Robust Decision Making using a Risk-Averse Utility Set*

Jian Hu Sanjay Mehrotra
Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208
jianhu@northwestern.edu, mehrotra@iems.northwestern.edu

March 17, 2012

Abstract

Eliciting the utility of a decision maker is difficult. In this paper, we develop a flexible decision making framework, which uses the concept of utility robustness to address the problem of ambiguity and inconsistency in utility assessments. The ideas are developed by giving a probabilistic interpretation to utility and marginal utility functions. Boundary and additional conditions are used to describe a utility set that characterizes a decision maker's risk attitude. Reformulation and convergence results are given for the discrete and continuous specifications of the utility set. A portfolio investment decision problem is used to illustrate the basic ideas, and demonstrate the usefulness of the proposed decision making framework.

Key Words: Utility Function, Marginal Utility Function, Expected Utility Decision Making, Robust Optimization, Portfolio Optimization

*The research is partially supported by the National Science Foundation grant CMMI-1131386 and Office of Naval Research grant ONR-N00014210051

1 Introduction

Utility is a fundamental concept in decision making. It is used in a wide variety of disciplines, such as economics, energy, finance, health care, management, marketing, etc. Complete and comprehensive investigations in the literature have systemized utilitarian theories, axioms, and postulates, which explain economic or psychologic behavior in terms of consumption of various goods and services, possession of wealth, and spending of leisure time (see e.g., Stigler (1950a,b, 1972); Keeney and Raiffa (1976); and reference therein). However, a normative and rational decision maker finds it difficult to exactly describe his utility function. In this paper, we study a framework of decision making under utility uncertainty. More specifically, we study the problem

$$\max_{x\in\mathcal{X}} \pi(x), \tag{1.1}$$

where $\pi(x)$ is described by an inner minimization problem

$$\pi(x) := \min_{u\in\mathcal{U}} \Big\{ \Pi(x,u) := \mathbb{E}[u(\xi(x))] \Big\}. \tag{1.2}$$

Here, $\mathcal{X} \in \mathbb{R}^n$ is the decision feasible region, $\xi(x)$ represents a measurable random function for each $x \in \mathcal{X}$, $\mathcal{U}$ is a set of risk-averse utility functions discussed further in Section 1.1.

According to the expected utility theory in von Neumann and Morgenstern (1947), there exists a utility function characterizing the relative attitude of a decision maker toward risk arising from the uncertainty. However, the correspondence of the utility function and the attitude is rather obscure, which Karmarkar (1978) and Chajewska et al. (2000) ascribed to cognitive difficulty. Although standard gamble methods and paired gamble methods such as preference comparison, probability equivalence, value equivalence, and certainty equivalence have been proposed to elicit a decision maker's utility (see e.g. Farquhar (1984), Wakker and Deneffe (1996) and references therein), the questionnaire in these methods are difficult for untrained clients to answer. Different methods yield inconsistent utility constructions (Hershey and Schoemaker (1985), Froberg and Kane (1989), and Nord (1992)). Moreover, a limited number of survey questions may be insufficient to completely quantify a real-life decision problem which often has a very large decision space (Chajewska et al. (2000)).

In addition to the non-parametric utility assessment approaches, parametric estimation is also common. For instance Pratt (1964) and Arrow (1965) independently derived measures for absolute and relative risk aversion, and established widely accepted constant absolute risk aversion (CARA) or relative risk aversion (CRRA) utility functions, i.e., the negative exponential utility function with constant coefficient of absolute risk aversion γ_a

$$u_a(t) = 1 - e^{-\gamma_a t}, \tag{1.3}$$

and the power utility function with constant coefficient of relative risk aversion γ_r

$$u_r(t) = \frac{t^{1-\gamma_r}}{1-\gamma_r}. \tag{1.4}$$

These coefficients of risk aversion have been studied experimentally (e.g. Szpiro (1986), Riley Jr. and Chow (1992), Schooley and Worden (1996), Holt and Laury (2002), and Brunello (2002)). For example, Holt and Laury (2002) indicated that a risk averse decision maker exhibits (1.4) with $0.41 \le \gamma_r \le 0.68$.

The concept of stochastic dominance provides an alternative approach. Here, a class of utility functions preserves a decision maker's primary preference (see e.g., Müller and Stoyan (2002) and references therein). This allows us to model risk attitudes, even though an exact form of the utility function is not known; for example, we can model risk-aversion by second order stochastic dominance saying that a random output ξ_1 is preferred to another random output ξ_2 if $\mathbb{E}[u(\xi_1)] \geq \mathbb{E}[u(\xi_2)]$ for all increasing concave utility functions u. Different notions of stochastic dominance can be defined corresponding to different classes of functions u. Unfortunately, however, stochastic dominance is too conservative because it permits a large class of utilities.

Problem (1.1) inherits the feature of stochastic dominance in that a class of utility functions is used to address ambiguity and inconsistency of utility elicitation. However, this class of utilities is restricted to a set by specifying additional requirements. Consequently, the utility set $\mathcal{U}$ specified in the following section provides a flexible framework which allows a decision maker to specify risk attitudes by providing more information beyond a general classification of attitude by stochastic dominance.

1.1 A Set of Risk Averse Utilities

Monotonicity and boundedness are usually assumed to be basic properties of utility functions (Farquhar (1984)). We assume throughout that, for all $x \in \mathcal{X}$, a random outcome $\xi(x)$ has a bounded support in $\Theta := [0, \theta]$. Let $\mathfrak{U}_1$ be the set of all increasing utility functions well defined on Θ, which are right-continuous with left limits (RCLL) and satisfy the boundary condition:

$$u(0) = 0, \quad u(\theta) = 1. \tag{1.5}$$

We use $\mathfrak{U}_2$ to denote the subset of $\mathfrak{U}_1$ which consists of all concave utility functions in $\mathfrak{U}_1$. Note that condition (1.5) is a commonly used normalization of utility. By the following proposition, the strategic equivalence that the preference ranking of any two alternatives is not changed in the transformation is preserved while making this normalization.

Proposition 1.1 (Theorem 4.1 in Keeney and Raiffa (1976)) *Two utility functions u_1 and u_2 with the same domain are strategically equivalent if and only if there exits two constants h and $k > 0$ such that, for all x in the domain, $u_1(x) = h + ku_2(x)$.*

We use functions $\underline{a}$, $\overline{a}$, $\underline{b}$, and $\overline{b}$ to depict bounds on the marginal utility u' as

$$\underline{a} \leq u' \leq \overline{a} \tag{1.6}$$

and bounds on the utility u as

$$\underline{b} \leq u \leq \overline{b}. \tag{1.7}$$

Additional conditions on utility are constructed by

$$\int \varphi^i du \leq c_i, \quad i = 1, \dots, m, \tag{1.8}$$

using Lebesgue integrable functions φ^i, $i = 1, \dots, m$, i.e., for the Lebesgue measure ℓ,

$$\int |\varphi^i| d\ell < \infty, \quad i = 1, \dots, m. \tag{1.9}$$

The boundary and general conditions (1.6)-(1.8) provide a flexible characterization of a decision maker's risk attitudes, addressing classical non-parametric and parametric utility assessments in a natural way. This will be shown in Section 2 with the help of serval examples.

We now specify the utility set $\mathcal{U}$ in problem (1.2) by combining conditions (1.6), (1.7), and (1.8):

$$\mathcal{U} := \{u \in \mathfrak{U}_2 \mid u \text{ satisfies conditions (1.6), (1.7), and (1.8)}\}. \tag{1.10}$$

1.2 Contributions and Organization of this Paper

In Section 2.1 we consider utility function from a novel probabilistic perspective. In particular, we interpret the utility function as a c.d.f. of an underlying random utility index. A fundamental complement relationship between random outcome and utility function is deduced from this interpretation. On this basis, we study the boundary and general conditions on utility describing the class $\mathcal{U}$ in (1.10) in Sections 2.2 and 2.3. An investment decision problem is used to illustrate the constructions of conditions by analyzing an investor's risk preference in these sections. In Sections 3 and 4 we study the decision making framework (1.1) using utility set motivated in Section 2. Section 3 provides a linear programming tractable formulation of a particular case where the bounds of utility, $\underline{b}$ and $\overline{b}$, are piecewise linear increasing concave function, and the bounds on marginal utility, $\underline{a}$ and $\overline{a}$, as well as all φ^i in the general conditions are step functions. For the general continuous case addressed in Section 4, we study an approximation problem, which is proved to asymptotically converge to the true counterpart. Based on the discussion in the previous sections, Section 5 further studies the robust investment decision making model and illustrates the usefulness of the framework (1.1) with the help of a numerical example. While this paper presents results for the risk averse case, a comparison paper (Hu and Mehortra (2012)) studies the case of general utility set.

2 Construction of Utility Set

In this section a probabilistic interpretation of utility is given, and on this basis, we study boundary and general conditions used to describe the utility set $\mathcal{U}$. Several examples are used to illustrate the use of classical non-parameter and parameter utility assessments in our framework.

2.1 Probabilistic Interpretation of Utility

In economics, utility is a measure of the relative happiness and satisfaction gained from a good or service. Marginal utility is the derivative of utility, which quantifies positive (or negative) impact on consumers resulting from an increase (or decrease) in the consumption of that good or service. As mentioned earlier, these traditional definitions are based on economic and psychologic analyses of consumers' normative behavior. We now consider the concept of utility and marginal utility from a probabilistic perspective.

We identify a utility function $u \in \mathfrak{U}_1$ as the cumulative probability function (c.d.f.) of a random variable ζ independent of random output ξ. The random variable ζ is called the random utility index of a decision maker. Denote the set Υ_1 of all ζ's obtained from this one-to-one mapping on $\mathfrak{U}_1$. The expected utility of random output ξ,

$$\mathbb{E}[u(\xi)] = \mathbb{E}[\Pr\{\xi \geq \zeta | \xi\}] = \mathbb{E}[\mathbb{E}[\mathbf{1}\{\xi \geq \zeta\}|\xi]] = \mathbb{E}[\mathbf{1}\{\xi \geq \zeta\}] = \Pr\{\xi \geq \zeta\}, \tag{2.1}$$

can be viewed to quantify the probability of ξ over ζ, i.e, against the random utility index ζ, a decision maker pursues a maximum benefit. A random outcome $\hat{\xi}_1$ is preferred to another random outcome $\hat{\xi}_2$ if

$$\Pr\left\{\hat{\xi}_1 \geq \zeta\right\} \geq \Pr\left\{\hat{\xi}_2 \geq \zeta\right\}. \tag{2.2}$$

In order to see the practical importance of equality (2.1), consider an example where two different supply policies meet a random market demand. Let $\hat{\xi}_1$ and $\hat{\xi}_2$ be the random outcomes produced by these supply polices and ζ represent the random demand. The relationship (2.2) suggests that a better policy should satisfy the market demand with a higher probability. This insight takes us beyond the traditional approach that constructs a utility function u characterizing consumers' happiness and satisfaction and thus selects the policy with the maximum expected utility. In this supply-demand case, the utility function constructed by the market demand ζ is "objective", depending on market data and analysis which excludes customers' "subjective" opinions and feelings.

Based on (2.1), the following proposition gives another interesting and useful complementary interpretation of the random outcome ξ and the random utility index ζ.

