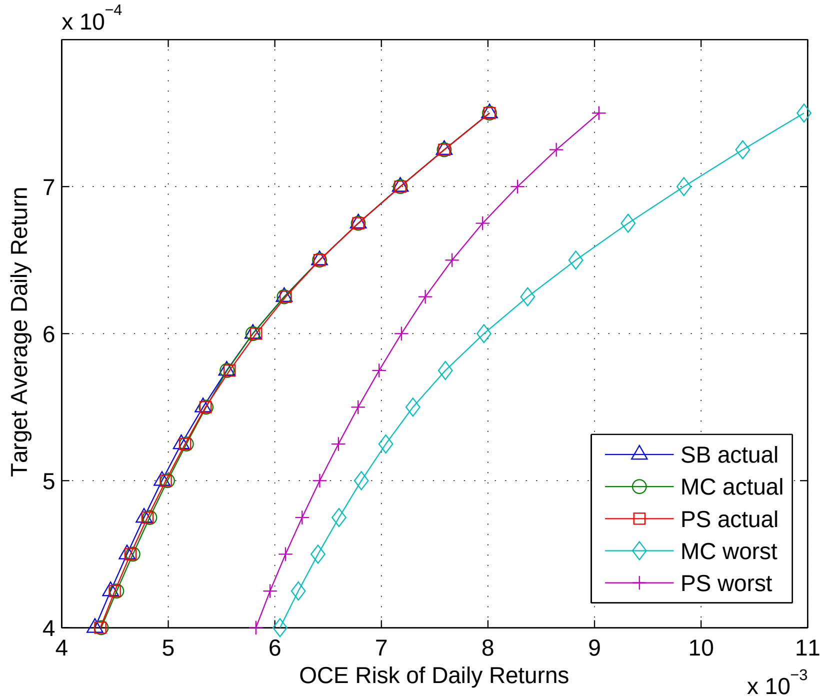


Figure 3: The efficient frontier of the rebalanced portfolio over September 1997 to August 2007.



(a) March 1998 to February 1999



(b) March 2002 to February 2003

Figure 4: 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.

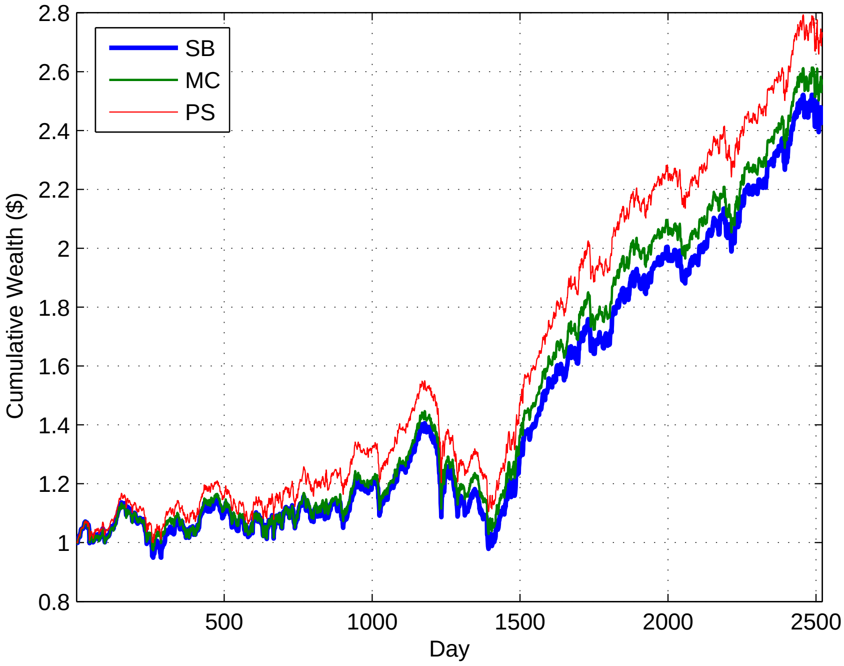


Figure 5: The cumulative wealth of the trading strategies over the period September 1997 to August 2007 for target mean return, $\mu_t = 0.06\%$.

