Decision Making Based on a Nonparametric Shape-Preserving Perturbation of a Reference Utility Function

Jian Hu

Department of Industrial and Manufacturing System Engineering

University of Michigan - Dearborn

jianhu@umich.edu

November 25, 2013

Abstract

This paper develops a robust optimization based decision-making framework using a nonparametric perturbation of a reference utility function. The perturbation preserves the risk-aversion property but solves the problem of ambiguity and inconsistency in eliciting the reference utility function. We study the topology of the perturbation, and show that in the decision-making framework the price of perturbation is increasing and concave. When the reference utility is given at discrete points, we reformulate this optimization problem as a second-order cone program. The Monte Carlo sampling method is used to solve the general case that a reference utility is a continuous function, and the convergence of this method is proved. The usefulness of the robust utility optimization framework is illustrated with the help of a portfolio investment decision problem.

Key Words: Expected Utility Maximization, Robust Optimization, Nonparametric Perturbation, Sensitivity Analysis, Portfolio Optimization

1 Introduction

Expected utility maximization is one of classical risk-averse optimization techniques, in which utility function characterizes the decision maker's attitude toward the risk arising from uncertainty. Complete and comprehensive investigations in economics and decision theory 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, it is still rather obscure to accurately formulate the risk attitude due to cognitive difficulty and incomplete information (Karmarkar (1978) and Weber (1987)). The standard gamble methods and paired gamble methods, such as preference comparison, probability equivalence, value equivalence, and certainty equivalence, are classical nonparametric utility assessments (see e.g. Farquhar (1984), Wakker and Deneffe (1996) and references therein). The questionnaire in these methods are difficult for untrained clients to answer. A limited number of survey questions may be insufficient to completely quantify a real decision problem which often has a very large decision space (Chajewska et al. (2000)). The decision maker often gives inconsistent evaluations resulting in nonconcavity of an elicited utility function (Jacobson and Petrie (2007)). Moreover, multiple methods yield differing utility constructions (Hershey and Schoemaker (1985), Fromberg and Kane (1989), and Nord (1992)). In parametric assessments, constant absolute risk aversion (CARA), relative risk aversion (CRRA), and hyperbolic absolute risk aversion (HARA) utility functions are often used as underlying functions. The coefficients of risk aversion in these functions have been studied experimentally (see e.g. Szpiro (1986), Riley Jr. and Chow (1992), Schooley and Worden (1996), Holt and Laury (2002), and Brunello (2002)). In comparison with nonparametric methods, parametric approaches facilitate the survey and avoid the issue of inconsistency; however, truly representing the decision maker's risk preference very relies on the appropriate choice of an underlying utility function.

A solution to the ambiguity in utility assessments is to specify an uncertainty set of utility functions implicitly preserving the decision maker's preference (Weber (1987)). The concept of stochastic dominance serves a good example (see e.g., Müller and Stoyan (2002) and references therein). For an instance, 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 sets of utility functions. This core idea also leads to the studies of a minimax regret optimization problem (Boutilier et al. (2006)) and robust optimization frameworks (Hu and Mehrotra (2012a,b)). Particularly, using knowledge learned from classical utility assessments, Hu and Mehrotra (2012a,b) described shape constraints based on the functional bounds on utility and marginal utility functions, and certain additional rather general auxiliary conditions.

We present a robust decision making framework, based on a nonparametric shape-preserving perturbation region of a given reference utility function u_r, as follows:

$$\max_{x \in \mathcal{X}} f(x), \tag{1}$$

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

$$f(x) := \min_{u \in \mathcal{U}} \left\{ F(x, u) := \mathbb{E}[u(\xi(x))] \right\}. \tag{2}$$

Here, $\mathcal{X}$ is a decision feasible region, $\mathcal{U}$ represents a perturbation region of u_r, and $\xi(x)$ represents a random profit. In the paper, we assume that, for all $x \in \mathcal{X}$, $\xi(x)$ has a bounded support in $\Theta := [0, \theta]$, and follows a discrete distribution as

$$\Pr \left\{ \xi(x) = \xi_k(x) \right\} = p_k, \quad k = 1, \dots, K. \tag{3}$$

1

In general, the Sample Average Approximation (SAA) method can be used to solve the case that $\xi(\cdot)$ has a continuous distribution. Under some mild assumptions, model (1) is ensured to be a convex programming problem. The convergence theory has been well studied in the literature (see e.g. Birge and Louveaux (1997), Shapiro et al. (2009), and references therein).

Different from the approaches in Hu and Mehrotra (2012a,b) which construct a utility set by specifying general properties, we allow the decision maker to recommend a reference utility function, and use its perturbation region against the ambiguity of eliciting this reference. Furthermore, we preserve the risk-aversion property in the perturbation region. Another advantage of the framework (1) provides a nonparametric approach to the sensitivity analysis of a given utility function in the expected utility maximization problem. This nonparametric perturbation can be used to test the sensitivity of the fixed mathematical form of an underlying utility function given by parametric assessment methods.

This paper is organized as follows. Section 2 constructs a perturbation region $\mathcal{U}$ of the reference utility function u_r specified by a perturbation tolerance condition and a perturbation symmetry condition. We show the compactness and convexity of $\mathcal{U}$, and discuss the price of perturbation which is defined using the mapping relationship between the optimal value of problem (1) and the size of $\mathcal{U}$. In Section 3, we study the case that u_r is given at discrete points. The framework (1) in this case is reformulated as a second-order cone program. In Section 4, we use the Monte Care sampling method to address the general case that u_r is a continuous function. The convergence in this approximation approach is proved. In Section 5, the usefulness of the framework (1) is illustrated by a portfolio investment decision problem. Numerical results are used to verify the impact of the perturbation on the optimal value and solutions of the framework (1). The test of convergence property and computational performance of the Monte Carlo sampling approach is also given.

2   Perturbation region

We first state the following notions. Denote by $\mathfrak{L}^\infty$ the metric space of continuous functions defined on Θ based on the infinite norm written as $\|\cdot\|_\infty$. Let $(\Theta, \mathfrak{G}, \mu)$ be a complete measure space where $\mathfrak{G}$ is the smallest σ-algebra of Θ and μ is a positive signed measure on Θ with $\mu(\Theta) = 1$. Denote by $\mathcal{L}^2(\mu)$ the space of all measurable functions on Θ for which $\|u\|_\mu := \left(\int_\Theta u(t)^2 \mu(dt)\right)^{1/2} < \infty$. It is well known that $\mathcal{L}^2(\mu)$ is a Hilbert space. $\mathfrak{U}$ represents a set of all increasing concave utility functions well defined on Θ satisfying boundary conditions:

$$u(0) = 0, \qquad u(\theta) = 1, \tag{4}$$

and also,

$$u(t) \leq u_{\sup}(t) := \frac{t^{1-\rho}}{1-\rho}, \qquad \text{for all } t \in \Theta. \tag{5}$$

Here, $\rho > 0$ is less than but very close to 1. Condition (4) is a commonly used normalization of utility, and the preference ranking of any two alternatives is not changed in this normalization (Keeney and Raiffa (1976)). Condition (5) ensures the continuity of all utility functions in $\mathfrak{U}$, ruling out lower semicontinuous concave functions with a jump at 0. The upper bound $u_{\sup}(t)$ is very generally defined as a CRRA (power) utility function without normalization. This definition permits the right derivative of the utility functions in $\mathfrak{U}$ at 0 to be infinite so that normalized CRRA utility functions $\left(\frac{t}{\theta}\right)^{1-\gamma}$ with $\gamma \leq \rho$ are still kept in consideration. Note that, in addition to the

upper bound u_{sup}, condition (4) implies that the linear function

$$u_{\text{inf}}(t) := t/\theta \tag{6}$$

is the lower bound of the utility functions in $\mathfrak{U}$. Also, condition (5) ensures the equicontinuity of the utility functions in $\mathfrak{U}$, and hence, $\mathfrak{U}$ is a compact set.