Proposition 2.1 *Let ζ be the random utility index with respect to a continuous utility function $u \in \mathfrak{U}_1$. Then a complementary relationship between ξ and ζ is given by*

$$\mathbb{E}[u(\xi)] + \mathbb{E}[F_\xi(\zeta)] = \mathbb{E}[F_\zeta(\xi)] + \mathbb{E}[F_\xi(\zeta)] = 1, \tag{2.3}$$

where F_ξ is the c.d.f. of ξ and $F_\zeta = u$ is the c.d.f. of ζ.

Proof: It follows from Theorem 4.30 in Rudin (1976) that the monotonic function F_ξ at most has countable points of discontinuity, denoted as t_k, $k = 1, \ldots, K$. Note that K can be ∞. Then $\Pr\{\xi = t_k\} =: \mu_k > 0$ and $\Pr\{\xi = t\} = 0$ if $t \neq t_k$ for all k. Since ζ is independent of ξ, we have

$$\Pr\{\xi = \zeta\} = \mathbb{E}[\Pr\{\xi = \zeta|\zeta\}] = \int_\Theta \Pr\{\xi = t\}\, du(t) = \sum_{k=1}^K \int_\Theta \mu_k \mathbf{1}\{\xi = t_k\} du(t) = 0, \tag{2.4}$$

where the last equality holds because u is continuous. It then follows from (2.1) and (2.4) that

$$\mathbb{E}[u(\xi)] = \Pr\{\xi \geq \zeta\} = 1 - \Pr\{\zeta > \xi\} = 1 - \Pr\{\zeta \geq \xi\} = 1 - \mathbb{E}[F_\xi(\zeta)].$$

□

The complementary relationship (2.3) indicates that the random outcome ξ and the random utility index ζ are a negative symmetrical "pair". In the above example of random supply and demand, the relationship (2.3) provides a symmetry in handling risk due to increased randomness in supply or demand.

In the probabilistic interpretation, a marginal utility for random output ξ is the probability density (mass) function of $\zeta \in \Upsilon_1$. A random utility index $\zeta \in \Upsilon_1$ corresponds to a risk attitude. The risk-neutral preference is described by a uniformly distributed ζ. A constant p.d.f. indicates a decision maker's indifference to risk. The p.d.f. of a risk-averse utility index is a decreasing function, which gives more evaluation at the side of low gain so as to avoid winning a high gain at a large risk. An example of such as a utility index is a random variable with a left triangle distribution on Θ given as

$$f_\zeta(t) = \frac{2(\theta - t)}{\theta^2}, \quad u(t) = F_\zeta(t) = \frac{2t}{\theta} - \frac{t^2}{\theta^2}, \quad t \in \Theta,$$

where $f_\zeta(\cdot)$ is the p.d.f. of ζ. On the contrary, the risk-seeking decision maker is interested in sacrificing investment security for high profit. This attitude is depicted by an increasing p.d.f. emphasizing a high gain. For instance, for a risk seeker, ζ may be a random variable with a right triangle distribution on Θ given as

$$f_\zeta(t) = \frac{2t}{\theta^2}, \quad u(t) = F_\zeta(t) = \frac{t^2}{\theta^2}, \quad t \in \Theta.$$

2.2 Motivation and Interpretation of the Boundary Conditions

Let us study the boundary conditions (1.6) on marginal utility and (1.7) on utility. $u \in \mathfrak{U}_2$ is a continuous increasing concave function on Θ, and is thus differentiable on Θ a.e.. Moreover, the left and right derivatives of u exists for every point in Θ. Without loss of generality, we define the derivative of a continuous concave function f on Θ in this paper as follows:

$$f'(t) = \begin{cases} f'(t^+) & t \in [0, \theta), \\ f'(\theta^-) & t = \theta. \end{cases} \tag{2.5}$$

The derivative f' in this definition satisfies RCLL on Θ except the right boundary point θ, at which f' is left continuous.

In the boundary condition (1.6) on marginal utility, $\underline{a}$ and $\overline{a}$ are assumed to be two positive extended real-valued decreasing functions on Θ. In the context of probability, condition (1.6) describes a preference relationship of probability measure that a finite signed measure μ_1 is said to be preferred to another finite signed measure μ_2 (written as $\mu_1 \succeq \mu_2$) if $\mu_1(\mathcal{A}) \geq \mu_2(\mathcal{A})$ for any Borel measured set $\mathcal{A} \subseteq \Theta$. If $\underline{a}$ and $\overline{a}$ are p.d.f's, the integration of (1.6) tells us that

$$\int_0^t \underline{a}(s)ds \leq u(t) \leq \int_0^t \overline{a}(s)ds, \quad t \in \Theta,$$

which indicates that the utility function u stochastically dominates $\int_0^t \underline{a}(s)ds$ and is dominated by $\int_0^t \overline{a}(s)ds$ in first order. Obviously, the boundary condition (1.6) gives a stricter requirement than first order stochastic dominance.

We assume throughout this paper that $\underline{b}$ and $\overline{b}$ in the boundary condition (1.7) are increasing concave functions on Θ satisfying $\underline{b}(0) \leq 0 \leq \overline{b}(0)$ and $\underline{b}(\theta) \leq 1 \leq \overline{b}(\theta)$. Condition (1.7) is actually built using first order stochastic dominance. Since $\underline{\underline{b}} := \max\{0, \underline{b}\}$ and $\overline{\overline{b}} =: \min\{1, \overline{b}\}$ are c.d.f's, condition (1.7) is equivalently represented as that u dominates $\overline{\overline{b}}$ and is dominated by $\underline{\underline{b}}$ in first order. In a special case, a decision maker may choose a benchmark utility function $u_{\#} \in \mathfrak{U}_2$, using which boundary conditions (1.6) and (1.7) may be expressed by

$$\rho_1 u'_{\#} \leq u' \leq \rho_2 u'_{\#}, \tag{2.6}$$

$$\rho_3 u_{\#} \leq u \leq \rho_4 u_{\#}, \tag{2.7}$$

where $\rho_1,\ \rho_3 \in [0, 1]$ and $\rho_2,\ \rho_4 \geq 1$.

We now use a portfolio investment decision problem to illustrate the construction of the boundary conditions (1.6) and (1.7) of a decision maker's utility. Assume that a decision maker's risk attitude is well characterized by the CRRA utility function; however, he hesitates in choosing a coefficient of relative risk aversion.

Example 2.2 *Boundary conditions of investment utility constructed by CRRA.*

John plans to invest \$1 million into the index fund market using help from Mary, an index fund manager. His best hope is a 200% money return, while he understands that it is possible for him to lose all investment. Hence, Mary wants to learn John's risk preference in the range between 0 and \$2 million.

In financial markets investors exhibit different individual risk preferences depending on such factors as age, wealth, income, and education (e.g. Riley Jr. and Chow (1992) and Lhabitant (2006)). A large volume of research suggests that CRRA describe an individual investment attitude for forming a portfolio with risky and risk-free assets (e.g. Schooley and Worden (1996), Holt and

Laury (2002), and Brunello (2002)). The CRRA utility function in this example is a power utility function with constant coefficient of relative risk aversion $\gamma_r \in [0, 1)$

$$u_r(t) = (t/2)^{1-\gamma_r}, \quad t \in [0, 2].$$

According to the report in Holt and Laury (2002), we classify γ_r as: (i) risk neutrality by $0 < \gamma_r < 0.15$, (ii) slight risk aversion by $0.15 < \gamma_r < 0.41$, (iii) risk aversion by $0.41 < \gamma_r < 0.68$, and (iv) very risk aversion by $0.68 < \gamma_r < 1$.

John clearly claims his risk-averse attitude; however, knowing little about the concept of utility, he is not sure about the exact value of γ_r. Hence, Mary ranges γ_r from 0.41 to 0.68. We can describe this case by the boundary condition of John's utility as

$$(t/2)^{1-0.41} \leq u(t) \leq (t/2)^{1-0.68}, \quad t \in [0, 2]. \tag{2.8}$$

An alternative choice is to construct bounds of the marginal utility u' as

$$0.295(t/2)^{-0.41} \leq u'(t) \leq 0.16(t/2)^{-0.68}, \quad t \in [0, 2]. \tag{2.9}$$

Note that condition (2.8) is equivalent to

$$\int_0^t 0.295(s/2)^{-0.41} ds \leq \int_0^t u'(s) ds \leq \int_0^t 0.16(s/2)^{-0.68} ds, \quad t \in [0, 2],$$

which is implied by condition (2.9). □

2.3 Motivation and Interpretation of the General Conditions

We now motivate the general conditions (1.8) to describe a utility set. A decision maker is often asked to evaluate the expected utility of a random output $\hat{\xi}$ directly. He may not be very confident in a single-valued estimation; however, he is inclined to suggest a range $[\underline{\beta}, \overline{\beta}]$ allowing for errors, i.e.,

$$\underline{\beta} \leq \mathbb{E}[u(\hat{\xi})] \leq \overline{\beta},$$

which, by the complementary relationship (2.3), is equivalently represented by

$$1 - \overline{\beta} \leq \int_\Theta F_{\hat{\xi}} du = 1 - \mathbb{E}[u(\hat{\xi})] \leq 1 - \underline{\beta}.$$

In this case, we choose

$$\begin{array}{ll} \varphi^i := F_{\hat{\xi}}, & c_i := 1 - \underline{\beta}, \\ \varphi^{i+1} := -F_{\hat{\xi}}, & c_{i+1} := \overline{\beta} - 1. \end{array} \tag{2.10}$$

We now present a practical use of (2.10). Let us first give an alternative way to obtain bounds of u if a decision maker is only interested in some discrete points in Θ, e.g., $0 < t_1 < \cdots < t_q < \Theta$. This particular case often happens when using non-parametric utility assessments such as standard and paired gamble methods surveyed by Farquhar (1984). We only need to be concerned about the bounds of u at these points. Let these bounds be given as

$$\underline{\beta}_k \leq u(t_k) \leq \overline{\beta}_k, \quad k = 1, \ldots, q.$$

In this case, $\hat{\xi}^k$ in (2.10) are chosen to be constants t_k and their c.d.f's $F_{\hat{\xi}^k} = \mathbf{1}\{t \geq t_k\}$ are step functions. Let us further develop John's utility set in Example 2.3 to illustrate this construction.