A Proofs

Proof of Proposition 1: By definition, $Z_h \leq Z_{opt}$, since Z_{opt} is the optimal objective value for the utility function f. Let $(y_0^*, \boldsymbol{y}^*)$ be the true optimal solution under f. Then

$$\begin{aligned}
Z_{opt} &= \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(f(y_0^* r + \boldsymbol{y}^{*\prime}\tilde{\boldsymbol{z}})\right) \\
&\leq \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(u(y_0^* r + \boldsymbol{y}^{*\prime}\tilde{\boldsymbol{z}})\right) + \epsilon \\
&\leq \sup_{(y_0,\boldsymbol{y})\in Y} \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(u(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}})\right) + \epsilon \\
&= \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(u(\hat{y}_0 r + \hat{\boldsymbol{y}}'\tilde{\boldsymbol{z}})\right) + \epsilon \\
&\leq \inf_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(f(\hat{y}_0 r + \hat{\boldsymbol{y}}'\tilde{\boldsymbol{z}})\right) + \epsilon \\
&= Z_h + \epsilon
\end{aligned}$$

□

Proof of Theorem 2.1: Using Proposition 2, the moment problem over n random variables with given mean and covariance matrix

$$\inf_{\tilde{\boldsymbol{z}}\sim(\boldsymbol{\mu},\boldsymbol{Q})} \mathrm{E}_{\mathbb{P}}\left(\min_k\{a_k(c+\boldsymbol{y}'\tilde{\boldsymbol{z}})+b_k\}\right), \tag{A.1}$$

is equivalent to the moment problem over a single random variable with given mean and variance

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

Let $\mu_x = c + \boldsymbol{y}'\boldsymbol{\mu}$ and $\sigma_x^2 = \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}$. The dual formulation (see Isii [18]) is given as

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

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. We focus on the case with $\sigma_x > 0$. From Isii [18], strong duality holds under this regularity condition. The case when $\sigma_x = 0$ is easily handled as we see later. 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 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

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

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 an equivalent formulation for (A.3)

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

When $\sigma_x = 0$, the optimal solution sets both z and t equal to $\epsilon \downarrow 0$. As $\epsilon \downarrow 0$, the dual objective reduces to the primal objective of $\min_k (a_k\mu_x + b_k)$. Linearizing the objective, we obtain

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

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

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

From strong conic program duality for SOCP (A.5) (see Nesterov and Nemirovski [26]), the equivalent primal formulation is:

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

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:

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

□

Proof of Theorem 2.2: From strong duality results of Isii [18], the problem

$$\inf_{\tilde{\boldsymbol{z}} \sim_{\mathcal{W}} \boldsymbol{\mu}} \quad \mathrm{E}\left(\min_{k \in \{1,\ldots,K\}} \{a_k(c + \boldsymbol{y}'\tilde{\boldsymbol{z}}) + b_k\} \right),$$

is equivalent to the dual formulation

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

The optimal value for s_0 is then

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

The dual formulation thus reduces to

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

□

Proof of Theorem 2.3: We first show that that $\hat{u}_{\mathcal{L}}(c+\boldsymbol{y}'\tilde{\boldsymbol{z}})$ 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 $(\boldsymbol{y},\boldsymbol{d})$,

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

which is straightforward, since $\pi_{\mathcal{L}}(\boldsymbol{y},\boldsymbol{d})$ is the convolution of L lower bounds and by positive homogeneity, $\pi_l(\boldsymbol{0},\boldsymbol{0})=0$. 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 $(\boldsymbol{y},\boldsymbol{d})$

$$\begin{aligned}\sum_{l\in\mathcal{L}}\pi_l(\boldsymbol{y_l},\boldsymbol{d_l}) &\leq \sum_{l\in\mathcal{L}}\mathrm{E}_{\mathbb{P}}\left(\min_k\{a_k\boldsymbol{y_l}'\tilde{\boldsymbol{z}}+d_{l,k}\}\right),\\ &= \mathrm{E}_{\mathbb{P}}\left(\sum_{l\in\mathcal{L}}\min_k\{a_k\boldsymbol{y_l}'\tilde{\boldsymbol{z}}+d_{l,k}\}\right).\end{aligned}$$

However, since

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

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

$$\sum_{l\in\mathcal{L}}\pi_l(\boldsymbol{y_l},\boldsymbol{d_l}) \leq \pi_{\mathcal{L}}(\boldsymbol{y},\boldsymbol{d}) \leq \mathrm{E}_{\mathbb{P}}\left(\min_k\left\{a_k\boldsymbol{y}'\tilde{\boldsymbol{z}}+d_k\right\}\right).$$