Proposition 2.1 *All $u \in \mathfrak{U}$ are equicontinuous on Θ.*

Proof: For a given $\gamma > 0$, let

$$\delta := ((1 - \rho)\gamma)^{\frac{1}{1-\rho}}.$$

Because of the properties of increasing and concavity of $u \in \mathfrak{U}$, we have that, for all $t \in \Theta$,

$$u(t + \delta) - u(t) \leq u(\delta) - u(0) \leq u_{\text{sup}}(\delta) - u_{\text{sup}}(0) = \gamma.$$

$\square$

We now show the set $\mathfrak{U}$ is compact on the metric spaces $\mathfrak{L}^\infty$ and $\mathcal{L}^2(\mu)$ in the following theorem.

Theorem 2.2 *$\mathfrak{U}$ is a convex and compact set on both the spaces $\mathfrak{L}^\infty$ and $\mathcal{L}^2(\mu)$.*

Proof: The proof of the convexity is trivial. To show the compactness of $\mathfrak{U}$ on the metric space $\mathfrak{L}^\infty$, we need to verify that, for every infinite sequence in $\mathfrak{U}$, there is a convergent subsequence in $\mathfrak{L}^\infty$. An analogy is used to prove the compactness on $\mathcal{L}^2(\mu)$.

Let $\{u_n \in \mathfrak{U}\}_{n \geq 1}$ be a sequence of utility functions. By definition, the sequence $\{u_n\}$ is uniformly bounded by 1. By Proposition 2.1, we know the equicontinuity of $\{u_n\}$. Hence, it follows from the Arzelà-Ascoli theorem (see e.g. Dunford and Schwartz (1958)) that there exists a subsequence $\{u_{n_k}\}$ that converges uniformly to a function denoted by $\tilde{u}$. It is trivial to show that $\tilde{u}$ is increasing and concave, and satisfies conditions (4) and (5). Hence, $\tilde{u} \in \mathfrak{U}$, and $\mathfrak{U}$ is a compact set on $\mathfrak{L}^\infty$.

We next check the convergence of the subsequence $\{u_{n_k}\}$ on $\mathcal{L}^2(\mu)$. Since $\|u_{n_k} - \tilde{u}\|_\infty \leq 1$, by the Lebesgue dominated convergence theorem (see e.g. Royden (1988)) we have

$$\lim_{k \to \infty} \|u_{n_k} - \tilde{u}\|_\mu = \left\| \lim_{k \to \infty} u_{n_k} - \tilde{u} \right\|_\mu = 0.$$

$\square$

We now specify a perturbation region of a given reference utility function u_r. u_r is assumed to satisfy the boundary conditions (4) and (5), but is unnecessary to be an increasing concave function for the inconsistency in utility assessments. Using a perturbation tolerance condition

$$g_1(u, \mu) := \|u - u_r\|_\mu \leq \epsilon_1, \tag{7}$$

and a perturbation symmetry condition

$$g_2(u, \mu) := \left| \int_\Theta u(t) - u_r(t) \; \mu(dt) \right| \leq \epsilon_2, \tag{8}$$

we define a perturbation region of u_r as

$$\mathcal{U}(\epsilon) := \{ u \in \mathfrak{U} : u \text{ satisifes conditions (7) and (8) } \}. \tag{9}$$

Here, $\epsilon = (\epsilon_1, \epsilon_2) \in \mathbb{R}_+^2$ is a given constant vector where ϵ_1 is a perturbation tolerance and ϵ_2 gives an asymmetrical level of the perturbation. The tolerance condition (7) describes a neighborhood of u_r defined on $\mathcal{L}^2(\mu)$. The symmetry condition (8) requests that a perturbation should be around u_r at both up and down sides. For $u_r \in \mathfrak{U}$, without condition (8), an optimal utility solution of the inner minimization problem (2) is always below u_r for every $x \in \mathcal{X}$. The following proposition shows the convexity of functions $g_1^2(u, \mu)$ and $g_2(u, \mu)$. The proof is trivial, and hence, omitted.

Proposition 2.3 *For a given measure μ, $g_1^2(u,\mu)$ and $g_2(u,\mu)$ are convex in u.*

In conditions (7) and (8) μ is used to measure the relative importance of different subintervals of Θ. A particular case is that μ is a discrete measure, based on which condition (7) generalizes the least squares model fitting criterion. Actually, nonparametric utility assessments can only generate finitely many utility points corresponding to the decision maker's answers to a questionnaire, and use a piecewise linear curve to link the all points. Using this piecewise linear utility function as the reference u_r, what we really need to consider is a perturbation region defined using the discrete utility points directly given by the decision maker. Let $(t_i, u_r(t_i)) \in \Theta \times [0,1]$, $i = 1, \ldots, I$, be the utility points. In this case, the measure μ should be assigned on t_i, i.e.,

$$\mu(t_i) > 0, \ i = 1, \ldots, I, \ \text{and} \ \mu(\Theta \setminus \{t_1, \ldots, t_I\}) = 0. \tag{10}$$

Conditions (7) and (8) are represented as

$$g_1(u,\mu) = \left(\sum_{i=1}^{I} \mu(t_i)(u(t_i) - u_r(t_i))^2 \right)^{1/2} \leq \epsilon_1 \tag{11}$$

and

$$g_2(u,\mu) = \left| \sum_{i=1}^{I} \mu(t_i)(u(t_i) - u_r(t_i)) \right| \leq \epsilon_2. \tag{12}$$

Condition (11) is constructed using the weighted least squares model fitting criterion, which requires utility functions to well fit the points $(t_i, u_r(t_i))$ with the allowable tolerance ϵ_1. In what follows, we give some properties of the utility set $\mathcal{U}(\epsilon)$.

Theorem 2.4 *The utility set $\mathcal{U}(\epsilon)$ given in (9) is convex and compact on the space $\mathcal{L}^2(\mu)$.*

Proof: We first show convexity of the set $\mathcal{U}(\epsilon)$. Let $u_1, u_2 \in \mathcal{U}(\epsilon)$, and $u := \lambda u_1 + (1-\lambda)u_2$ for a given $\lambda \in [0,1]$. By Proposition 2.3, we have

$$g_1^2(u,\mu) \leq \lambda g_1^2(u_1,\mu) + (1-\lambda)g_1^2(u_2,\mu) = \epsilon_1^2,$$

and also $g_2(u,\mu) \leq \epsilon_2$. Moreover, by Theorem 2.4, we have $u \in \mathfrak{U}$. Hence, it follows that $u \in \mathcal{U}(\epsilon)$, and $\mathcal{U}(\epsilon)$ is a convex set.