Example 2.3 *The general conditions built by the certainty equivalent method.*

Mary wants to learn more about John's risk attitude beyond that described by the boundary conditions (2.8) and (2.9). Mary decides to use the certainty equivalent method to assist John to analyze his risk preference. John is asked to answer 25%, 50%, and 75% certainty equivalents of lottery of either losing anything or wining 200% investment return. Instead of a point estimation, John is allowed to suggest a range of each certainty equivalent, which hedges estimation errors. This approach also alleviates difficulties in precisely answering questions needed by the certainty equivalent method. Note that a similar situation arise in a group decision setting. A natural approach is to treat personal choices as i.i.d. samples and a confidence interval is thus built to be a range of values representing this group. Assume that $[0.16, 0.24]$, $[0.46, 0.54]$, and $[0.96, 1.04]$ are the ranges of the 25%, 50%, and 75% certainty equivalents given by John. General conditions on utility are then built by

$$\begin{array}{ll} u(0.16) \leq 0.25, & u(0.24) \geq 0.25, \\ u(0.46) \leq 0.5, & u(0.54) \geq 0.5, \\ u(0.96) \leq 0.75, & u(1.04) \geq 0.75, \end{array} \tag{2.11}$$

which are formulated by (1.8) as

$$\begin{array}{ll} \int_0^2 \mathbf{1}\{t \geq 0.16\} du(t) \leq 0.75, & \int_0^2 -\mathbf{1}\{t \geq 0.24\} du(t) \leq -0.75, \\ \int_0^2 \mathbf{1}\{t \geq 0.46\} du(t) \leq 0.5, & \int_0^2 -\mathbf{1}\{t \geq 0.54\} du(t) \leq -0.5, \\ \int_0^2 \mathbf{1}\{t \geq 0.96\} du(t) \leq 0.25, & \int_0^2 -\mathbf{1}\{t \geq 1.04\} du(t) \leq -0.25. \end{array}$$

□

Let us consider another use of (1.8) in the pairwise comparison of two random outcomes discussed in (2.2). Farquhar (1984) describes it as the preference comparison in paired gamble methods. A preference of $\hat{\xi}_1$ over $\hat{\xi}_2$ is represented by

$$\mathbb{E}[u(\hat{\xi}_1)] \geq \mathbb{E}[u(\hat{\xi}_2)],$$

which, by the complementary relationship (2.3), is equivalent to

$$\int_\Theta \left(F_{\hat{\xi}_1} - F_{\hat{\xi}_2} \right) du \leq 0. \tag{2.12}$$

By letting

$$\varphi^i := F_{\hat{\xi}_1} - F_{\hat{\xi}_2}, \quad c_i = 0, \tag{2.13}$$

the inequality (2.12) is represented by (1.8).

Example 2.4 *The general conditions built by the preference comparison method.*

Besides the certainty equivalent method in Example 2.3, Mary likes to use the preference comparison Method, which she feels is very easy for her customers to answer without needing special training. For example, John is now requested to choose his favorite investment outcome in each group of lotteries:

I (a) 0 and \$1 million with the same probability 0.5; (b) \$0.2 million and \$1.2 million with probabilities 0.7 and 0.3 respectively;

II (c) \$1 million and \$2 million with the same probability 0.5; (d) \$0.8 million and \$1.8 million with probabilities 0.3 and 0.7 respectively.

Assume that (b) and (d) are John's preferences in groups I and II, described by general conditions as

$$\begin{aligned} 0.7u(0.2) + 0.3u(1.2) &\geq 0.5u(0) + 0.5u(1), \\ 0.3u(0.8) + 0.7u(1.8) &\geq 0.5u(1) + 0.5u(2). \end{aligned} \tag{2.14}$$

By (2.12), we rewrite (2.14) as

$$\int_0^2 (0.7 \times \mathbf{1}\{t \geq 0.2\} + 0.3 \times \mathbf{1}\{t \geq 1.2\} - 0.5 \times \mathbf{1}\{t \geq 1\} - 0.5 \times \mathbf{1}\{t \geq 0\}) \ du(t) \leq 0,$$

$$\int_0^2 (0.3 \times \mathbf{1}\{t \geq 0.8\} + 0.7 \times \mathbf{1}\{t \geq 1.8\} - 0.5 \times \mathbf{1}\{t \geq 1\} - 0.5 \times \mathbf{1}\{t \geq 2\}) \ du(t) \leq 0.$$

□

With the probabilistic interpretation, our construction of the utility set $\mathcal{U}$ has several common features to the one given in Shapiro and Ahmed (2004). We interpret the boundary condition (1.6) on marginal utility as the preference relationship of probability measures, which Shapiro and Ahmed (2004) use to describe an uncertain distribution set. The general conditions (1.8) also correspond to the moment inequality constraints given in Shapiro and Ahmed (2004). However, the boundary condition (1.7) on utility and the restriction of $\mathcal{U}$ to a class of risk averse utilities, which is also natural in our setting, are not studied by Shapiro and Ahmed (2004), because of the difference in problem motivations. Shapiro and Ahmed (2004) give a Lagrangian based Sample Average Approximation approach to solve their model. In contrast, we present a linear programming reformulation based approach for our problem in Sections 3 and 4. Such a reformulation is possible because the utility functions considered in this paper are one-dimensional so that a natural discretization of the problem is allowable.

3 Risk-Averse Utility Decision Making with Piecewise Linear Specification

In this section we study the decision making framework (1.1). We assume throughout this section that the random function $\xi(\cdot)$ follows a discrete distribution. Let

$$P(\xi(x) = \xi_j(x)) = p_j, \quad j = 1, \ldots, M. \tag{3.1}$$

Let $\underline{b}$ and $\overline{b}$ be piecewise linear functions and $\underline{a}$, $\overline{a}$ and φ^i be step functions. Let $0 = t_0 < \cdots < t_L = \theta$ be the ordered sequence of the break points of $\underline{b}$, $\overline{b}$, and the points of discontinuity of $\underline{a}$, $\overline{a}$, and φ^i.

More specifically, these functions are described as

$$\underline{a}(t) := \sum_{j=0}^{L-1} \underline{a}(t_j)\mathbf{1}\{t_j \le t < t_{j+1}\}, \tag{3.2}$$

$$\overline{a}(t) := \sum_{j=0}^{L-1} \overline{a}(t_{j+1})\mathbf{1}\{t_j < t \le t_{j+1}\}, \tag{3.3}$$

$$\underline{b}(t) := \sum_{j=0}^{L-1} \left(\frac{\underline{b}(t_{j+1}) - \underline{b}(t_j)}{t_{j+1} - t_j} t + \frac{t_{j+1}\underline{b}(t_j) - t_j\underline{b}(t_{j+1})}{t_{j+1} - t_j} \right) \mathbf{1}\{t_j \le t < t_{j+1}\}, \tag{3.4}$$

$$\overline{b}(t) := \sum_{j=0}^{L-1} \left(\frac{\overline{b}(t_{j+1}) - \overline{b}(t_j)}{t_{j+1} - t_j} t + \frac{t_{j+1}\overline{b}(t_j) - t_j\overline{b}(t_{j+1})}{t_{j+1} - t_j} \right) \mathbf{1}\{t_j \le t < t_{j+1}\}, \tag{3.5}$$

$$\varphi^i(t) := \sum_{j=0}^{L-1} \varphi^i(t_j)\mathbf{1}\{t_j \le t < t_{j+1}\}, \quad i = 1, \ldots, m. \tag{3.6}$$

In the two general constructions given in (2.10) and (2.13), the functions φ^i are step functions if all random outcomes $\hat{\xi}_i$ selected in these constructions have discrete distributions. The following theorem shows that, for each $x \in \mathcal{X}$, $\pi(x)$ can be evaluated by solving a linear program. This result also implies that optimal utility is a piecewise linear function.

Theorem 3.1 *For a given $x \in \mathcal{X}$, the problem* (1.2) *with the utility set $\mathcal{U}$ given in* (1.10) *with further specification in* (3.2)-(3.6) *is equivalent to*

$$\pi(x) = \max_{\mu,\nu,\alpha,\beta,\tau,\delta,\eta,\lambda} \; -\mu_{L-1} - \nu_{L-1} + \sum_{j=0}^{L-1} \left(\overline{a}(t_{j+1})\alpha_j + \max\{\underline{a}(t_j), 0\}\beta_j \right)$$
$$+ \sum_{j=1}^{L-1} \left(\overline{b}(t_j)\tau_j + \underline{b}(t_j)\delta_j \right) + \sum_{i=1}^{m} \left[\varphi^i(t_{L-1}) - c_i \right] \eta_i + \sum_{k=1}^{M} \lambda_{kL} \tag{3.7a}$$

$$s.t. \; \sum_{j=1}^{L} t_j \lambda_{kj} \le p_k \xi_k(x), \quad k = 1, \ldots, M, \tag{3.7b}$$

$$\sum_{j=1}^{L} \lambda_{kj} \le p_k, \quad k = 1, \ldots, M, \tag{3.7c}$$

$$\mu_{j-1} + \nu_{j-1} - \mu_j - \nu_j + \tau_j + \delta_j + \sum_{i=1}^{m} \left[\varphi^i(t_j) - \varphi^i(t_{j-1}) \right] \eta_i - \sum_{k=1}^{M} \lambda_{kj} \le 0,$$
$$j = 1, \ldots, L-1, \tag{3.7d}$$

$$-(t_{j+1} - t_j)\mu_j - (t_j - t_{j-1})\nu_{j-1} + \alpha_j + \beta_{j-1} \le 0, \quad j = 0, \ldots, L, \tag{3.7e}$$

$$\mu_j \le 0, \; \nu_j \ge 0, \; \alpha_j \le 0, \; \beta_j \ge 0, \quad j = 0, \ldots, L-1, \tag{3.7f}$$

$$\tau_j \le 0, \; \delta_j \ge 0, \quad j = 1, \ldots, L-1, \tag{3.7g}$$

$$\eta_i \ge 0, \quad i = 1, \ldots, m, \tag{3.7h}$$

$$\lambda_{kj} \ge 0, \quad k = 1, \ldots, M, \; j = 1, \ldots, L, \tag{3.7i}$$

where $\mu_L = \nu_{-1} = \alpha_L = \beta_{-1} = 0$ are constants.

Let $w_0(x) \equiv 0$, $w_L(x) \equiv 1$, and $w_j(x)$, $j = 1, \ldots, L-1$, be the optimal dual solutions with respect to constraint (3.7d)*. Then*

$$u^*(x,t) := \sum_{j=0}^{L-1} \left(\frac{w_{j+1}(x) - w_j(x)}{t_{j+1} - t_j} t + \frac{t_{j+1} w_j(x) - t_j w_{j+1}(x)}{t_{j+1} - t_j} \right) \mathbf{1}\{t_j \le t < t_{j+1}\}, \tag{3.8}$$

is an optimal utility of problem (1.2)*.*

Proof: We first prove that problem (1.2) can be equivalently represented by

$$
\begin{aligned}
\pi(x) = \min_{r,s,y,z} \; & \sum_{k=1}^{M} p_k (s_k \xi_k(x) + r_k) && (3.9a)\\
\text{s.t. } & y_{j+1} - y_j - z_j(t_{j+1} - t_j) \le 0, \quad j = 0, \ldots, L-1, && (3.9b)\\
& y_{j+1} - y_j - z_{j+1}(t_{j+1} - t_j) \ge 0, \quad j = 0, \ldots, L-1, && (3.9c)\\
& z_j \le \overline{a}(t_{j+1}), \quad j = 0, \ldots, L-1, && (3.9d)\\
& z_{j+1} \ge \max\{\underline{a}(t_j),\ 0\}, \quad j = 0, \ldots, L-1, && (3.9e)\\
& y_j \le \overline{b}(t_j), \quad j = 1, \ldots, L-1, && (3.9f)\\
& y_j \ge \underline{b}(t_j), \quad j = 1, \ldots, L-1, && (3.9g)\\
& \sum_{j=1}^{L-1} (\varphi^i(t_j) - \varphi^i(t_{j-1})) y_j \ge \varphi^i(t_{L-1}) - c_i, \quad i = 1, \ldots, m, && (3.9h)\\
& t_j s_k + r_k - y_j \ge 0, \quad k = 1, \ldots, M,\ j = 1, \ldots, L, && (3.9i)\\
& s_k \ge 0,\ r_k \ge 0, \quad k = 1, \ldots, M, && (3.9j)
\end{aligned}
$$

where $y_0 = 0$ and $y_L = 1$ are constants.