This shows that $\pi_{\mathcal{L}}(\boldsymbol{y},\boldsymbol{d})$ remains a lower bound of (2.2) for all $\mathbb{P}\in\cap_l\mathbb{F}_l$.
To see that $\pi_{\mathcal{L}}(\boldsymbol{y},\boldsymbol{d})$ is positive homogenous, note that for all $\lambda>0$,

$$\begin{aligned}\pi_{\mathcal{L}}(\lambda\boldsymbol{y},\lambda\boldsymbol{d}) &= \sup_{\boldsymbol{y_l},\boldsymbol{d_l}} \sum_{l\in\mathcal{L}}\pi_l\left(\boldsymbol{y_l},\boldsymbol{d_l}\right)\\ &\qquad \text{s.t. } \sum_{l\in\mathcal{L}}\boldsymbol{y_l}=\lambda\boldsymbol{y},\\ &\qquad\qquad \sum_{l\in\mathcal{L}}\boldsymbol{d_l}=\lambda\boldsymbol{d},\\ &= \sup_{\boldsymbol{y_l},\boldsymbol{d_l}} \sum_{l\in\mathcal{L}}\pi_l\left(\lambda\boldsymbol{y_l},\lambda\boldsymbol{d_l}\right)\\ &\qquad \text{s.t. } \sum_{l\in\mathcal{L}}\boldsymbol{y_l}=\boldsymbol{y},\\ &\qquad\qquad \sum_{l\in\mathcal{L}}\boldsymbol{d_l}=\boldsymbol{d}.\end{aligned}$$

From positive homogeneity of π_l for all l, it follows that $\pi_{\mathcal{L}}(\lambda\boldsymbol{y},\lambda\boldsymbol{d})=\lambda\pi_{\mathcal{L}}(\boldsymbol{y},\boldsymbol{d})$ 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,

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

□

Proof of Theorem 2.4: We can show that $\hat{u}_3(c+\boldsymbol{y}'\tilde{\boldsymbol{z}})$ 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 (2.12) and for all $(\tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ satisfying the first two moments and support $\Re_+^{2n}$,

$$\mathrm{E}_{\mathbb{P}}\left(\min_{k\in\{1,\ldots,K\}}\left\{a_k(c+\boldsymbol{y}'\tilde{\boldsymbol{z}}^+-\boldsymbol{y}'\tilde{\boldsymbol{z}}^-)+b_k\right\}\right)\geq \bar{\pi}_1(\boldsymbol{y_1^p},\boldsymbol{y_1^m},\boldsymbol{d_1})+\bar{\pi}_2(\boldsymbol{y_2^p},\boldsymbol{y_2^m},\boldsymbol{d_2}).$$

Using the fact that $\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{z}}^+ - \tilde{\boldsymbol{z}}^-$, and by taking the supremum of the righthand side over all feasible solutions, we find that $\hat{u}_3(\boldsymbol{y}, \boldsymbol{a}c+\boldsymbol{b})$ is a lower bound of the worst-case expected utility under partitioned statistics. We can also find a feasible solution to problem (2.12) by setting $(\boldsymbol{y_1^p},\boldsymbol{y_1^m},\boldsymbol{d_1}) = (\boldsymbol{y}, -\boldsymbol{y}, \boldsymbol{a}c+\boldsymbol{b})$ and $(\boldsymbol{y_2^p},\boldsymbol{y_2^m},\boldsymbol{d_2}) = (\boldsymbol{0},\boldsymbol{0},\boldsymbol{0})$. From the equivalence (2.11) and $\boldsymbol{\mu} = \boldsymbol{\mu^p} - \boldsymbol{\mu^m}$, it follows that

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

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

□

Proof of Theorem 2.13: We can formulate $\hat{u}_{1s}$ as

$$\inf_{(\boldsymbol{\mu},\boldsymbol{Q})\in\mathcal{U}}\inf_{\mathbb{P}\in\mathbb{F}_1}\mathrm{E}_{\mathbb{P}}\left(\min_k\{a_k(c+\boldsymbol{y}'\tilde{\boldsymbol{z}})+b_k\}\right), \tag{A.7}$$

where $\mathcal{U}$ is the uncertainty set on the moments and $\mathbb{F}_1$ is the set of all distributions with known mean and covariance. Therefore, for any feasible $(\boldsymbol{\mu},\boldsymbol{Q})$ using Theorem 2.1, we can express the inner problem of (A.7) as an SOCP. Thus, $\hat{u}_{1s}$ can be reformulated as

$$\begin{aligned}\inf_{(\boldsymbol{\mu},\boldsymbol{Q})\in\mathcal{U}}\sup_{z,t,w,s}\quad & w-s\\ \text{s.t.}\quad & w\leq a_k(c+\boldsymbol{\mu}'\boldsymbol{y})+b_k-a_k^2z+a_kt,\quad \forall k=1,\ldots,K,\\ & s\geq\frac{\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}+t^2}{4z},\\ & z\geq 0.\end{aligned} \tag{A.8}$$