We next prove the compactness of $\mathcal{U}(\epsilon)$. Let us first check the closeness of the set

$$\widehat{\mathcal{U}} := \left\{ u \in \mathcal{L}^2(\mu) : \text{u satsifies condition (8).} \right\}$$

Now suppose that $\widehat{\mathcal{U}}$ is not closed. There exists a limit point $\hat{u} \notin \widehat{\mathcal{U}}$ such that, in any open neighborhood of $\hat{u}$, there are points belonging to $\widehat{\mathcal{U}}$. Since $\hat{u} \notin \widehat{\mathcal{U}}$, there is a $\tau > 0$ such that

$$\left| \int_{\Theta} \hat{u}(t) - u_r(t) \ \mu(dt) \right| > \epsilon_2 + \tau. \tag{13}$$

We write an open ball of $\hat{u}$ on $\mathcal{L}^2(\mu)$ for the given τ as

$$\mathcal{W}_\tau(\hat{u}) := \left\{ u \in \mathcal{L}^2(\mu) : \|u - \hat{u}\|_\mu < \tau \right\}.$$

4

Let $u \in \widehat{\mathcal{U}} \cap \mathcal{W}_\tau(\hat{u})$. It follows that

$$\left| \int_\Theta \hat{u}(t) - u_r(t) \ \mu(dt) \right| \leq \left| \int_\Theta \hat{u}(t) - u(t) \ \mu(dt) \right| + \left| \int_\Theta u(t) - u_r(t) \ \mu(dt) \right|$$

$$\leq \int_\Theta |\hat{u}(t) - u(t)| \ \mu(dt) + \epsilon_2$$

$$\leq \left(\int_\Theta |\hat{u}(t) - u(t)|^2 \ \mu(dt) \right)^{1/2} + \epsilon_2$$

$$< \tau + \epsilon_2,$$

which contracts the inequality (13). Hence, $\widehat{\mathcal{U}}$ is a closed subset on $\mathcal{L}^2(\mu)$. It implies that $\mathcal{U}(\epsilon)$ is closed since $\mathcal{U}(\epsilon)$ is the intersection of the closed set $\widehat{\mathcal{U}}$, the closed ball $\{u \in \mathcal{L}^2(\mu) : \|u - u_r\|_\mu \leq \epsilon_1\}$, and the compact set $\mathfrak{U}$ shown in Theorem 2.2. Therefore, $\mathcal{U}(\epsilon)$ is compact since $\mathcal{U}(\epsilon)$ is a closed subset of the compact set $\mathfrak{U}$. $\qquad \square$

We now discuss the price of perturbation robustness of problem (1). This issue is mentioned by Bertsimas et al. (2011) for trading off the conservativeness of robust optimization and the efficiency of decision policies. For testing the impact of enlarging the perturbation, we adjust tolerance constant ϵ_1 and denote the price function of perturbation as

$$\varphi(\epsilon_1) = -\max_{x \in \mathcal{X}} \min_{u \in \mathcal{U}(\epsilon_1, \epsilon_2)} F(x, u) \tag{14}$$

for a fixed asymmetric level ϵ_2. We assume that, in the discussion, ϵ_1 and ϵ_2 satisfy the following assumption.

(A1). There exists $\gamma > 0$ such that the set $\mathcal{U}(\epsilon - \gamma e)$ is nonempty (e denotes the vector whose elements are all 1's).

Assumption (A1) ensures that conditions (7) and (8) are strictly satisfied, i.e., there exists $u \in \mathcal{U}(\epsilon)$ such that $g_1(u, \mu) < \epsilon_1$ and $g_2(u, \mu) < \epsilon_2$. Under Assumption (A1), we show that the price of perturbation is an increasing concave function of ϵ_1.

Theorem 2.5 *The price function $\varphi(\epsilon_1)$ is increasing and concave at ϵ_1 with which, for a fixed ϵ_2, $\mathcal{U}(\epsilon_1, \epsilon_2)$ satisfies Assumption (A1) ,*

Proof: Let

$$\mathcal{K} := \{(u, \epsilon) : \|u\|_\mu \leq \epsilon\}$$

denote the second-order cone on $\mathcal{L}_2(\mu)$. Using the fact that the second-order cone is self-dual, we write the dual of $\mathcal{K}$ as

$$\mathcal{K}^* := \left\{ (v, \eta) : \int_\Theta v(t) u(t) \mu(dt) + \eta \epsilon \geq 0 \text{ for all } (u, \epsilon) \in \mathcal{K} \right\} = \{(v, \eta) : \|v\|_\mu \leq \eta\}.$$

Since $(u - u_r, \epsilon_1) \in \mathcal{K}$, we can rewrite condition (7) in the form of conic inequality as

$$-(u - u_r, \epsilon_1) \preceq_\mathcal{K} 0.$$

The Lagrangian of problem (2) is thus constructed as

$$L(u, v, \eta; x, \epsilon_1) = F(x, u) - \int_\Theta v(t)(u(t) - u_r(t))\mu(dt) - \eta\epsilon_1 + \alpha(g_2(u, \mu) - \epsilon_2)$$

5

for $(v, \eta) \in \mathcal{K}^*$ and $\alpha \geq 0$. Note that $F(x, u)$ and $g_2(u, \mu)$ are linear and convex functionals of u, respectively. Also by the assumption, there exists a strictly feasible point satisfying conditions (7) and (8). Hence, the strong duality holds between problem (2) and its Lagrangian dual problem

$$\max_{\substack{(v, \eta) \in \mathcal{K}^* \\ \alpha \geq 0}} \min_{u \in \mathfrak{U}} L(u, v, \eta; x, \epsilon_1).$$

We thus have

$$c(\epsilon_1) = - \max_{\substack{(v, \eta) \in \mathcal{K}^* \\ \alpha \geq 0 \\ x \in \mathcal{X}}} \min_{u \in \mathfrak{U}} L(u, v, \eta; x, \epsilon_1)$$

$$= - \max_{\substack{(v, \eta) \in \mathcal{K}^* \\ \alpha \geq 0 \\ x \in \mathcal{X}}} -\eta\epsilon_1 + \left\{ \min_{u \in \mathfrak{U}} F(x, u) - \int_{\Theta} v(t)(u(t) - u_r(t))\mu(dt) + \alpha(g_2(u, \mu) - \epsilon_2) \right\}.$$

It shows that $c(\epsilon_1)$ is an increasing concave function of ϵ_1. $\qquad\square$

3 Model with a discrete measure μ

We now discuss problem (1) with the perturbation region $\mathcal{U}(\epsilon)$ given in (9). Let μ be the discrete measure denoted as in (10). Let $0 = t_0 \leq t_1 \leq \cdots \leq t_{I+1} = \theta$ be the ordered sequence. The following theorem reformulates problem (1) as a second-order cone program.