Let $y_j := u(t_j)$, $j = 1, \ldots, L$. Note $y_0 = u(0) \equiv 0$, and $y_L = u(\theta) \equiv 1$ for all $u \in \mathfrak{U}_2$. Since

$$\int_\Theta \varphi^i du = \sum_{j=0}^{L-1} \varphi^i(t_j)(y_{j+1} - y_j) = \varphi^i(t_{L-1}) - \sum_{j=1}^{L-1} y_j (\varphi^i(t_j) - \varphi^i(t_{j-1})),$$

we may rewrite the general conditions (1.8) in the utility set $\mathcal{U}$ as (3.9h).

For a given $\hat{u} \in \mathcal{U}$, we define a piecewise linear increasing concave function

$$\hat{u}_L(t) := \sum_{j=0}^{L-1} \left(\frac{\hat{u}(t_{j+1}) - \hat{u}(t_j)}{t_{j+1} - t_j} t + \frac{t_{j+1} \hat{u}(t_j) - t_j \hat{u}(t_{j+1})}{t_{j+1} - t_j} \right) \mathbf{1}\{t_j \le t < t_{j+1}\}.$$

It is straightforward to see that $\hat{u}_L$ satisfies conditions (3.9h) and $\hat{u}_L \le \hat{u} \le \overline{b}$ on Θ. Since $\hat{u} \ge \underline{b}$, we have $\hat{u}(t_j) \ge \underline{b}(t_j)$ for all $j = 0, \ldots, L$. For a given $\hat{t} \in [t_j, t_{j+1}]$ and some j, we have

$$\hat{u}_L(\hat{t}) = \frac{(\hat{t} - t_j)\hat{u}(t_j) + (t_{j+1} - \hat{t})\hat{u}(t_{j+1})}{t_{j+1} - t_j} \ge \frac{(\hat{t} - t_j)\underline{b}(t_j) + (t_{j+1} - \hat{t})\underline{b}(t_{j+1})}{t_{j+1} - t_j} = \underline{b}(\hat{t}),$$

and by concavity of $\hat{u}'$,

$$\hat{u}_L'(\hat{t}) = \frac{\hat{u}(t_{j+1}) - \hat{u}(t_j)}{t_{j+1} - t_j} \le \hat{u}'(t_j) \le \overline{a}(t_{j+1}) = \overline{a}(\hat{t}).$$

On the other hand

$$\hat{u}'_L(\hat{t}) \geq \hat{u}'(t_{j+1}^-) \geq \underline{a}(t_j) = \underline{a}(\hat{t}).$$

It follows that $\hat{u}_L \in \mathcal{U}$.

The above discussions also lead to the fact that an optimal solution u^* of problem (1.2) belongs to the set $\widetilde{\mathcal{U}}$ of piecewise linear increasing concave functions with break points in the set $\{t_0, \ldots, t_L\}$. Otherwise, we can construct $u^*_L \in \widetilde{\mathcal{U}} \subseteq \mathcal{U}$ which is below u^*. Therefore, problem (1.2) is equivalent to

$$\min_{u \in \widetilde{\mathcal{U}}} \Pi(x, u).$$

We now describe the set $\widetilde{\mathcal{U}}$. Conditions (3.9b) and (3.9c) imply

$$z_0 \geq \frac{y_1 - y_0}{t_1 - t_0} \geq z_1 \cdots \geq z_{L-1} \geq \frac{y_L - y_{L-1}}{t_L - t_{L-1}} \geq z_L,$$

which ensure the increasing and concavity property of $u \in \widetilde{\mathcal{U}}$ when all $z_j \geq 0$. Conditions (3.9d)-(3.9e) represent the boundedness of u' and also imply that all y_j and z_j are nonnegative. Conditions (3.9f)-(3.9g) consider the boundedness of u. Since u is a piecewise linear increasing concave function with break points t_j and values y_j, $j = 0, \ldots, L$, it can be written as the minimization of all subgradients at points (t_j, y_j)

$$\begin{aligned} u(v) = \min_{s,r} \ & vs + r \\ \text{s.t. } & t_j s + r - y_j \geq 0, \quad j = 1, \ldots, L, \\ & s \geq 0, \ r \geq 0. \end{aligned}$$

This gives the objective (3.9a) and conditions (3.9i) and (3.9j).

Let μ, ν, α, β, τ, δ, η, and λ to be the dual variables of condition (3.9b)-(3.9i) respectively. Then (3.7) is the dual problem of (3.9). Note that in this transformation we consider the implied nonnegativity of y_j and z_j.

Let y^*_j be the y-component of an optimal solution of (3.9). Note $y^*_0 = y_0 = 0$ and $y^*_L = y_L = 1$. For this given x, an optimal utility of the problem (1.2) is represented by

$$u^*(x, t) := \sum_{j=0}^{L-1} \left(\frac{y^*_{j+1} - y^*_j}{t_{j+1} - t_j} t + \frac{t_{j+1} y^*_j - t_j y^*_{j+1}}{t_{j+1} - t_j} \right) \mathbf{1}\{t_j \leq t < t_{j+1}\}.$$

We also note that the above discussion tells us that $w_j(x) = y^*_j$ for every j. □

Corollary 3.2 *The problem* (1.1) *with the utility set* $\mathcal{U}$ *given in* (1.10) *with further specification* *in* (3.2)-(3.6) *is equivalent to*

$$
\begin{aligned}
\max_{x,\mu,\nu,\alpha,\beta,\tau,\delta,\eta,\lambda} \quad & -\mu_{L-1} - \nu_{L-1} + \sum_{j=0}^{L-1} \left(\overline{a}(t_{j+1})\alpha_j + \max\{\underline{a}(t_j), 0\}\beta_j\right) \\
& \quad + \sum_{j=1}^{L-1} \left(\overline{b}(t_j)\tau_j + \underline{b}(t_j)\delta_j\right) + \sum_{i=1}^{m} \left[\varphi^i(t_{L-1}) - c_i\right]\eta_i + \sum_{k=1}^{M} \lambda_{kL} \\
s.t. \quad & \sum_{j=1}^{L} t_j \lambda_{kj} \le p_k \xi_k(x), \quad k = 1, \ldots, M, \\
& \sum_{j=1}^{L} \lambda_{kj} \le p_k, \quad k = 1, \ldots, M, \\
& \mu_{j-1} + \nu_{j-1} - \mu_j - \nu_j + \tau_j + \delta_j + \sum_{i=1}^{m} \left[\varphi^i(t_j) - \varphi^i(t_{j-1})\right]\eta_i - \sum_{k=1}^{M} \lambda_{kj} \le 0, \\
& \qquad j = 1, \ldots, L-1, \\
& -(t_{j+1} - t_j)\mu_j - (t_j - t_{j-1})\nu_{j-1} + \alpha_j + \beta_{j-1} \le 0, \quad j = 0, \ldots, L, \\
& \mu_j \le 0, \ \nu_j \ge 0, \ \alpha_j \le 0, \ \beta_j \ge 0, \quad j = 0, \ldots, L-1, \\
& \tau_j \le 0, \ \delta_j \ge 0, \quad j = 1, \ldots, L-1, \\
& \eta_i \ge 0, \quad i = 1, \ldots, m, \\
& \lambda_{kj} \ge 0, \quad k = 1, \ldots, M, \ j = 1, \ldots, L, \\
& x \in \mathcal{X},
\end{aligned}
$$

where $\mu_L = \nu_{-1} = \alpha_L = \beta_{-1} = 0$ *are constants.*

Remark 3.3 *The set* $\mathcal{U}$ *combines the boundary conditions* (1.6) *and* (1.7) *on utility and the general conditions* (1.8) *on utility. However, not all of these conditions are necessary in a practical case. The formulation* (3.10) *in the following particular cases are given by fixing certain variables:*

1. *if* $u' \le \overline{a}$ *in the boundary condition* (1.6) *on marginal utility is not used, then* $\alpha_j = 0$, $j = 0, \ldots, L-1$;
2. *if* $u' \ge \underline{a}$ *in the boundary condition* (1.6) *on marginal utility is not used, then* $\beta_j = 0$, $j = 0, \ldots, L-1$;
3. *if* $u \le \overline{b}$ *in the boundary condition* (1.7) *on utility is not used, then* $\tau_j = 0$, $j = 1, \ldots, L-1$;
4. *if* $u \ge \underline{b}$ *in the boundary condition* (1.7) *on utility is not used, then* $\delta_j = 0$, $j = 1, \ldots, L-1$;
5. *if the uncertainty condition* (1.8) *on utility is not used, then* $\eta_i = 0$, $i = 1, \ldots, m$.

4 Risk-Averse Utility Decision Making with a General Specification

We now consider the general case under the following assumptions:

(A1). $\xi(x)$ is concave in a neighborhood of $\mathcal{X}$ bounded in Θ a.s..

(A2). $\underline{a}$ is lower continuous, and $\overline{a}$ is upper continuous; either $\overline{a}(0) < \infty$ or $\overline{b}(0) = 0$ holds.

(A3). φ^i for $i = 1, \ldots, m$ are bounded functions with finitely many points of discontinuity at which φ^i are RCLL.

We give a constructive approach to find a solution of (1.1) under Assumptions (A1)-(A3). This approach discretizes and thus approximates problem (1.1), and uses the results in Section 3. More specifically, we show that the sequence of optimal values and sets of optimal solutions of the approximation problems converge to the true counterpart.

Let $\mathfrak{P}(L) := \{t_0, \ldots, t_L\}$ be a partition of Θ ($0 = t_0 < \cdots < t_L = \theta$), which is nested with increasing L, i.e., $\mathfrak{P}(L_1) \subset \mathfrak{P}(L_2)$ if $L_1 < L_2$. Also, Let the longest subinterval in $\mathfrak{P}(L)$, $\max_{j\in\{0,\ldots,L-1\}} t_{j+1} - t_j$, approach 0 as L increases to ∞. Let $\varphi^i_\Delta(t_j) := \inf_{t\in[t_j,t_{j+1})} \varphi^i(t)$. On $\mathfrak{P}(L)$, we denote

$$\underline{a}_L(t) := \begin{cases} \underline{a}(t^-_{j+1}) & t_j \le t < t_{j+1},\ j = 0, \ldots, L-1, \\ \underline{a}(\theta) & t = t_L, \end{cases} \tag{4.1}$$

$$\overline{a}_L(t) := \begin{cases} \overline{a}(0) & t = t_0, \\ \overline{a}(t^+_j) & t_j < t \le t_{j+1}, \quad j = 0, \ldots, L-1, \end{cases} \tag{4.2}$$

$$\underline{b}_L(t) := \begin{cases} \frac{\underline{b}(t_{j+1}) - \underline{b}(t_j)}{t_{j+1} - t_j} t + \frac{t_{j+1}\underline{b}(t_j) - t_j \underline{b}(t_{j+1})}{t_{j+1} - t_j} & t_j \le t < t_{j+1},\ j = 0, \ldots, L-1, \\ \underline{b}(\theta) & t = t_L. \end{cases} \tag{4.3}$$

and for $i = 1, \ldots, m$,

$$\varphi^i_L(t) := \begin{cases} \varphi^i_\Delta(t_j) & t_j \le t < t_{j+1},\ j = 0, \ldots, L-1, \\ \varphi^i(\theta) & t = t_L. \end{cases} \tag{4.4}$$