For the moment uncertainty set specified by upper and lower bounds on the mean and covariance, $\hat{u}_{1s}$ is thus equivalent to

$$\begin{aligned}\sup_{z,t,w,s}\quad & w-s\\ \text{s.t.}\quad & w\leq a_k(c+\boldsymbol{\mu}'\boldsymbol{y})+b_k-a_k^2z+a_kt,\quad \forall\underline{\boldsymbol{\mu}}\leq\boldsymbol{\mu}\leq\overline{\boldsymbol{\mu}},\quad \forall k=1,\ldots,K,\\ & s\geq\frac{\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y}+t^2}{4z},\quad \forall\underline{\boldsymbol{Q}}\leq\boldsymbol{Q}\leq\overline{\boldsymbol{Q}},\quad \boldsymbol{Q}\succeq\boldsymbol{0},\\ & z\geq 0.\end{aligned} \tag{A.9}$$

To see why, assume that $(\boldsymbol{\mu}^*,\boldsymbol{Q}^*,z^*,t^*,w^*,s^*)$ optimizes the problem (A.8). Further suppose that for some $(\boldsymbol{\mu}_0,\boldsymbol{Q}_0)\in\mathcal{U}$, at least one of the inequalities is violated for (z^*,t^*,w^*,s^*). In other words, at least one of the following inequalities is valid

$$\begin{aligned}
w^* &> a_k(c + \boldsymbol{\mu}_0'\boldsymbol{y}) + b_k - a_k^2 z^* + a_k t^*, \quad \forall k = 1, \ldots, K, \\
s^* &< \frac{\boldsymbol{y}'\boldsymbol{Q}_0\boldsymbol{y} + t^*}{4z^*}.
\end{aligned}$$

Then we can find $w < w^*$ or $s > s^*$ such that the inner problem of (A.8) is still feasible under $(\boldsymbol{\mu}_0, \boldsymbol{Q}_0)$. Furthermore this feasible solution under $(\boldsymbol{\mu}_0, \boldsymbol{Q}_0)$ attains a lower value than $w^* - s^*$. This violates the assumption that $(\boldsymbol{\mu}^*, \boldsymbol{Q}^*)$ achieves the optimal value to (A.8).
We can write the second inequality of (A.9) as

$$\begin{aligned}
s &\geq p + \frac{t^2}{4z}, \\
p &\geq \max_{\boldsymbol{Q}} \ \frac{1}{4z}\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y} \\
&\quad \text{s.t.} \ \ \underline{\boldsymbol{Q}} \leq \boldsymbol{Q} \leq \overline{\boldsymbol{Q}}, \\
&\qquad\quad \boldsymbol{Q} \succeq \mathbf{0}.
\end{aligned}$$

Consider the maximization problem above. The dual form can be written as:

$$\begin{aligned}
\min_{\underline{\boldsymbol{\Sigma}}, \overline{\boldsymbol{\Sigma}}} \quad & \overline{\boldsymbol{Q}} \bullet \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{Q}} \bullet \underline{\boldsymbol{\Sigma}} \\
\text{s.t.} \quad & \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{\Sigma}} - \frac{\boldsymbol{y}\boldsymbol{y}'}{4z} \succeq \mathbf{0}, \\
& \underline{\boldsymbol{\Sigma}}, \overline{\boldsymbol{\Sigma}} \geq 0.
\end{aligned}$$

Since by assumption that there exists a matrix in the uncertainty set that is positive definite, the maximization problem is strictly feasible and bounded. Hence from strong duality for semidefinite programs, the primal and dual formulations are equivalent. Using Schur's complement, the positive semidefinite constraint in the minimization problem can be written as

$$\begin{pmatrix} 4z & \boldsymbol{y}' \\ \boldsymbol{y} & \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{\Sigma}} \end{pmatrix} \succeq \mathbf{0}.$$