Theorem 3.1 *If Assumption (A1) holds, then problem (1) is equivalent to*

$$\max_{x, \phi, \psi, \pi, \tau, \kappa} \alpha \left(\sum_{i=1}^{I} \mu(t_i)u_r(t_i) + \epsilon_2 \right) + \beta \left(\sum_{i=1}^{I} \mu(t_i)u_r(t_i) - \epsilon_2 \right) - \pi_{I+1} - \tau_{I+1} + \sum_{i=1}^{I} \kappa_i u_{\sup}(t_i)$$

$$+ \sum_{k=1}^{K} \phi_{I+1,k} + \sum_{i=1}^{I} \psi_i u_r(t_i)\sqrt{\mu(t_i)} - \epsilon_1 \eta \tag{15a}$$

$$s.t. \ -\eta^2 + \left(\sum_{i=1}^{I} \psi_i^2 \right) \leq 0 \tag{15b}$$

$$\sum_{i=1}^{I+1} (t_i - t_0)\phi_{ik} \leq p_k(\xi_k(x) - t_0), \quad k = 1, \ldots, K, \tag{15c}$$

$$\sum_{i=1}^{I+1} \phi_{ik} \leq p_k, \quad k = 1, \ldots, K, \tag{15d}$$

$$\pi_i(t_i - t_{i-1}) + \tau_{i+1}(t_{i+1} - t_i) \geq 0, \quad i = 1, \ldots, I+1, \tag{15e}$$

$$\alpha\mu(t_i) + \beta\mu(t_i) - \sum_{k=1}^{K} \phi_{ik} + \psi_i\sqrt{\mu(t_i)} + \pi_i - \pi_{i+1} + \tau_i - \tau_{i+1} + \kappa_i \leq 0, \quad i = 1, \ldots, I,$$

$$\tag{15f}$$

$$\phi_{ik} \geq 0, \quad i = 1, \ldots, I+1, \ k = 1, \ldots, K, \tag{15g}$$

$$\pi_i \geq 0, \quad i = 1, \ldots, I+1, \tag{15h}$$

$$\tau_1 = \tau_{I+2} = 0, \quad \tau_i \leq 0, \quad i = 2, \ldots, I+1, \tag{15i}$$

$$\kappa_i \leq 0, \quad i = 1, \ldots, I, \tag{15j}$$

$$\alpha \leq 0, \ \beta \geq 0, \ \eta \geq 0, \ x \in \mathcal{X}. \tag{15k}$$

Let $z_0 = 0$, $z_{I+1} = 1$, and z_i, $i = 1, \ldots, I$, be the optimal dual solutions with respect to constraint (15f). Then

$$u^*(t) := \sum_{i=1}^{I} \left(\frac{z_{i+1} - z_i}{t_{i+1} - t_i} t + \frac{t_{i+1} z_i - t_i z_{i+1}}{t_{i+1} - t_i} \right) \mathbf{1} \{ t_i \leq t < t_{i+1} \} \tag{16}$$

is an optimal utility of problem (1). Here, $\mathbf{1} \{\cdot\}$ is an indicator function.

Proof: For a given $x \in \mathcal{X}$, we first prove that the inner minimization problem (2) can be equivalently represented by a second order cone program

$$\min_{s,z,r,h} \sum_{k=1}^{K} p_k(s_k(\xi_k(x) - t_0) + r_k) \tag{17a}$$

$$\text{s.t. } (t_i - t_0)s_k + r_k - z_i \geq 0, \quad i = 1, \ldots, I+1, \ k = 1, \ldots, K, \tag{17b}$$

$$\sum_{i=1}^{I} \mu(t_i)(z_i - u_r(t_i))^2 \leq \epsilon_1^2, \tag{17c}$$

$$\sum_{i=1}^{I} \mu(t_i)(z_i - u_r(t_i)) \leq \epsilon_2, \tag{17d}$$

$$\sum_{i=1}^{I} \mu(t_i)(z_i - u_r(t_i)) \geq -\epsilon_2, \tag{17e}$$

$$z_i - z_{i-1} - h_i(t_i - t_{i-1}) \geq 0, \quad i = 1, \ldots, I+1, \tag{17f}$$

$$z_i - z_{i-1} - h_{i-1}(t_i - t_{i-1}) \leq 0, \quad i = 2, \ldots, I+1, \tag{17g}$$

$$z_0 = 0, \ z_{I+1} = 1, \ 0 \leq z_i \leq u_{\sup}(t_i), \quad i = 1, \ldots, I, \tag{17h}$$

$$s_k \geq 0, \ r_k \geq 0, \quad k = 1, \ldots, K \tag{17i}$$

$$h_i \geq 0, \quad i = 1, \ldots, I. \tag{17j}$$

We denote by $\widetilde{\mathcal{U}}$ the subset of $\mathcal{U}(\epsilon)$ consisting of all piecewise linear increasing concave functions with break points t_i for $i = 0, \ldots, I+1$. Let $\tilde{u}$ be an optimal solution of problem (2). We construct a piecewise linear increasing concave function

$$\tilde{u}_I := \sum_{i=0}^{I} \left(\frac{\tilde{u}(t_{i+1}) - \tilde{u}(t_i)}{t_{i+1} - t_i} t + \frac{t_{i+1}\tilde{u}(t_i) - t_i\tilde{u}(t_{i+1})}{t_{i+1} - t_i} \right) \mathbf{1} \{ t_i \leq t \leq t_{i+1} \}.$$

$\tilde{u}$ and $\tilde{u}_I$ have the same values at all t_i, and at other points, the values of $\tilde{u}_I$ is not greater than those of $\tilde{u}$. It is straightforward to see that $\tilde{u}_I \in \mathfrak{U}$ and $\tilde{u}_I$ satisfies conditions (11) and (12). Hence, $\tilde{u}_I \in \mathcal{U}(\epsilon)$ and $\tilde{u}_I$ is also a minimizer of problem (2). It implies that problem (2) is equivalent to

$$\min_{u \in \widetilde{\mathcal{U}}} F(x, u). \tag{18}$$

We now describe the set $\widetilde{\mathcal{U}}$ and show the equivalence of problems (17) and (18). Conditions (17c) - (17e) are the copy of conditions (11) and (12). Conditions (17f) and (17g) imply

$$\frac{z_1 - z_0}{t_1 - t_0} \geq h_1 \geq \cdots \geq h_I \geq \frac{z_{I+1} - z_I}{t_{I+1} - t_I} \geq h_{I+1},$$

which ensure that valid utility functions are increasing and concave when all $h_i \geq 0$. For a piecewise linear increasing concave function u with break points t_i and values z_i, $i = 0, \ldots, I+1$, it can be written as the minimization of all subgradients at points (t_i, z_i)

$$
\begin{aligned}
u(v) = \min_{s,r} \ & (v - t_0)s + r \\
\text{s.t. } & (t_i - t_0)s + r - z_i \geq 0, \quad i = 1, \ldots, I+1, \\
& s \geq 0, \ r \geq 0.
\end{aligned}
$$

This gives the objective (17a) and conditions (17b) and (17i).