Note that $\underline{a}_L$, $\overline{a}_L$, $\underline{b}_L$, and φ^i_L are the approximation functions of $\underline{a}$, $\overline{a}$, $\underline{b}$, and φ^i based on the partition $\mathfrak{P}(L)$. Note that $\underline{b}_L$ uniformly converges to $\underline{b}$, and under Assumptions (A2), $\underline{a}_L$ and $\overline{a}_L$ uniformly converge to $\underline{a}$ and $\overline{a}$. Assumption (A3) also implies that φ^i_L converges to φ^i a.e.. Let us next construct a piecewise linear envelope function $\overline{b}_L$ of $\overline{b}$ on $\mathfrak{P}(L)$ which uniformly converges to $\overline{b}$ from above. We denote

$$\begin{aligned} \hat{t}_0 &:= 0, & \hat{y}_0 &:= \overline{b}(t_1) - \overline{b}'(t_1)t_1, \\ \hat{t}_{L+1} &:= \theta, & \hat{y}_{L+1} &:= \overline{b}(\theta), \end{aligned}$$

and for any two successive points t_{j-1} and t_j in $\mathfrak{P}(L)$, $j = 1, \ldots, L$,

$$\begin{aligned} \hat{t}_j &:= \frac{\overline{b}(t_j) - \overline{b}(t_{j-1}) - \overline{b}'(t_j)t_j + \overline{b}'(t_{j-1})t_{j-1}}{\overline{b}'(t_{j-1}) - \overline{b}'(t_j)}, \\ \hat{y}_j &:= \frac{\overline{b}'(t_{j-1})\overline{b}(t_j) - \overline{b}'(t_j)\overline{b}(t_{j-1}) - \overline{b}'(t_{j-1})\overline{b}'(t_j)(t_j - t_{j-1})}{\overline{b}'(t_{j-1}) - \overline{b}'(t_j)}, \end{aligned}$$

using which $\overline{b}_L$ is represented as

$$\overline{b}_L(t) := \begin{cases} \frac{\hat{y}_{j+1} - \hat{y}_j}{\hat{t}_{j+1} - \hat{t}_j} t + \frac{\hat{t}_{j+1}\hat{y}_j - \hat{t}_j \hat{y}_{j+1}}{\hat{t}_{j+1} - \hat{t}_j} & \hat{t}_j \le t < \hat{t}_{j+1},\ j = 0, \ldots, L, \\ \hat{y}_{L+1} & t = \hat{t}_{L+1}. \end{cases} \tag{4.5}$$

Based on (4.1), (4.2), (4.3), (4.4), and (4.5), we let

$$\mathcal{U}(L) := \left\{ u \in \mathfrak{U}_2 \;\middle|\; \begin{array}{l} \underline{a}_L \le u' \le \overline{a}_L, \\ \underline{b}_L \le u \le \overline{b}_L, \\ \int_\Theta \varphi^i_L du \le c_i,\ i = 1, \ldots, m. \end{array} \right\} \tag{4.6}$$

Theorem 3.1 gives a linear programming formulation of the problem

$$\pi_L(x) := \min_{u \in \mathcal{U}(L)} \Pi(x, u). \tag{4.7}$$

In Theorem 4.11 below we show the asymptotic convergence of the optimal value and the set of optimal solutions of the approximation problem

$$\max_{x\in\mathcal{X}} \pi_L(x) \tag{4.8}$$

to those of the true problem (1.1, 1.2, and 1.10). Before we state this theorem, we need to prove some preliminary results.

Lemma 4.1 *If Assumption (A1) holds, then $\pi(x)$ and $\pi_L(x)$ are continuous concave functions over $\mathcal{X}$.*

Proof: By Assumption (A1) and the monotonic increase of $u \in \mathcal{U}$, it follows that $u(\xi(x))$ is concave a.s. in a neighborhood of $\mathcal{X}$ denoted by $\mathcal{N}(\mathcal{X})$, and thus $\Pi(x,u)$ are concave in $\mathcal{N}(\mathcal{X})$. For given $x_1,\ x_2 \in \mathcal{N}(\mathcal{X})$ and $\lambda \geq 0$, we have

$$\begin{aligned}
\pi(\lambda x_1 + (1-\lambda)x_2) &= \min_{u\in\mathcal{U}} \Pi(\lambda x_1 + (1-\lambda)x_2, u)\\
&\geq \min_{u\in\mathcal{U}} \lambda\Pi(x_1,u) + (1-\lambda)\Pi(x_2,u)\\
&\geq \lambda \min_{u\in\mathcal{U}} \Pi(x_1,u) + (1-\lambda)\min_{u\in\mathcal{U}} \Pi(x_2,u)\\
&= \lambda\pi(x_1) + (1-\lambda)\pi(x_2).
\end{aligned}$$

Since $\pi(x)$ is concave in $\mathcal{N}(\mathcal{X})$, $\pi(x)$ is continuous in $\mathcal{X} \subset \mathcal{N}(\mathcal{X})$.

A similar proof can also be applied for the continuity and concavity of $\pi_L(x)$. □

Lemma 4.2 *Let*

$$g^i(u) := \int_\Theta \varphi^i du - c_i, \quad i = 1,\ldots,m, \tag{4.9}$$

$$g_L^i(u) := \int_\Theta \varphi_L^i du - c_i, \quad i = 1,\ldots,m. \tag{4.10}$$

If Assumption (A3) holds, then for $i = 1,\ldots,m$,

$$\lim_{L\to\infty} \sup_{u\in\mathfrak{U}_2} |g_L^i(u) - g^i(u)| = 0.$$

Proof: By assumption (A3), each $\varphi^i(t)$ is right continuous at $t = 0$. In addition, $\varphi^i(t)$ has finitely many points of discontinuity. These facts imply that there exists $\epsilon > 0$ such that $\varphi^i(t)$ is continuous at $(0,\epsilon)$. It follows that, for any given $\delta > 0$, there exists $\hat{L} > 0$ such that for all $L \geq \hat{L}$,

$$\left|\varphi_L^i(t) - \varphi^i(t)\right| \leq \delta, \quad t \in [0,\epsilon).$$

Thus, we have

$$\sup_{u\in\mathfrak{U}_2} \int_0^\epsilon \left|\varphi_L^i - \varphi^i\right| du \leq \delta \sup_{u\in\mathfrak{U}_2} \int_0^\epsilon du \leq \delta.$$

Increasing and concavity property of any $u \in \mathfrak{U}_2$ implies that $u'(t) \leq u'(\epsilon) \leq 1/\epsilon$ for $t \in [\epsilon,\theta]$, which leads to

$$\sup_{u\in\mathfrak{U}_2} \int_\epsilon^\theta \left|\varphi_L^i - \varphi^i\right| du = \sup_{u\in\mathfrak{U}_2} \int_\epsilon^\theta \left|\varphi_L^i - \varphi^i\right| u' d\ell \leq \frac{1}{\epsilon}\int_\epsilon^\theta \left|\varphi_L^i - \varphi^i\right| d\ell.$$

Thus, for $L \geq \hat{L}$, we have

$$\begin{aligned}\sup_{u \in \mathfrak{U}_2} |g_L^i(u) - g^i(u)| &= \sup_{u \in \mathfrak{U}_2} \left| \int_\Theta \varphi_L^i du - \int_\Theta \varphi^i du \right| \\ &\leq \sup_{u \in \mathfrak{U}_2} \int_0^\epsilon \left| \varphi_L^i - \varphi^i \right| du + \sup_{u \in \mathfrak{U}_2} \int_\epsilon^\theta \left| \varphi_L^i - \varphi^i \right| du \\ &\leq \delta + \frac{1}{\epsilon} \int_\epsilon^\theta \left| \varphi_L^i - \varphi^i \right| d\ell.\end{aligned}$$

Also by Assumption (A3), it follows that φ_L^i converges to φ^i a.e., and the difference between φ_L^i and φ^i is bounded. Using the Lebesgue Dominated Theorem, we have $\int_\epsilon^\theta \left|\varphi_L^i - \varphi^i\right| d\ell \to 0$ as $L \to \infty$. $\square$

The following four lemmas are standard results in probability theory. They are stated here for completeness.

Lemma 4.3 (The Prohorov-Varadarajan theorem) *(or see Theorem 9.2.4 in Athreya and Lahiri (2006)) Let $\{F_n\}_{n\geq 1}$ be a sequence of one-dimensional c.d.f's. The sequence $\{F_n\}_{n\geq 1}$ is tight if and only if given any subsequence $\{F_{n_i}\}_{i\geq 1}$ of $\{F_n\}_{n\geq 1}$, there exists a further subsequence $\{F_{m_i}\}_{i\geq 1}$ of $\{F_{n_i}\}_{i\geq 1}$ and a c.d.f. F such that F_{m_i} weakly converges to F as i increases to ∞.*

Lemma 4.4 (Polyā's Theorem) *(or see Theorem 9.1.4 in Athreya and Lahiri (2006)) Let $\{F_n\}_{n\geq 1}$ be a sequence of one-dimensional c.d.f. which converges to a continuous c.d.f. F. Then the sequence $\{F_n\}_{n\geq 1}$ uniformly converges to F as n increase to ∞.*

Lemma 4.5 (Theorem 9.3.1 in Athreya and Lahiri (2006)) *The sequence $\{F_n\}_{n\geq 1}$ of one-dimensional c.d.f. converges to a c.d.f. F if and only if for every open set $\mathcal{B} \subseteq \mathbb{R}$,*

$$\liminf_{n\to\infty} F_n(\mathcal{B}) \geq F(\mathcal{B}).$$

Lemma 4.6 (The second Helly-Bray theorem) *(or see Theorem 9.2.3 in Athreya and Lahiri (2006)) The sequence $\{F_n\}_{n\geq 1}$ of one-dimensional c.d.f. converges to a c.d.f. F if and only if*

$$\int \phi dF_n \to \int \phi dF,$$

for all $\phi \in \mathcal{C}_B(\mathbb{R}) := \{g \mid g : \mathbb{R} \to \mathbb{R}$ is continuous and bounded$\}$.

Lemma 4.7 *For any sequence $\{u_L \in \mathcal{U}(L)\}$, there exists a subsequence $\{u_{L_k}\}$ which weakly converges to a c.d.f. $\hat{u}$. If Assumption (A2) holds, then there exists $\hat{u} \in \mathfrak{U}_2$ such that $\{u_{L_k}\}$ uniformly converges to $\hat{u}$.*

Proof: The sequence $\{u_L\}$ is tight ($\int_\Theta du_L = 1$) so that by Lemma 4.3, we have a subsequence $\{u_{L_k}\}$ which weakly converges to a c.d.f. $\hat{u}$.