We can handle the first inequality constraint of (A.9) in a similar manner. Note that since $a_k \geq 0$ for all $k = 1, \ldots, K$, we can write it as

$$\begin{aligned}
w &\leq a_k(c + v) + b_k - a_k^2 z + a_k t, \quad \forall k = 1, \ldots, K \\
v &\leq \min_{\boldsymbol{\mu}} \ \boldsymbol{y}'\boldsymbol{\mu} \\
&\quad \text{s.t.} \ \ \underline{\boldsymbol{\mu}} \leq \boldsymbol{\mu} \leq \overline{\boldsymbol{\mu}}.
\end{aligned}$$

The maximization problem has a dual form of

$$\begin{aligned}
\max_{\underline{\boldsymbol{\omega}}, \overline{\boldsymbol{\omega}}} \quad & \underline{\boldsymbol{\mu}}'\underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\mu}}'\overline{\boldsymbol{\omega}} \\
\text{s.t.} \quad & \underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\omega}} = \boldsymbol{y}, \\
& \underline{\boldsymbol{\omega}}, \overline{\boldsymbol{\omega}} \geq \mathbf{0}.
\end{aligned}$$

Therefore, combining everything, we find that $\hat{u}_{1s}$ has the following equivalent form:

$$
\begin{array}{lll}
\sup\limits_{z,t,w,s,p,v,\underline{\boldsymbol{\omega}},\overline{\boldsymbol{\omega}},\underline{\boldsymbol{\Sigma}},\overline{\boldsymbol{\Sigma}}} & w - s & \\
\text{s.t.} & w \leq a_k(c+v) + b_k - a_k^2 z + a_k t, \quad \forall k = 1,\dots,K \\
& v \leq \underline{\boldsymbol{\mu}}'\underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\mu}}'\overline{\boldsymbol{\omega}} \\
& 4z(s-p) \geq t^2, \\
& p \geq \overline{\boldsymbol{Q}} \bullet \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{Q}} \bullet \underline{\boldsymbol{\Sigma}} \\
& \underline{\boldsymbol{\omega}} - \overline{\boldsymbol{\omega}} = \boldsymbol{y}, \\
& \begin{pmatrix} 4z & \boldsymbol{y}' \\ \boldsymbol{y} & \overline{\boldsymbol{\Sigma}} - \underline{\boldsymbol{\Sigma}} \end{pmatrix} \succeq \boldsymbol{0}, \\
& z \geq 0 \ \underline{\boldsymbol{\omega}}, \overline{\boldsymbol{\omega}} \geq \boldsymbol{0}, \ \underline{\boldsymbol{\Sigma}}, \overline{\boldsymbol{\Sigma}} \geq \boldsymbol{0}.
\end{array}
$$

□

Proof of Proposition 5: Consider the random variable $\tilde{x}$ with mean μ_x and variance $\sigma_x^2 > 0$.
(a) From Proposition 4, the worst-case OCE for the two-piece utility function, is obtained by solving the single variable maximization problem

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

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 4, the worst-case OCE for the three-piece utility function is obtained by solving the single variable maximization problem

$$\hat{S}_u(\tilde{x}) = \sup_{v \in \Re} \left\{ v + \hat{u}(\tilde{x} - v) \right\},$$