Using Assumption (A1), we now prove the strong duality of problem (17) and its dual problem. We know that there exists $\bar{u} \in \mathcal{U}(\epsilon - \gamma e)$. Let us first assume that $\bar{u}$ is not $u_{\inf}$ which is mentioned to be the linear function t/θ. It implies that $\bar{u}(t) > u_{\inf}(t)$ for all $t \in \mathrm{int}\Theta$. Hence, there exists $\hat{u} \in \mathfrak{U}$ such that $\bar{u}(t) > \hat{u}(t) > u_{\inf}(t)$ for all for all $t \in \mathrm{int}\Theta$ and $\hat{u}$ is strictly concave at t_i, for $i = 1, \ldots, I$, i.e.,

$$
\frac{\hat{u}(t_i) - \hat{u}(t_{i-1})}{t_i - t_{i-1}} > \frac{\hat{u}(t_{i+1}) - \hat{u}(t_i)}{t_{i+1} - t_i}.
$$

Using $\lambda \in (0, \gamma)$, we construct

$$
u_\lambda = (1 - \lambda)\bar{u} + \lambda\hat{u}.
$$

By Theorem 2.4, $u_\lambda \in \mathfrak{U}$. It is easy to see that $u_\lambda(t) < u_{\sup}(t)$ for all $t \in \mathrm{int}\Theta$. Also, u_λ is strictly concave at t_i, since

$$
\begin{aligned}
\frac{u_\lambda(t_i) - u_\lambda(t_{i-1})}{t_i - t_{i-1}} &= (1 - \lambda)\frac{\bar{u}(t_i) - \bar{u}(t_{i-1})}{t_i - t_{i-1}} + \lambda\frac{\hat{u}(t_i) - \hat{u}(t_{i-1})}{t_i - t_{i-1}} \\
&> (1 - \lambda)\frac{\bar{u}(t_{i+1}) - \bar{u}(t_i)}{t_{i+1} - t_i} + \lambda\frac{\hat{u}(t_{i+1}) - \hat{u}(t_i)}{t_{i+1} - t_i} \\
&= \frac{u(t_{i+1}) - u(t_i)}{t_{i+1} - t_i}.
\end{aligned}
$$

We next claim that u_λ strictly satisfies conditions (7) and (8), since

$$
g_1(u_\lambda, \mu) \leq \|u_\lambda - \bar{u}\|_\mu + \|\bar{u} - u_r\|_\mu \leq \lambda\|\hat{u} - \bar{u}\|_\mu + \epsilon_1 - \gamma \leq \epsilon_1 + (\lambda - \gamma) < \epsilon_1,
$$

and

$$
\begin{aligned}
g_2(u_\lambda, \mu) &\leq \left| \int_\Theta u_\lambda(t) - \bar{u}(t)\mu(dt) \right| + \left| \int_\Theta \bar{u}(t) - u_r(t)\mu(dt) \right| \\
&\leq \lambda \int_\Theta |\hat{u}(t) - \bar{u}(t)|\mu(dt) + \epsilon_2 - \gamma \\
&< \epsilon_2.
\end{aligned}
$$

We have proven that $u_\lambda \in \mathcal{U}(\epsilon)$ and $z_{\lambda,i} = u_\lambda(t_i)$ strictly satisfies conditions (17c) - (17h). Also, condition (17b) can be strictly satisfied automatically, since s_k and r_k, $k = 1, \ldots, K$, are unrestricted above. It implies that problem (17) is strictly feasible such that the strong duality holds. There is no duality gap between problem (17) and its dual problem which is problem (15) for the given x where ϕ, ψ, η, α, β, π, τ, and κ are the dual variables of conditions (17b) - (17h), respectively.

For the case that $\bar{u} = u_{\inf}$, we can construct a strictly concave function $\hat{u}$ such that $u_{\inf} < \hat{u} < u_{\sup}$. A similar proof shows that the strong duality of problem (17) and its dual problem holds. Finally, problem (15) maximizes the dual maximization problem for all $x \in \mathcal{X}$. $\qquad \square$

Remark 3.2

1. *Problem (1) can be studied with respect to a general region $\Theta := [\theta_1, \theta_2]$. In this case, let $t_0 = \theta_1$ and $t_{I+1} = \theta_2$ in the reformulation (15).*

2. *If ρ is chosen to satisfy $u_{\sup}(t_1) = \frac{t_1^{1-\rho}}{1-\rho} > 1$, then the boundary constraint (5) is redundant. Correspondingly, let $\kappa_i = 0$ for all $i = 1, \ldots, I$, in the reformulation (15).*

4   Model with a continuous measure μ

We now address the general case that μ is a continuous measure. Let ζ_i, $i = 1, \ldots, N$, be the i.i.d Monte Carlo samples under measure μ. Using these samples, we construct a sample measure μ_N with $\mu_N(\zeta_i) = 1/N$, for $i = 1, \ldots, N$, and $\mu_N(\Theta \setminus \{\zeta_1, \ldots, \zeta_N\}) = 0$, and denote a utility set as

$$\mathcal{U}_N(\epsilon) := \{u \in \mathfrak{U} : g_1(u, \mu_N) \le \epsilon_1, \ g_2(u, \mu_N) \le \epsilon_2\}. \tag{19}$$

An approximation approach of problem (1) based on the Monte Carlo sampling method is specified as

$$\max_{x \in \mathcal{X}} f_N(x) \tag{20}$$

where

$$f_N(x) := \min_{u \in \mathcal{U}_N(\epsilon)} F(x, u) \tag{21}$$

is the approximation of the inner minimization problem (2). Problem (20) can be solved using the reformulation of the second-order cone programming given in Theorem 3.1. We now study the convergence of this Monte Carlo sampling approach. In addition to Assumption (A1), the following mild assumptions are needed for our convergence analysis:

(A2). $\mathcal{X}$ is a nonempty compact set.

(A3). $\xi(x)$ is concave in an open neighborhood of $\mathcal{X}$ w.p.1.

Under these assumptions, we now consider the relationship between the set $\mathcal{U}_N(\epsilon)$ and its counter part $\mathcal{U}(\epsilon)$. Let us denote the distance of the utility function u and the utility set $\mathcal{U}$ as

$$d(u, \ \mathcal{U}) := \begin{cases} \inf_{v \in \mathcal{U}} \|u - v\|_\infty & \text{if } \mathcal{U} \text{ is nonempty,} \\ \infty & \text{otherwise,} \end{cases}$$

and the deviation of two utility set $\mathcal{U}_1$ and $\mathcal{U}_2$ as

$$\mathbb{D}(\mathcal{U}_1, \mathcal{U}_2) := \begin{cases} \sup_{u \in \mathcal{U}_1} d(u, \mathcal{U}_2) & \text{if } \mathcal{U}_1 \text{ is nonempty,} \\ \infty & \text{otherwise.} \end{cases}$$

Then the Hausdorff distance of sets $\mathcal{U}_1$ and $\mathcal{U}_2$ is represented as

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

Theorem 4.3 stated in the following shows that, under Assumption (A1), the Hausdorff distance of $\mathcal{U}_N(\epsilon)$ and $\mathcal{U}(\epsilon)$ converges to 0 a.e. as the sample size N increases to ∞. It implies that $\mathcal{U}_N(\epsilon)$ is nonempty for sufficiently large N. Before stating Theorem 4.3, we need discuss two technical lemmas used in the proof of the theorem.

Lemma 4.1 $g_1(u, \mu_N)$ *and* $g_2(u, \mu_N)$ *converges to* $g_1(u, \mu)$ *and* $g_2(u, \mu)$ *uniformly on the set* $\mathfrak{U}$ *a.e., respectively, as* N *increases to* ∞.

Proof: Let us prove the uniform convergence of $g_1(u, \mu)$. The proof of the uniform convergence of $g_2(u, \mu_N)$ is similar. Let

$$G_1(u, t) := (u(t) - u_r(t))^2, \quad \text{and} \quad \hat{g}_1(u, \mu) := \int_\Theta G_1(u, t)\mu(dt).$$

For any $t \in \Theta$, $G_1(\cdot, t)$ is continuous on $\mathfrak{U}$, since for every $\gamma > 0$, we have $|G_1(u_1, t) - G_1(u_2, t)| \leq \gamma$ for all $u_1, u_2 \in \mathfrak{U}$ and $\|u_1 - u_2\|_\infty \leq \gamma^{1/2}$. Theorem 2.2 indicates the compactness of $\mathfrak{U}$ on the space $\mathfrak{L}^\infty$. Also, due to the boundedness of the utility functions in $\mathfrak{U}$, we have that $G_1(u, t) < 1$ for all $u \in \mathfrak{U}$ and $t \in \Theta$.