The uniform convergence of $\{u_{L_k}\}$ to $\hat{u}$ under Assumption (A2) follows by Lemma 4.4, if $\hat{u}$ is continuous on $\mathbb{R}$. Since $\hat{u} \equiv 0$ on $(-\infty, 0)$ and $\hat{u} \equiv 1$ on $[\theta, \infty)$, the continuity of $\hat{u}$ on the real line follows from the continuity of $\hat{u}$ on Θ. Let us first show that $\hat{u}$ is continuous on $(\epsilon, \theta]$ for an arbitrarily chosen $\epsilon \in (0, \theta)$. By contradiction, we assume that $\hat{u}$ has at least a point $t_1 \in (\epsilon, \theta]$ of discontinuity. Let $\delta := \hat{u}(t_1) - \hat{u}(t_1^-) > 0$. It follows by Lemma 4.5 that, for any $t_2 \in (t_1, \infty)$ and $t_3 \in (\epsilon, t_1)$,

$$\liminf_{k\to\infty} u_{L_k}(t_2^-) - u_{L_k}(t_3) \geq \hat{u}(t_2^-) - \hat{u}(t_3) \geq \hat{u}(t_1) - \hat{u}(t_1^-) = \delta. \tag{4.11}$$

The increasing and concavity property of u_{L_k} ensures that $u'_{L_k}(t) \le u'_{L_k}(\epsilon) \le 1/\epsilon$ for $t \in (\epsilon, \infty)$. Hence, u_{L_k} is Lipschitz continuous on (ϵ, ∞) with the same Lipschitz constant $1/\epsilon$. We choose t_2 and t_3 such that $t_2 - t_3 < \delta\epsilon$, and then have

$$\limsup_{k\to\infty} u_{L_k}(t_2^-) - u_{L_k}(t_3) \le 1/\epsilon|t_2 - t_3| < \delta. \tag{4.12}$$

The contradiction between (4.11) and (4.12) shows that $\hat{u}$ is continuous in $(\epsilon, \theta]$. Since ϵ is arbitrary, it certifies the continuity of $\hat{u}$ in $(0, \theta]$. This conclusion leads to the pointwise convergence of u_{L_K} to $\hat{u}$ on $(0, \theta]$.

We next check the continuity of $\hat{u}$ at 0. The second part of Assumption (A2) is used to ensure this property. If $\overline{a}(0) < \infty$ holds, the uniform convergence of $\overline{a}_{L_k}$ to $\overline{a}$ ensured in the construction (4.2) implies that there exists constants $\alpha > 0$ and $\hat{k} > 0$ such that $\overline{a}_{L_k}(0) < \alpha$ for all $k \ge \hat{k}$. It follows that $u'_{L_k}(t) \le u'_{L_k}(0) \le \overline{a}_{L_k}(0) < \alpha$ for $t \in \Theta$ and $k \ge \hat{k}$. $\hat{u}(t) \le \alpha t$ for $t \in (0, \theta]$ is implied by the pointwise convergence of u_{L_K} to $\hat{u}$. It shows that the c.d.f. $\hat{u}$ is continuous at 0 as $\hat{u}(0) = \alpha 0 = 0$. Let us now consider the case that $\overline{b}(0) = 0$. It follows by the pointwise convergence of u_{L_K} to $\hat{u}$ and the uniform convergence of $\overline{b}_{L_k}$ to $\overline{b}$ ensured in the construction (4.5) that, for $t \in (0, \theta]$,

$$\hat{u}(t) = \lim_{k\to\infty} u_{L_k}(t) \le \lim_{k\to\infty} \overline{b}_{L_k}(t) = \overline{b}(t).$$

The fact that $\overline{b}$ is continuous at 0 results in the continuity of the c.d.f. $\hat{u}$ at 0 as $\hat{u}(0) = \overline{b}(0) = 0$.

So far, we have shown the uniform convergence of $\{u_{L_k}\}$ to $\hat{u}$ under Assumption (A2). It directly follows that $\hat{u}(0) = 0$ and $\hat{u}(\theta) = 1$. Since any u_{L_k} is an increasing concave function, for two arbitrarily chosen points $0 \le t_1 < t_2 \le \theta$, we have

$$\hat{u}(t_1) = \lim_{k\to\infty} u_{L_k}(t_1) \le \lim_{k\to\infty} u_{L_k}(t_2) = \hat{u}(t_2),$$

and for $\lambda > 0$,

$$\hat{u}(\lambda t_1 + (1-\lambda)t_2) = \lim_{k\to\infty} u_{L_k}(\lambda t_1 + (1-\lambda)t_2) \ge \lambda \lim_{k\to\infty} u_{L_k}(t_1) + (1-\lambda)u_{L_k}(t_2) = \lambda\hat{u}(t_1) + (1-\lambda)\hat{u}(t_2).$$

It means that $\hat{u}$ is an increasing concave function and satisfies the boundary condition (1.5). Thus, $\hat{u} \in \mathfrak{U}_2$. □

The following lemma shows the convergence of $\mathcal{U}(L)$ to $\mathcal{U}$ in the sense of a weak Hausdorff distance described below. Based on the set $\mathfrak{S}$ of signed measures on Θ of finite total mass, we describe the distance of utility function u to class $\mathcal{U} \subseteq \mathcal{U}_2$ as

$$d_s(u, \mathcal{U}) := \begin{cases} \inf_{\hat{u}\in\mathcal{U}} \left|\int_\Theta u ds - \int_\Theta \hat{u} ds\right| & \mathcal{U} \ne \emptyset, \\ \infty & \text{o.w.}, \end{cases} \tag{4.13}$$

for $s \in \mathfrak{S}$. It is easy to check that $d_s(u, 0) = \left|\int_\Theta u ds\right|$ for each $s \in \mathfrak{S}$ has the norm satisfying properties of positivity, subadditivity, and homogeneity, which ensure that $d_s(u_1, u_2)$ is a distance metric. Let the deviation and Hausdorff distance of class $\mathcal{U}_1$ and $\mathcal{U}_2$ in the weak topology be

$$\mathbb{D}_s(\mathcal{U}_1, \mathcal{U}_2) := \begin{cases} \sup_{u\in\mathcal{U}_1} d_s(u, \mathcal{U}_2) & \text{if } \mathcal{U}_1 \ne \emptyset, \\ 0 & \text{o.w.}, \end{cases}, \tag{4.14}$$

$$\mathbb{H}_s(\mathcal{U}_1, \mathcal{U}_2) := \max\{\mathbb{D}_s(\mathcal{U}_1, \mathcal{U}_2),\ \mathbb{D}_s(\mathcal{U}_2, \mathcal{U}_1)\}. \tag{4.15}$$

Lemma 4.8 *Suppose that Assumptions (A2) and (A3) hold. For a given $x \in \mathcal{X}$, $\mathbb{H}_s(\mathcal{U}(L), \mathcal{U}) \to 0$ for all $s \in \mathfrak{S}$ as $L \to \infty$.*

Proof: Recall the functionals $g^j(u)$ and $g^j_L(u)$ denoted as in (4.9) and (4.10). Let us denote classes

$$\begin{aligned}\mathcal{U}^1 &:= \{u \in \mathfrak{U}_2 \mid \underline{a} \le u' \le \overline{a},\ \underline{b} \le u \le \overline{b}\},\\ \mathcal{U}^2 &:= \{u \in \mathfrak{U}_2 \mid g^j(u) \le 0,\ j = 1,\dots,m\},\\ \mathcal{U}^1(L) &:= \{u \in \mathfrak{U}_2 \mid \underline{a}_L \le u' \le \overline{a}_L,\ \underline{b}_L \le u \le \overline{b}_L\},\\ \mathcal{U}^2(L) &:= \{u \in \mathfrak{U}_2 \mid g^j_L(u) \le 0,\ j = 1,\dots,m\}.\end{aligned}$$

It follows that $\mathcal{U} = \mathcal{U}^1 \cap \mathcal{U}^2$ and $\mathcal{U}(L) = \mathcal{U}^1(L) \cap \mathcal{U}^2(L)$.

We first claim $\mathbb{D}_s(\mathcal{U},\mathcal{U}(L)) \to 0$ for all $s \in \mathfrak{S}$ as $L \to \infty$. The fact that $\underline{a} \ge \underline{a}_L$, $\overline{a} \le \overline{a}_L$, $\underline{b} \ge \underline{b}_L$, and $\overline{b} \le \overline{b}_L$ implies $\mathcal{U}^1 \subseteq \mathcal{U}^1(L)$. Since $\varphi^i_L \le \varphi^i$ and any given $u \in \mathcal{U}^2$ is increasing, it follows that

$$g^i_L(\hat{u}) = \int_\Theta \varphi^i_L d\hat{u} - c_i \le \int_\Theta \varphi^i d\hat{u} - c_i = g^i(\hat{u}) \le 0,$$

which shows that $u \in \mathcal{U}^2(L)$, and then $\mathcal{U}^2 \subseteq \mathcal{U}^2(L)$. Hence, we have $\mathcal{U} \subseteq \mathcal{U}(L)$, for which $\mathbb{D}_s(\mathcal{U},\mathcal{U}(L)) \equiv 0$ for all $s \in \mathfrak{S}$ and $L > 0$.

We now argue $\mathbb{D}_s(\mathcal{U}(L),\mathcal{U}) \to 0$ for all $s \in \mathfrak{S}$ as $L \to \infty$ by contradiction. Assume that there exists $\hat{s} \in \mathfrak{S}$ such that $\mathbb{D}_{\hat{s}}(\mathcal{U}(L),\mathcal{U}) \nrightarrow 0$. Then, there is $\epsilon > 0$ with the property that for any $k > 0$ there exists $L_k \ge k$ satisfying $\mathbb{D}_{\hat{s}}(\mathcal{U}(L_k),\mathcal{U}) > \epsilon$, i.e., there exists $u_{L_k} \in \mathcal{U}(L_k)$ such that $d_{\hat{s}}(u_{L_k},\mathcal{U}) > \epsilon$. Under Assumption (A2), Lemma 4.7 shows that k and L_k can be chosen to make $\{u_{L_k}\}$ uniformly converge to a c.d.f. $\hat{u} \in \mathfrak{U}_2$.

Because of the uniform convergence of $\underline{b}_{L_k}$ and $\overline{b}_{L_k}$ to $\underline{b}$ and $\overline{b}$ respectively, we have

$$\hat{u} = \lim_{k\to\infty} u_{L_k} \ge \lim_{k\to\infty} \underline{b}_{L_k} = \underline{b},$$

and analogously $\hat{u} \le \overline{b}$. We now show that $\underline{a} \le u' \le \overline{a}$. Suppose that there exists $\hat{t}_1 \in \Theta$ such that $\underline{a}(t_1) - \hat{u}'(t_1) > \delta$ for some $\delta > 0$. Because of the right continuity of $u'(t_1)$ for $t_1 \in [0,\theta)$, there exists $t_2 > t_1$ such that $\underline{a} - \hat{u}' > \delta/2$ on $[t_1,t_2]$. If $t_1 = \theta$, we choose $t_2 < \theta$ for the left continuity of $\hat{u}'(\theta)$. Since $\underline{a}_L$ uniformly converges to $\underline{a}$ from below under Assumption (A2), it follows that $\underline{a}_{L_k} - \hat{u}' > \delta/4$ on $[t_1,t_2]$ for large enough k, and then

$$\begin{aligned}\hat{u}(t_2) - \hat{u}(t_1) &= \int_{t_1}^{t_2} \hat{u}' dt\\ &< \lim_{k\to\infty} \int_{t_1}^{t_2} \underline{a}_{L_k} dt - \delta(t_2 - t_1)/4\\ &\le \lim_{k\to\infty} \int_{t_1}^{t_2} u'_{L_k} dt - \delta(t_2 - t_1)/4\\ &= \lim_{k\to\infty} u_{L_k}(t_2) - u_{L_k}(t_1) - \delta(t_2 - t_1)/4\\ &= \hat{u}(t_2) - \hat{u}(t_1) - \delta(t_2 - t_1)/4,\end{aligned}$$

which is a contradiction. Hence, we have $\hat{u}' \ge \underline{a}$., and analogously, $\hat{u}' \le \overline{a}$. So far, we have shown that $u \in \mathfrak{U}_2$, $\underline{b} \le u \le \overline{b}$, and $\underline{a} \le u' \le \overline{a}$. It ensures that $\hat{u} \in \mathcal{U}^1$.