where

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

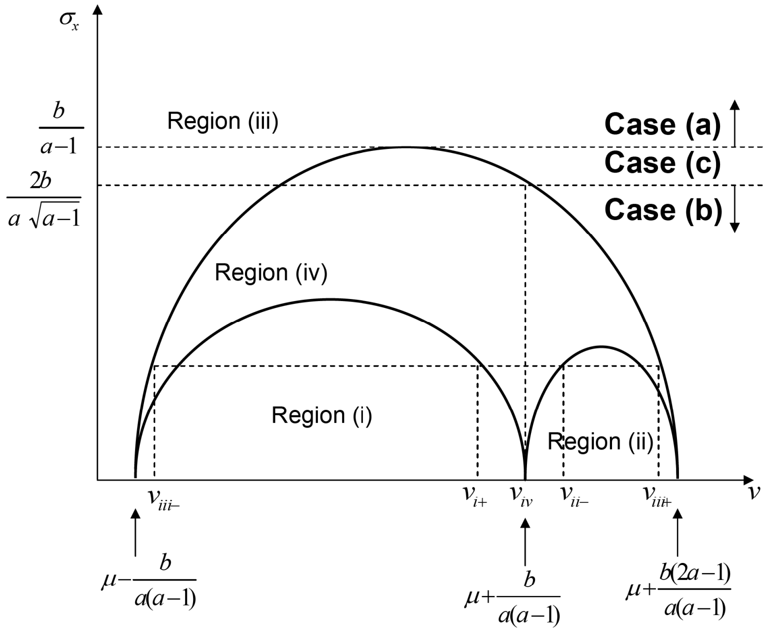


Figure 6: Characterization of the regions for three-piece utility function.

We partition the four different regions (i)-(iv) in the (v, σ_x) space as indicated in Figure 6. In fact, we show next that only regions (iii) or (iv) can occur in the optimal solution.

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

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

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): $\boldsymbol{\sigma_x \leq 2b/(a\sqrt{a-1})}$:
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 6).

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

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

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 6)

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

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 6). Lastly over the region (iv), the objective function for the worst-case OCE is given as

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

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

In this case, v_{iii} definitely lies in region (iii) and is feasible while v_{iv} always lies in region (iii) (see Figure 6(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.

□

Proof of Proposition 6: The portfolio allocation problem is given as

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

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

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

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

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

The optimal solution to this problem is given as:

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

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:

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

Combining these two solutions, we get the optimal portfolio:

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

□

Proof of Theorem 3.1: To prove that $\hat{\rho}_u : \mathcal{X} \mapsto \Re$ 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 $c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}}$, $c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}} \in \mathcal{X}$ such that

$$c_1 + \boldsymbol{y_1}'\boldsymbol{z} \geq c_2 + \boldsymbol{y_2}'\boldsymbol{z}, \qquad \forall \boldsymbol{z} \in \mathcal{W}.$$

Equivalently, we have

$$\inf_{\boldsymbol{z} \in \mathcal{W}} (c_1 - c_2 + (\boldsymbol{y_1} - \boldsymbol{y_2})'\boldsymbol{z}) \geq 0.$$

Then for all $v \in \Re$, we have

$$\begin{aligned} \pi(\boldsymbol{y_1}, \boldsymbol{a}(c_1 + v) + \boldsymbol{b}) &= \pi(\boldsymbol{y_2} + (\boldsymbol{y_1} - \boldsymbol{y_2}), \boldsymbol{a}(c_2 + v) + \boldsymbol{b} + \boldsymbol{a}(c_1 - c_2)), \\ &\geq \pi(\boldsymbol{y_2}, \boldsymbol{a}(c_2 + v) + \boldsymbol{b}) + \pi(\boldsymbol{y_1} - \boldsymbol{y_2}, \boldsymbol{a}(c_1 - c_2)), \end{aligned}$$

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

$$\begin{aligned} \pi(\boldsymbol{y_1} - \boldsymbol{y_2}, \boldsymbol{a}(c_1 - c_2)) &\geq \min_{k \in \{1,\ldots,K\}} \left\{ \inf_{\boldsymbol{z} \in \mathcal{W}} a_k (c_1 - c_2 + (\boldsymbol{y_1} - \boldsymbol{y_2})'\boldsymbol{z}) \right\} \\ &\geq 0 \end{aligned}$$

Hence

$$\inf_v \{v - \pi(\boldsymbol{y_2}, \boldsymbol{a}(c_2 + v) + \boldsymbol{b})\} \geq \inf_v \{v - \pi(\boldsymbol{y_1}, \boldsymbol{a}(c_1 + v) + \boldsymbol{b})\},$$

or equivalently, $\hat{\rho}_u(c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}}) \geq \hat{\rho}_u(c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}})$. 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 $d \in \Re$,

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

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

$$\hat{\rho}_u\left(\lambda(c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}}) + (1-\lambda)(c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}})\right) = \inf_v \left\{v - \pi\left(\lambda \boldsymbol{y_1} + (1-\lambda)\boldsymbol{y_2}, \boldsymbol{a}(\lambda c_1 + (1-\lambda)c_2 + v) + \boldsymbol{b}\right)\right\}.$$