The above discussion shows that the conditions required by Proposition 7 in Shapiro (2003) are all satisfied in this case. Therefore, we have that $\hat{g}_1(u, \mu_N)$ converges to $\hat{g}_1(u, \mu)$ uniformly on $\mathfrak{U}$ a.e. as N increases to ∞. Since $g_1(u, \mu_N) = (\hat{g}_1(u, \mu_N))^{1/2}$, the uniform convergence of $g_1(u, \mu_N)$ follows. $\qquad\square$

Lemma 4.2 *If Assumption (A1) holds, then* $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}(\epsilon - \tau e)) \downarrow 0$ *as* $\tau \downarrow 0$.

Proof: We argue the convergence by a contradiction. Assume that there exists $\lambda > 0$ such that $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}(\epsilon - \tau e)) > \lambda$ for any $\tau > 0$. Assume that $\lambda \in (0, 1)$ without loss of generality. By Assumption (A1), $\mathcal{U}(\epsilon - \gamma e)$ is nonempty. We choose

$$\delta := \frac{\epsilon_1 - \sqrt{\epsilon_1^2 - 2\epsilon_1\gamma\lambda + \gamma^2\lambda}}{\gamma},$$

and denote

$$\widetilde{\mathcal{U}}_\lambda := \{(1 - \lambda)u_1 + \lambda u_2 : u_1 \in \mathcal{U}(\epsilon), \ u_2 \in \mathcal{U}(\epsilon - \gamma e)\}.$$

Be the definition of δ, we have

$$\epsilon_1 - \delta\gamma \geq \sqrt{\epsilon_1^2 - 2\epsilon_1\gamma\lambda + \gamma^2\lambda^2} = \epsilon - \lambda\gamma.$$

It shows that $0 < \delta \leq \lambda$. Choose an arbitrary $u_\lambda \in \widetilde{\mathcal{U}}_\lambda$, and write $u_\lambda = (1 - \lambda)u_1 + \lambda u_2$ for some $u_1 \in \mathcal{U}(\epsilon)$ and $u_2 \in \mathcal{U}(\epsilon - \gamma e)$. Because the convexity of $g_1^2(u, \mu)$ and $g_2(u, \mu)$ shown in Proposition 2.3, we have

$$g_1^2(u_\lambda, \mu) \leq (1 - \lambda)g_1^2(u_1, \mu) + \lambda g_1^2(u_2, \mu) \leq (1 - \lambda)\epsilon_1^2 + \lambda(\epsilon_1 - \gamma)^2 = (\epsilon_1 - \delta\gamma)^2,$$

and

$$g_2(u_\lambda, \mu) \leq (1 - \lambda)\epsilon_2 + \lambda(\epsilon_2 - \gamma) = \epsilon_2 - \lambda\gamma \leq \epsilon_2 - \delta\gamma.$$

By Theorem 2.4, it follows $u_\lambda \in \mathfrak{U}$. It implies $u_\lambda \in \mathcal{U}(\epsilon - \delta\gamma e)$, and hence, $\widetilde{\mathcal{U}}_\lambda \subseteq \mathcal{U}(\epsilon - \delta\gamma e) \subseteq \mathcal{U}(\epsilon)$. Then $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}(\epsilon - \delta\gamma e)) \leq \mathbb{D}(\mathcal{U}(\epsilon), \widetilde{\mathcal{U}}_\lambda)$. On the other hand, since

$$d(u_\lambda, \mathcal{U}(\epsilon)) \leq \|u_\lambda - u_1\|_\infty = \lambda\|u_1 - u_2\|_\infty \leq \lambda,$$

we have $\mathbb{D}(\mathcal{U}(\epsilon), \widetilde{\mathcal{U}}_\lambda) \leq \lambda$, and thus $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}(\epsilon - \delta\gamma e)) \leq \lambda$. $\qquad\square$

Theorem 4.3 *If Assumption (A1) holds,* $\mathbb{H}(\mathcal{U}_N(\epsilon), \mathcal{U}(\epsilon)) \to 0$ *a.e. as* $N \to \infty$.

Proof: We first prove $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}_N(\epsilon)) \to 0$ as $N \to \infty$ a.e. According to Lemma 4.1, for $\tau > 0$, there exists $N_\tau > 0$ such that, for all $N \geq N_\tau$,

$$\sup_{u \in \mathfrak{U}} |g_k(u, \mu_N) - g_k(u, \mu)| \leq \tau, \quad \text{a.e.,} \quad k = 1, \, 2.$$

Hence, for any $u \in \mathcal{U}(\epsilon - \tau e)$, we have that

$$g_k(u, \mu_N) \leq \tau + g_k(u, \mu) \leq \tau + \epsilon_k - \tau = \epsilon_k.$$

It implies that $\mathcal{U}(\epsilon - \tau e) \subseteq \mathcal{U}_N(\epsilon)$. Using the triangular inequality

$$\mathbb{D}(\mathcal{U}(\epsilon), \, \mathcal{U}_N(\epsilon)) \leq \mathbb{D}(\mathcal{U}(\epsilon), \, \mathcal{U}(\epsilon - \tau e)) + \mathbb{D}(\mathcal{U}(\epsilon - \tau e), \, \mathcal{U}_N(\epsilon)),$$

and the fact shown in Lemma 4.2 that $\mathbb{D}(\mathcal{U}(\epsilon), \, \mathcal{U}(\epsilon - \tau e)) \downarrow 0$, by letting $N \to \infty$ and $\tau \to 0$, we have that $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}_N(\epsilon)) \to 0$ a.e..

On the other hand, we argue that $\mathbb{D}(\mathcal{U}_N(\epsilon), \mathcal{U}(\epsilon)) \to 0$ as $N \to \infty$ a.e. by a contradiction. It has shown that, for any $\tau > 0$, $\mathcal{U}(\epsilon - \tau e) \subseteq \mathcal{U}_N(\epsilon)$ for sufficiently large N. Thus $\mathcal{U}_N(\epsilon)$ is nonempty by Assumption (A1). Suppose that there exists a measurable set $Q \in \mathfrak{G}$ with $\mu(Q) > 0$ such that, given any $\omega \in Q$, we have $\delta > 0$ with the property that for any $j > 0$ there exists $N_j \geq j$ to satisfy $\mathbb{D}(\mathcal{U}_{N_j}(\epsilon), \mathcal{U}(\epsilon)) > \delta > 0$ i.e., there exists $u_{N_j} \in \mathcal{U}_{N_j}(\epsilon)$ such that $d(u_{N_j}, \mathcal{U}(\epsilon)) > \delta > 0$. Thus, by the compactness of $\mathfrak{U}$ shown in Theorem 2.2, we can choose j and N_j to obtain a sequence $\{u_{N_j}\}$ converging to some utility function $\hat{u} \in \mathfrak{U}$ a.e.. Since $d(u_{N_j}, \mathcal{U}(\epsilon)) > \delta$, it follows that $d(\hat{u}, \mathcal{U}(\epsilon)) \geq \delta$. Now, for $k = 1$ and 2,

$$\begin{aligned}|g_k(u_{N_j}, \mu_{N_j}) - g_k(\hat{u}, \mu)| &\leq |g_k(u_{N_j}, \mu_{N_j}) - g_k(u_{N_j}, \mu)| + |g_k(u_{N_j}, \mu) - g_k(\hat{u}, \mu)| \\ &\leq \sup_{u \in \mathfrak{U}} |g_k(u, \mu_{N_j}) - g_k(u, \mu)| + \|u_{N_j} - \hat{u}\|_\infty.\end{aligned}$$