We now show that $\hat{u} \in \mathcal{U}^2$. Under Assumption (A3), there are finitely many points of discontinuity of φ^i for any i, denoted by $0 < z_1 < \cdots < z_{N-1} \le \theta$. Also, let $z_0 = 0$ and $z_N = \theta$. Let us denote

$$\psi^i_j(t) := \begin{cases} \varphi^i(z_j) & t < z_j,\\ \varphi^i(t) & z_j \le t < z_{j+1},\\ \varphi^i(z^-_{j+1}) & t \ge z_{j+1},\end{cases}$$

which is a bounded continuous function on $\mathbb{R}$ since φ^i satisfies RCLL. Since the measure under c.d.f. $u \in \mathfrak{U}_2$ is 0 at these points of discontinuity, we rewrite

$$g^i(u) = \sum_{j=1}^{N} \int_{z_{j-1}}^{z_j} \varphi^i du - c_i = \sum_{j=1}^{N} \int_{z_{j-1}}^{z_j} \psi_j^i du - c_i.$$

It follows that

$$\left|g^i(u_{L_k}) - g^i(\hat{u})\right| \leq \sum_{j=1}^{N} \left| \int_{z_{j-1}}^{z_j} \psi_j^i du_{L_k} - \int_{z_{j-1}}^{z_j} \psi_j^i d\hat{u} \right|.$$

We have shown that u_{L_k} uniformly converges to $\hat{u}$. By Lemma 4.6, $\int_{z_{j-1}}^{z_j} \psi_j^i du_{L_k} \to \int_{z_{j-1}}^{z_j} \psi_j^i d\hat{u}$ as $k \to \infty$. It ensures that

$$|g^i(u_{L_k}) - g^i(\hat{u})| \to 0.$$

Lemma 4.2 also shows that, as $k \to \infty$,

$$\sup_{u \in \mathfrak{U}_2} |g^i_{L_k}(u) - g^i(u)| \to 0.$$

Hence,

$$\begin{aligned}
\lim_{k\to\infty} \left|g^i_{L_k}(u_{L_k}) - g^i(\hat{u})\right| &\leq \lim_{k\to\infty} |g^i_{L_k}(u_{L_k}) - g^i(u_{L_k})| + |g^i(u_{L_k}) - g^i(\hat{u})| \\
&\leq \lim_{k\to\infty} \sup_{u\in\mathfrak{U}_2} |g^i_{L_k}(u) - g^i(u)| + |g^i(u_{L_k}) - g^i(\hat{u})| \\
&= 0.
\end{aligned}$$

Thus, we have

$$g^i(\hat{u}) = \lim_{k\to\infty} g^i_{L_k}(u_{L_k}) \leq 0.$$

Hence, $\hat{u} \in \mathcal{U}^2$. The above discussions lead to $\hat{u} \in \mathcal{U}^1 \cap \mathcal{U}^2 = \mathcal{U}$, which contradicts the fact that $d_{\hat{s}}(\hat{u}, \mathcal{U}) \geq \epsilon > 0$. □

Lemma 4.9 *If Assumptions (A2) and (A3) hold, then $\pi_L(x) \to \pi(x)$ as $L \to \infty$ for all $x \in \mathcal{X}$.*

Proof: Let us denote by $\mathcal{V}$ and $\mathcal{V}(L)$ the classes of optimal solution functions of problems (1.2) and (4.7). Suppose that $\mathcal{U}$ is a nonempty class. For $u^* \in \mathcal{V}$, let $u_L = \arg\min_{u\in\mathcal{U}(L)} d_{F_{\xi(x)}}(u, u^*)$. Since $\mathbb{H}_s(\mathcal{U}(L), \mathcal{U}) \to 0$ for all $s \in \mathfrak{S}$ by Lemma 4.8, it follows that

$$0 = \lim_{L\to\infty} d_{F_{\xi(x)}}(u_L, u^*) = \left| \int_\Theta u_L dF_{\xi(x)} - \int_\Theta u^* dF_{\xi(x)} \right| = \left| \int_\Theta u_L d\ell - \pi(x) \right|,$$

which implies that

$$\lim_{L\to\infty} \inf \pi_L(x) \geq \pi(x).$$

On the other hand, for any weakly convergent sequence $\{\hat{u}_L\}$ with $u_L \in \mathcal{V}(L)$, the weak limit $\hat{u}$ is in $\mathcal{U}$. Hence,

$$\lim_{L\to\infty} \sup \pi_L(x) \leq \pi(x).$$

If $\mathcal{U}$ is empty, $\mathcal{U}(L)$ must be empty and $\pi_L(x) = \pi_(x) = -\infty$ for large enough L; otherwise, $\mathbb{D}_s(\mathcal{U}(L), \mathcal{U}) = \infty$ for all $L > 0$ so that $\mathbb{D}_s(\mathcal{U}(L), \mathcal{U}) \to \infty$, which contradicts the conclusion of Lemma 4.8. □

Lemma 4.10 *Let y_L and $\mathcal{Z}_L$ be the optimal objective value and the set of optimal solutions of problem* (4.8)*, and y^* and $\mathcal{Z}^*$ be those of problem* (1.1)*. Suppose (i) $\mathcal{X}$ is a nonempty compact set, (ii) the function $\pi(\cdot)$ is continuous on $\mathcal{X}$, and (iii) $\pi_L(\cdot)$ uniformly converges to $\pi(\cdot)$ on $\mathcal{X}$ as $L \to \infty$. Then, $y_L \to y^*$ and $\mathbb{D}(\mathcal{Z}_L, \mathcal{Z}^*) := \max_{x_1 \in \mathcal{Z}_L} \min_{x_2 \in \mathcal{Z}^*} \|x_1 - x_2\| \to 0$ as $L \to \infty$.*

Proof: The proof of this lemma follows the steps of the proof of Theorem 5.3 in Shapiro et al. (2009). It is given in Appendix A for completeness because the two results are differently stated. □

Theorem 4.11 *Let y_L and $\mathcal{Z}_L$ be the optimal objective value and the set of optimal solutions of problem* (4.8)*, and y^* and $\mathcal{Z}^*$ be those of problem* (1.1)*. Suppose that $\mathcal{X}$ is a nonempty compact set and Assumptions (A1)-(A3) hold. Then $y_L \to y^*$ and $\mathbb{D}(\mathcal{Z}_L, \mathcal{Z}^*) \to 0$ as $L \to \infty$.*

Proof: By Lemma 4.1, we know that, for each $x \in \mathcal{X}$, $\pi(x)$ and $\pi_L(x)$ are continuous in $\mathcal{X}$. Since $\mathcal{X}$ is compact, there exist finitely many points, $x_1, \ldots, x_r \in \mathcal{X}$, and corresponding neighborhoods $\mathcal{N}_1, \ldots, \mathcal{N}_r$ covering $\mathcal{X}$ such that, for any $\epsilon > 0$ and $j = 1, \ldots, r$,

$$\sup_{x \in \mathcal{N}_j \cap \mathcal{X}} |\pi(x) - \pi(x_j)| < \epsilon/3,$$
$$\sup_{x \in \mathcal{N}_j \cap \mathcal{X}} |\pi_L(x) - \pi_L(x_j)| < \epsilon/3.$$

Furthermore, Lemma 4.9 shows that $\pi_L(x_j) \to \pi(x_j)$ as $L \to \infty$, i.e.,

$$|\pi_L(x_j) - \pi(x_j)| < \epsilon/3, \quad j = 1, \ldots, r,$$

for L large enough. Suppose without loss of generality that a given $x \in \mathcal{X}$ is covered by the neighborhood $\mathcal{N}_j$ for some j. Then

$$|\pi_L(x) - \pi(x)| \leq |\pi_L(x) - \pi_L(x_j)| + |\pi_L(x_j) - \pi(x_j)| + |\pi(x) - \pi(x_j)| \leq \epsilon.$$

Therefore,

$$\sup_{x \in \mathcal{X}} |\pi_L(x) - \pi(x)| \leq \epsilon,$$

which shows that π_L uniformly converges to π over $\mathcal{X}$. It follows by Lemma 4.10 that $y_L \to y^*$ and $\mathbb{D}(Z_L, Z^*) \to 0$ as $L \to \infty$. □

5 Case Study: Portfolio Optimization Problem

We now discuss an application of robust decision making over a utility set in John's investment decision problem expressed in Examples 2.2, 2.3, and 2.4. The index fund manager Mary will help John make a portfolio investment among the $N(= 8)$ assets which are widely used indexes: CBOE 10 Year Treasury Note Yield Index (TNX), CBOE 30-Year Treasury Bond Yield Index (TYX), CBOE Gold Index (GOX), Dow Jones Industrial Average Index (DJI), iShares MSCI EAFE Index (EFA), NASDAQ Composite Index (IXIC), S&P 500 Index (GSPC), and Wilshire 5000 Total Market Index (W5000). Table 1 lists $M(= 37)$ monthly returns of these assets in the latest three years.

Let $\xi = (\xi_1, \ldots, \xi_N)$ be a random vector of monthly returns, which has equally-likely realizations $\xi^k = (\xi_1^k, \ldots, \xi_N^k)$, $k = 1, \ldots, M$, i.e., the probabilities of these realizations are $p_k = 1/M$. John plans to invest \$1 million into the index fund market, and then his wealth will be $(1 + x^T\xi)$ in a

Table 1: Asset Monthly Returns (%)

Months	TNX	TYX	GOX	DJI	EFA	IXIC	GSPC	W5000
1/2012	1.60	5.88	10.21	3.63	5.98	8.11	4.67	5.32
12/2011	-9.66	-5.56	-13.72	1.43	-2.17	-0.58	0.85	0.68
11/2011	-4.61	-4.38	2.67	0.76	-2.18	-2.39	-0.51	-0.68
10/2011	13.02	9.59	6.64	9.54	9.64	11.14	10.77	11.39
9/2011	-13.51	-18.66	-12.42	-6.03	-10.81	-6.36	-7.18	-7.89
8/2011	-21.00	-13.08	7.94	-4.36	-8.76	-6.42	-5.68	-6.17
7/2011	-11.08	-5.71	2.85	-2.18	-2.39	-0.62	-2.15	-2.29
6/2011	3.61	3.79	-6.01	-1.24	-1.20	-2.18	-1.83	-1.85
5/2011	-7.58	-4.31	-4.80	-1.88	-2.21	-1.33	-1.35	-1.43
4/2011	-4.35	-2.22	2.55	3.98	5.63	3.32	2.85	2.80
3/2011	1.17	0.45	1.21	0.76	-2.40	-0.04	-0.10	0.21
2/2011	0.89	-1.75	8.18	2.81	3.56	3.04	3.20	3.34
1/2011	2.11	4.82	-10.45	2.72	2.09	1.78	2.26	1.93
12/2010	18.21	6.34	5.76	5.19	8.31	6.19	6.53	6.53
11/2010	7.28	2.50	0.80	-1.01	-4.82	-0.37	-0.23	0.42
10/2010	3.57	8.40	3.69	3.06	3.79	5.86	3.69	3.89
9/2010	1.61	4.53	7.00	7.72	9.99	12.04	8.76	9.27
8/2010	-14.78	-11.31	8.03	-4.31	-3.81	-6.24	-4.74	-4.82
7/2010	-1.36	1.79	-4.32	7.08	11.62	6.90	6.88	6.79
6/2010	-10.61	-7.13	2.23	-3.58	-1.87	-6.55	-5.39	-3.53
5/2010	-9.84	-7.06	-2.05	-7.92	-11.20	-8.29	-8.20	-10.08
4/2010	-4.44	-4.03	10.22	1.40	-2.81	2.64	1.48	2.09
3/2010	6.39	4.19	-0.28	5.15	6.40	7.14	5.88	6.17
2/2010	-0.28	0.44	8.14	2.56	0.26	4.23	2.85	3.24
1/2010	-5.99	-2.80	-11.38	-3.46	-5.07	-5.37	-3.70	-3.44
12/2009	20.00	10.74	-10.82	0.80	0.72	5.81	1.78	3.05
11/2009	-5.60	-1.18	19.04	6.51	3.92	4.86	5.74	5.20
10/2009	2.42	4.69	-6.69	0.00	-2.53	-3.64	-1.98	3.68
9/2009	-2.65	-3.11	15.43	2.27	3.80	5.64	3.57	-0.55
8/2009	-2.86	-3.02	-0.26	3.54	4.51	1.54	3.36	8.97
7/2009	-0.57	0.00	5.98	8.58	10.03	7.82	7.41	-1.91
6/2009	1.73	-0.69	-13.68	-0.63	-1.41	3.42	0.02	7.42
5/2009	10.90	7.43	34.50	4.07	13.21	3.32	5.31	9.17
4/2009	16.42	13.48	-11.19	7.35	11.50	12.35	9.39	15.87
3/2009	-11.84	-4.30	11.94	7.73	8.39	10.94	8.54	-14.74
2/2009	7.04	3.33	1.58	-11.72	-10.38	-6.68	-10.99	-10.90
1/2009	26.79	33.83	2.05	-8.84	-13.73	-6.38	-8.57	15.23