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

$$\begin{aligned} & \pi\left(\lambda \boldsymbol{y_1} + (1-\lambda)\boldsymbol{y_2}, \boldsymbol{a}(\lambda c_1 + (1-\lambda)c_2 + v) + \boldsymbol{b}\right) \\ = \; & \pi\left(\lambda \boldsymbol{y_1} + (1-\lambda)\boldsymbol{y_2}, \lambda\left(\boldsymbol{a}(c_1 + v_1) + \boldsymbol{b}\right) + (1-\lambda)\left(\boldsymbol{a}(c_2 + v_2) + \boldsymbol{b}\right)\right), \\ \geq \; & \lambda\pi\left(\boldsymbol{y_1}, \boldsymbol{a}(c_1 + v_1) + \boldsymbol{b}\right) + (1-\lambda)\pi\left(\boldsymbol{y_2}, \boldsymbol{a}(c_2 + v_2) + \boldsymbol{b}\right). \end{aligned}$$

Thus,

$$\begin{aligned} & \hat{\rho}_u\left(\lambda(c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}}) + (1-\lambda)(c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}})\right) \\ \leq \; & \inf_{v_1, v_2} \left\{\lambda\left(v_1 - \pi(\boldsymbol{y_1}, \boldsymbol{a}(c_1 + v_1) + \boldsymbol{b})\right) + (1-\lambda)\left(v_2 - \pi(\boldsymbol{y_2}, \boldsymbol{a}(c_2 + v_2) + \boldsymbol{b})\right)\right\}, \\ = \; & \lambda \inf_{v_1} \left\{v_1 - \pi(\boldsymbol{y_1}, \boldsymbol{a}(c_1 + v_1) + \boldsymbol{b})\right\} + (1-\lambda)\inf_{v_2}\left\{v_2 - \pi(\boldsymbol{y_2}, \boldsymbol{a}(c_2 + v_2) + \boldsymbol{b})\right\}, \\ = \; & \lambda\hat{\rho}_u(c_1 + \boldsymbol{y_1}'\tilde{\boldsymbol{z}}) + (1-\lambda)\hat{\rho}_u(c_2 + \boldsymbol{y_2}'\tilde{\boldsymbol{z}}). \end{aligned}$$

Therefore, this defines a convex risk measure.
Finally, suppose $\boldsymbol{b} = \boldsymbol{0}$. For any $\lambda > 0$,

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

Since π is positive homogeneous, then

$$\begin{aligned} \hat{\rho}_u(c + \lambda \boldsymbol{y}'\tilde{\boldsymbol{z}}) &= \lambda \inf_v \left\{v - \pi(\boldsymbol{y}, \boldsymbol{a}(c + v))\right\}, \\ &= \lambda\hat{\rho}_u(y_0 r + \boldsymbol{y}'\tilde{\boldsymbol{z}}). \end{aligned}$$

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

References

[1] Arrow, K. J., and G. Debreu (1954): Existence of a Competitive Equilibrium for a Competitive Economy, *Econometrica* 22(3), 265-290.

[2] Acerbi, C., and D. Tasche (2002): On the Coherence of Expected Shortfall, *J. Bank. Financ.* 26, 1487-1503.

[3] Artzner, P., F. Delbaen, J. M. Eber, and D. Heath (1999): Coherent Measures of Risk, *Math. Financ.* 9(3), 203-228.

[4] Ben-Tal, A., and A. Nemirovski (1998): Robust Convex Optimization, *Math. Oper. Res.* 23(4), 769-805.

[5] Ben-Tal, A., and M. Teboulle (1986): Expected Utility, Penalty Functions and Duality in Stochastic Nonlinear Programming, *Management Sci.* 32(11), 1445-1466.

[6] Ben-Tal, A., and M. Teboulle (2007): An Old-New Concept of Convex Risk Measures: The Optimized Certainty Equivalent, *Math. Financ.* 17(3), 449-476.

[7] Buhlmann, H. (1970): *Mathematical Methods in Risk Theory*, Springer, Berlin.

[8] Calafiore, G. C. (2007): Ambiguous Risk measures and Optimal Robust Portfolios, *SIAM J. Optim.* 18(3), 853-877.

[9] Conrad, J., and G. Kaul (1988): Time-Variation in Expected Returns, *J. Business* 61(4), 409-425.

[10] Delage, E., and Y.Ye (2008): Distributionally Robust Optimization under Moment Uncertainty with Application to Data-Driven Problems, Working Paper.