From Lemma 4.1, $g_k(u, \mu_{N_j})$ uniformly converges to $g_k(u, \mu)$ on $\mathfrak{U}$ a.e.. Since $\mu(Q) > 0$, we can assume without loss of generality that, on the sample path ω, $g_k(u, \mu_{N_j})$ uniformly converges to $g_k(u, \mu)$. It follows that for that ω we have

$$\lim_{N_j \to \infty} |g_k(u_{N_j}, \mu_{N_j}) - g_k(\hat{u}, \mu)| = 0,$$

and hence we have

$$g_k(\hat{u}, \mu) = \lim_{N_k \to \infty} g_k(u_{N_k}, \mu_{N_k}) \leq \epsilon.$$

This contradicts that $\hat{u} \notin \mathcal{U}(\epsilon)$. Therefore, $\mathbb{D}(\mathcal{U}_N(\epsilon), \mathcal{U}(\epsilon)) \to 0$ as $N \to \infty$ a.e.. $\qquad \square$

We next show the concavity and continuity of the objective functions f and f_N of problems (1) and (20). If the feasible region $\mathcal{X}$ is convex, the concavity of f and f_N ensures that problems (1) and (20) are convex programming.

Theorem 4.4 *If Assumption (A3) holds, $f(x)$ and $f_N(x)$ are continuous and concave functions on $\mathcal{X}$. If Assumption (A1) also holds, $f_N(x)$ is equicontinuous for sufficiently large N a.e..*

Proof: By Assumption (A3) and the increasing and concavity of $u \in \mathcal{U}(\epsilon)$, it follows that $u(\xi(x))$ is concave in an open neighborhood of $\mathcal{X}$ denoted by $\mathcal{W}(\mathcal{X})$, and thus $F(x, u)$ are concave in $\mathcal{W}(\mathcal{X})$ for every $u \in \mathcal{U}(\epsilon)$. Since $f(x)$ is the minimization of $F(x, u)$ over $\mathcal{U}(\epsilon)$, it is easy to see that $f(x)$ is concave in $\mathcal{W}(\mathcal{X})$, and hence, $f(x)$ is continuous in $\mathcal{X} \subset \mathcal{W}(\mathcal{X})$. A similar proof can also be applied for the continuity and concavity of $f_N(x)$.

11

Theorem 4.3 implies that $\mathcal{U}_N(\epsilon)$ is nonempty for sufficiently large N. Hence, we now consider the equicontinuity of $f_N(x)$ of which N is large enough to ensure a nonempty $\mathcal{U}_N(\epsilon)$. For $x_1,\ x_2 \in \mathcal{X}$, we have that

$$
\begin{aligned}
|f_N(x_1) - f_N(x_2)| &\leq \left| \min_{u \in \mathcal{U}_N(\epsilon)} F(x_1, u) - \min_{u \in \mathcal{U}_N(\epsilon)} F(x_2, u) \right| \\
&\leq \max_{u \in \mathcal{U}_N(\epsilon)} |F(x_1, u) - F(x_2, u)| \\
&\leq \max_{u \in \mathcal{U}_N(\epsilon)} \sum_{k=1}^{K} p_k |u(\xi_k(x_1)) - u(\xi_k(x_2))|.
\end{aligned}
$$

By the equicontinuity of all $u \in \mathfrak{U}$ shown in Proposition 2.1, for any $\sigma > 0$, there exists $\delta > 0$ such that, if $|\xi_k(x_1) - \xi_k(x_2)| \leq \delta$ for all $k = 1, \ldots, K$, then $|u(\xi_k(x_1)) - u(\xi_k(x_2))| \leq \sigma$, and hence $|f_N(x_1) - f_N(x_2)| \leq \sigma$. The concavity of $\xi_k(x)$ given by Assumption (A3) implies the continuity of $\xi_k(x)$ on $\mathcal{X}$. It follows that there exists $\tau > 0$ such that, for all $x_1, x_2 \in \mathcal{X}$ with $\|x_1 - x_2\|_\infty \leq \tau$, we have that $|\xi_k(x_1) - \xi_k(x_2)| \leq \delta$, for all $k = 1, \ldots, K$. This prove the equicontinuity of f_N for sufficiently large N a.e.. $\qquad \square$

We finally discuss the asymptotic convergence of the approximation problem (20) to the counter part (1). In the following, Lemma 4.5 shows that the objective function $f_N(x)$ of problem (20) converges to the objective function $f(x)$ of problem (1) uniformly on the feasible region $\mathcal{X}$ a.e.. Based on this lemma, Theorem 4.6 ensures the the asymptotic convergence of problem (20).

Lemma 4.5 *If Assumptions (A1) - (A3) hold, $f_N(x)$ converges to $f(x)$ uniformly on $\mathcal{X}$ a.e. as N increases to ∞.*

Proof: We first prove that, for every $x \in \mathcal{X}$, $f_N(x)$ converges to $f(x)$ a.e. as N increases to ∞. Let Φ^* and Φ_N^* be the set of optimal solutions of problems (2) and (21). For $u^* \in \Phi^*$, let $u_N := \arg\min_{u \in \mathcal{U}_N(\epsilon)} \|u^* - u\|_\infty$. By Theorem 4.3 that shows $\mathbb{D}(\mathcal{U}(\epsilon), \mathcal{U}_N(\epsilon)) \to 0$ a.e., we have that $u_N \to u^*$ as $N \to \infty$ a.e.. It is trivial to see that $F(x, u)$ is continuous in u. Hence,

$$
\limsup_{N \to \infty} f_N(x) \leq \limsup_{N \to \infty} F(x, u_N) = F(x, u^*) = f(x) \quad \text{a.e..}
$$

Also, by $\mathbb{D}(\mathcal{U}_N(\epsilon), \mathcal{U}(\epsilon)) \to 0$ a.e., for any convergence sequence $\{\bar{u}_N\}$ with $\bar{u}_N \in \Phi_N^*$, the limit point $\bar{u}$ is in $\mathcal{U}(\epsilon)$. Hence,

$$
\liminf_{N \to \infty} f_N(x) \geq f(x), \quad \text{a.e..}
$$

Next, we check the uniform convergence of $f_N(x)$. Under Assumptions (A1) and (A3), Theorem 4.4 shows the continuity of $f(x)$ and the equicontinuity of $f_N(x)$ on $\mathcal{X}$. Using Assumption (A2), there exist finitely many points, $x_1, \ldots, x_r \in \mathcal{X}$, and corresponding neighborhoods $\mathcal{W}_1, \ldots, \mathcal{W}_r$ covering $\mathcal{X}$ such that, for any $\delta > 0$ and $j = 1, \ldots, r$,

$$
\sup_{x \in \mathcal{W}_j \cap \mathcal{X}} |f(x) - f(x_j)| < \delta/3
$$

and for sufficiently large N,

$$
\sup_{x \in \mathcal{W}_j \cap \mathcal{X}} |f_N(x) - f_N(x_j)| < \delta/3, \quad \text{a.e..}
$$