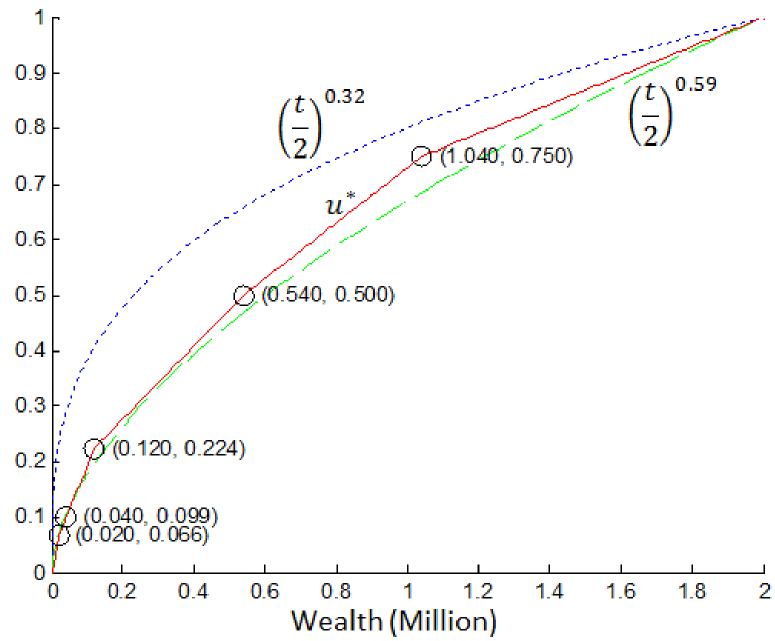

Figure 1: John's Wealth Utility Functions

month. Here $x \in \mathcal{X} := \{x \in \mathbb{R}^N_+ \mid x^T e \leq 1\}$ (e denotes the vector whose elements are all 1's) is John's portfolio choice. Example 2.2 provides boundary conditions (2.8) and (2.9) of John's risk preference based on CRRA utility function with a range of relative coefficients of risk aversion. In Examples 2.3 and 2.4, Mary further analyzes John's risk attitude using the certainty equivalent and preference comparison methods. Based on these analyses, Mary constructs the general conditions (2.11) and (2.14) on utility. In this portfolio management case, a risk averse utility set $\mathcal{U}$ defined by conditions (2.8), (2.11), and (2.14) is represented as

$$\mathcal{U} := \left\{ u \in \mathfrak{U}_2 \;\middle|\; \begin{array}{l} (t/2)^{0.59} \leq u(t) \leq (t/2)^{0.32}, \quad t \in [0,2], \\ u(0.16) \leq 0.25, \quad u(0.24) \geq 0.25, \\ u(0.46) \leq 0.50, \quad u(0.54) \geq 0.50, \\ u(0.54) \leq 0.75, \quad u(1.04) \geq 0.75, \\ 0.7u(0.2) + 0.3u(1.2) \geq 0.5u(1) + 0.5u(0), \\ 0.3u(0.8) + 0.7u(1.8) \geq 0.5u(1) + 0.5u(2). \end{array} \right\}.$$

To simply this study, we do not use the boundary condition (2.9) on marginal utility.

Our portfolio optimization model is described by

$$\max_{x \in \mathcal{X}} \min_{u \in \mathcal{U}} \sum_{k=1}^{M} p_k u(1 + x^T \xi^k). \tag{5.1}$$

To solve model (5.1), we approximate $\mathcal{U}$ by $\mathcal{U}(L)$ ($L = 100$) denoted as in (4.6) which is built on the partition equally dividing the interval $[0, 2]$ into L subintervals. Theorem 4.11 shows convergence of the optimal value and the set of optimal solution of the problem

$$\max_{x \in \mathcal{X}} \min_{u \in \mathcal{U}(L)} \sum_{k=1}^{M} p_k u(1 + x^T \xi^k) \tag{5.2}$$

to those of the true counterpart (5.1). When solving this approximation problem by the reformulation (3.7) given in Theorem 3.1, the optimal portfolio is

DJI:	0,	EFA:	0,	GOX:	0.207404,	GSPC:	0,
IXIC:	0.41178,	TNX:	0,	TYX:	0,	W5000:	0.380816.

Also, the optimal dual obtained from solving (3.7) gives us a saddle point to identify an optimal utility function, denoted as u^* in Figure 1. Observe that u^* is a piecewise linear increasing concave function between the lower bound function $(t/2)^{0.59}$ and the upper bound function $(t/2)^{0.32}$. The five break points of u^* are also labeled in Figure 1. This example suggests that the linear programming reformulation based approach is viable.

6 conclusions

Utility assessments typically result in ambiguity and inconsistency in the elicited utility function. To address this problem, we have developed a flexible decision making framework using a risk averse utility set described by boundary and general conditions on utility specifying a decision maker's risk attitude. This research is based on a novel probabilistic interpretation of utility, which regards utility and marginal utility functions as the c.d.f. and p.d.f of an underlying random utility index. We have also deduced a fundamental complementary relationship between random outcome and utility function from this interpretation. A tractable reformulation has been given for the discrete specification of a risk averse decision maker. We have also discussed the convergence of an approximation problem for the general continuous specification of a decision maker. A portfolio investment decision problem has illustrated a construction of the utility set using of classical non-parametric and parametric utility assessments. The computation results have demonstrated the usefulness of this decision making framework.

References

K. J. Arrow. *Aspects of the Theory of Risk Bearing.* Yrjö Jahnssonin Säätiö, 1965.

K. B. Athreya and S. N. Lahiri. *Measure theory and probability theory.* Springer, 2006.

G. Brunello. Absolute risk aversion and the returns to education. *Economics of Education Review*, 21(6):635–640, 2002.

U. Chajewska, D. Koller, and R. Parr. Making rational decisions using adaptive utility elicitation. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-00)*, pages 363–369, 2000.

P. H. Farquhar. Utility assessment methods. *Management Science*, 30(11):1283–1300, 1984.

D. G. Froberg and R. L. Kane. Methodology for measuring health-state preferences - ii: Scaling methods. *Journal of Clinical Epidemiology*, 42(5):459–471, 1989.

J. C. Hershey and P. J. H. Schoemaker. Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science*, 31(10):1213–1231, 1985.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *American Economic Review*, 92 (5):1644–1655, 2002.

J. Hu and S. Mehortra. Robust decision making using a general utility set. working paper, Dept. of IEMS, Northwestern University, 2012.

U. S. Karmarkar. Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance*, 21(1):61–72, 1978.

R. L. Keeney and H. Raiffa. *Decisions with multiple objectives: preferences and value tradeoffs.* John Wiley & Sons, New York, 1976.

F. Lhabitant. *Handbook of hedge funds.* John Wiley & Sons, 2006.

A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks.* John Wiley & Sons, Chichester, 2002.

E. Nord. Methods for quality adjustment of life years. *Social Science and Medicine*, 34(5):559–569, 1992.

J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32:122–136, 1964.

W. B. Riley Jr. and K. V. Chow. Asset allocation and individual risk aversion. *Financial Analysts Journal*, 48(6):32–37, 1992.

W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, 1976.

D. K. Schooley and D. D. Worden. Risk aversion measures: Comparing attitudes and asset allocation. *Financial Services Review*, 5(2):87–99, 1996.

A. Shapiro and S. Ahmed. On a class of minimax stochastic programs. *SIAM J. Optim.*, 14(4): 1237–1249, 2004.

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming : modeling and theory.* SIAM, 2009.

G. J. Stigler. The development of utility theory. i. *Journal of Political Economy*, 58(4):307–327, 1950a.

G. J. Stigler. The development of utility theory. ii. *Journal of Political Economy*, 58(5):373–396, 1950b.

G. J. Stigler. The adoption of the marginal utility theory. *History of Political Economy*, 4(2): 571–586, 1972.

G. G. Szpiro. Measuring risk aversion: An alternative approach. *The Review of Economics and Statistics*, 68(1):156–159, 1986.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, NJ, 2nd. edition, 1947.

P. Wakker and D. Deneffe. Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42(8):1131–1150, 1996.

A The proof of Lemma 4.10

The uniform convergence of $\pi_L(x)$ to $\pi(x)$ means that, for any $\epsilon > 0$, there exists $\hat{L}$ such that, for all $L \geq \hat{L}$,

$$\sup_{x \in \mathcal{X}} \|\pi_L(x) - \pi_(x)\| \leq \epsilon,$$

which implies that $|y_L - y^*| \leq \epsilon$. It follows that $y_L \to y^*$ as $L \to \infty$.

Arguing with a contradiction, we suppose that $\mathbb{D}(\mathcal{Z}_L, \mathcal{Z}) \nrightarrow 0$. There is $x_L \in \mathcal{Z}_L$ such that $d(x_L, \mathcal{Z}) := \min_{z \in \mathcal{Z}} \|x_L - z\| \geq \delta$ for some $\delta > 0$. For the compactness of $\mathcal{X}$, we can assume that x_L converges to a point $x^* \in \mathcal{X}$. It follows $x^* \notin \mathcal{Z}$ and hence $\pi(x^*) < y^*$. Moreover, $y_L = \pi_L(x_L)$ and

$$|\pi_L(x_L) - \pi(x^*)| \leq |\pi_L(x_L) - \pi(x_L)| + |\pi(x_L) - \pi(x^*)|,$$

whose first item $|\pi_L(x_L) - \pi(x_L)| \to 0$ by the uniform convergence of π_L to π under Assumption (iii) and second item $|\pi(x_L) - \pi(x^*)| \to 0$ by the continuity of π under Assumption (ii). That is, we have a contradiction that $y_L > y^*$.