[11] Detemple, J. B., R. Garcia, and M. Rindisbacher (2003): A Monte Carlo Method for Optimal Portfolios, *J. Finance* 58(1), 401-446.

[12] Erdoğan, E., and G. Iyengar (2006): Ambiguous Chance Constrained Problems and Robust Optimization, *Math. Prog.* 107, 37-61.

[13] Föllmer, H., and A. Schied (2002): Convex Measures of Risk and Trading Constraints, *Financ. Stochast.* 6, 429-447.

[14] French, K., and E. Fama: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[15] Fritelli, M., and E. R. Gianin (2002): Putting Order in Risk Measures, *J. Bank. Financ.* 26, 1473-1486.

[16] Garlappi, L., R. Uppal, and T. Wang (2007): Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach, *Rev. Financ. Stud.* 20(1), 41-81.

[17] Gilboa, I., and D. Schmeidler (1989): Maximin Expected Utility Theory with Non-Unique Prior, *J. Math. Econ.* 18, 141-153.

[18] Isii, K. (1963): On the Sharpness of Chebyshev-Type Inequalities, *Ann. Inst. Stat. Math.* 12, 185-197.

[19] Korn, R., and O. Menkens (2005): Worst-Case Scenario Portfolio Optimization: A New Stochastic Control Approach, *Math. Meth. Oper. Res.* 62, 123-140.

[20] Markowitz, H. (1952): Portfolio Selection, *J. Finance* 8, 77-91.

[21] Merton, R. C. (1971): Optimum Consumption and Portfolio Rules in a Conitinuous-Time Model, *J. Economic Theory* 3, 373-413.

[22] Murty, K. G., and S. N. Kabadi (1987): Some NP-Complete Problems in Quadratic and Nonlinear Programming, *Math. Prog.* 39, 117-129.

[23] Natarajan, K., and Z. Linyi (2007): A Mean-Variance Bound for a Three Piece Linear Function, *Prob. Eng. Inf. Sci.* 21, 611-621.

[24] Natarajan, K., D. Pachamanova, and M. Sim (2008): Incorporating Asymmetric Distributional Information in Robust Value-At-Risk Optimization, *Management Sci.* 54(3), 573-585.

[25] Natarajan, K., M. Song, and C. P. Teo (2008): Persistency Model and its Applications in Choice Modeling, To appear in *Management Sci.*

[26] Nesterov, Y., and A. Nemirovski (1994): Interior-Point Polynomial Methods in Convex Programming, *SIAM Stud. Appl. Math.* 13, Philadelphia, PA.

[27] Ocone, D., and I. Karatzas (1991): A Generalized Clark Representation Formula, with Applications to Optimal Portfolios, *Stochastics* 37(3), 187-220.

[28] Pflug, G. (2000): Some Remarks on the Value-At-Risk and the Conditional Value-At-Risk, *Probabilistic Constrained Optimization: Methodology and Applications,* S. Uryasev, ed., Kluwer Academic Publishers, Dordrecht.

[29] Pflug, G., and A. Ruszczynski (2004): A Risk Measure for Income Processes, *Risk Measures for the 21st Century,* G. Szegoe, ed., J. Wiley and Sons, New York.

[30] Popescu, I. (2007): Robust Mean-Covariance Solutions for Stochastic Optimization, *Oper. Res.* 55(1), 98-112.

[31] Rockafellar, R. T., and S. Uryasev (2000): Optimization of Conditional Value-At-Risk, *J. Risk* 2, 21-41.

[32] Rogosinsky, W. W. (1958): Moments of Non-Negative Mass, *Proc. Roy. Soc. London. Ser. A* 245, 1-27.

[33] Schwartz, R. A., and D. K. Whitcomb (1977): Evidence on the Presence and Causes of Serial Correlation in Market Model Residuals, *J. Financ. Quant. Anal.* 12(2): 291-313.

[34] Simonds, R. R., L. R. LaMotte, and Jr. A. McWhorter (1986): Testing for Nonstationarity of Market: An Exact Test and Power Considerations, *J. Financ. Quant. Anal.* 21(2), 209-220.

[35] Tütüncü, R. H., and M. Koenig (2004): Robust Asset Allocation, *Ann. Oper. Res.* 132, 157-187.

[36] Von-Neumann, J., and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.

[37] Wippern, R. F. (1971): Utility Implications of Portfolio Selection and Performance Appraisal Models, *J. Financ. Quant. Anal.* 6(3), 913-924.