Furthermore, by the convergence of f_N discussed above, we have

$$
|f_N(x_j) - f(x_j)| < \delta/3, \quad j = 1, \ldots, r, \quad \text{a.e.,}
$$

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

$$|f_N(x) - f(x)| \leq |f_N(x) - f_N(x_j)| + |f_N(x_j) - f(x_j)| + |f(x) - f(x_j)| \leq \delta.$$

Therefore,

$$\sup_{x \in \mathcal{X}} |f_N(x) - f(x)| \leq \delta, \quad \text{a.e.}.$$

which shows that f_N uniformly converges to f on $\mathcal{X}$ a.e.. $\qquad \square$

Theorem 4.6 *Let y^* and $\mathcal{Z}$ be the optimal values and the set of optimal solutions of problem* (1). *Let y_N and $\mathcal{Z}_N$ be those of the approximation problem* (20). *Suppose Assumptions (A1) - (A3) hold. Then $y_N \to y^*$ and $\mathbb{D}(\mathcal{Z}_N, \mathcal{Z}) \to 0$ a.e. as $N \to \infty$.*

Proof: Under Assumptions (A1) and (A2), problem (1) is feasible. Theorem 4.3 guarantees the feasibility of problem (20) for sufficiently large N. Under Assumption (A3), Theorem 4.4 shows the continuity of the objective function f of problem (1). Also, under Assumptions (A1) - (A3), Lemma 4.5 proves that the objective function $f_N(x)$ of problem (20) converges to $f(x)$ uniformly on $\mathcal{X}$ a.e. It follows from Theorem 5.3 in Shapiro et al. (2009) that $y_N \to y^*$ and $\mathbb{D}(\mathcal{Z}_N, \mathcal{Z}) \to 0$ a.e. as $N \to \infty$. $\qquad \square$

5 Case Study: Portfolio Optimization Problem

We now study an application of the robust decision making framework (1), using data from the portfolio investment problem discussed in Dentcheva and Ruszczyński (2003). The example has $J(=8)$ assets which are widely used market indexes: U.S. three-month treasury bills, U.S. long-term government bonds, S&P 500, Willshire 5000, NASDAQ, Lehman Brothers corporate bond index, EAFE foreign stock index, and gold. Dentcheva and Ruszczyński (2003) use $K(=22)$ yearly returns of these assets as equally probable realizations. Let $\xi = (\xi_1, \ldots, \xi_n)$ denote the random yearly return of these assets, which has realizations $\xi^k = (\xi_1^k, \ldots, \xi_n^k)$ with probability $p_k = 1/K$ for $k = 1, \ldots, K$. For the completeness, we give this historical data in Table 3 in the appendix.

An investor puts \$1 into the index fund market, and his wealth will be $\$(1 + \xi^T x)$ in a year. Using the framework (1), we present a robust portfolio investment problem as

$$\max_{x \in \mathcal{X}} \min_{u \in \mathcal{U}(\epsilon)} \sum_{k=1}^{K} p_k u \left(1 + \xi^{kT} x\right). \tag{22}$$

Here, $\mathcal{X} := \{x \in \mathbb{R}_+^n \mid x^T e = 1\}$ is the set of all available portfolios, $\mathcal{U}(\epsilon)$ is the perturbation region specified in (9). Assume that the investor's best hope is a 200% money return, while he understands that it is possible for him to lose all investment. Also assume that, to the best of our knowledge, his risk preference is most likely characterized as the CARA utility function of which the coefficient of absolute risk aversion (or the Arrow-Pratt measure of absolute risk-aversion (ARA)) is 1.5. Hence, in model (22), we choose $\Theta = [0, 2]$, and specify the reference utility function as

$$u_r(t) = \frac{1 - e^{-1.5t}}{1 - e^{-3}}, \quad t \in \Theta, \tag{23}$$

which is the dotted curve drawn in Figure 1. The data shows that the returns most varies in $[-30\%, \ 30\%]$, and hence, a portfolio decision is very sensitive to the characterization of investor's

risk preference in the middle of Θ. Using the Monte Carlo method to solve problem (22), we should choose μ to obtain more samples in the middle of Θ. In the study, we let μ be the beta distribution on Θ with shape parameters $\alpha = \beta = 6$, of which the pdf is

$$b(t; \alpha, \beta) := \left(\frac{\left(\frac{t}{\theta}\right)^{\alpha-1} \left(1 - \frac{t}{\theta}\right)^{\beta-1}}{\int_\Theta \left(\frac{s}{\theta}\right)^{\alpha-1} \left(1 - \frac{s}{\theta}\right)^{\beta-1} ds} \right) \mathbf{1} \{t \in \Theta\}.$$

The mean and vairance of this beta distribution are

$$\frac{\alpha\theta}{\alpha+\beta} = 1, \qquad \text{and} \qquad \frac{\alpha\beta\theta^2}{(\alpha+\beta)^2(\alpha+\beta+1)} = 0.0769.$$
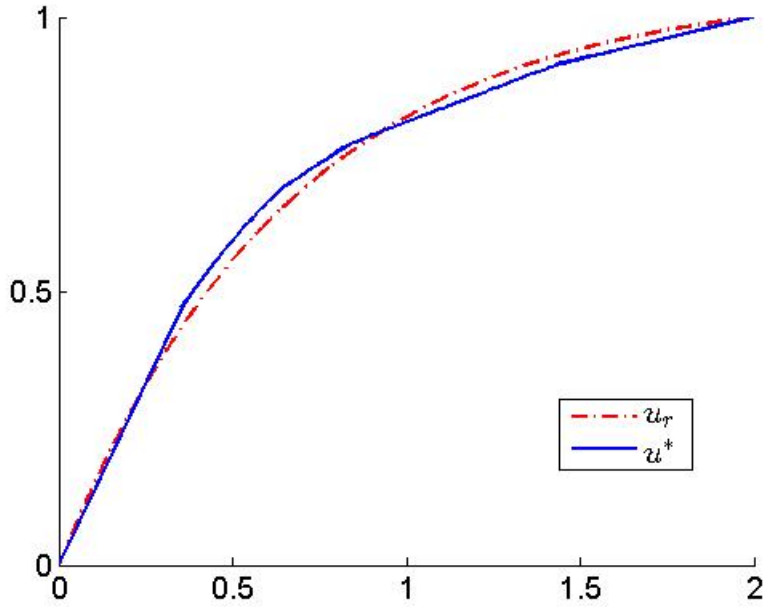

Figure 1: Reference and optimal utility functions

Table 1: Optimal Investment Portfolios

	Optimal Solution							
ϵ_1	A1	A2	A3	A4	A5	A6	A7	A8
0	0	0	0	39.32%	0	0	48.38%	12.34%
0.005	0	0	0	31.38%	0	0	61.46%	7.16%
0.01	0	0	0	21.87%	0	0	72.85%	5.28%
0.015	0	0	0	19.72%	0	0	75.84%	4.44%
0.02	0	0	0	17.98%	0	0	78.70%	3.33%

In the discussion, we permit the perturbation to have a slightly asymmetry by fixing $\epsilon_2 = 10^{-8}$, and approximate model (22) by (20) with the size of 5000 samples. When choosing $\epsilon_1 = 0.02$, the corresponding optimal utility functions u^* given in (16) is drawn as the solid curve in Figure 1. It
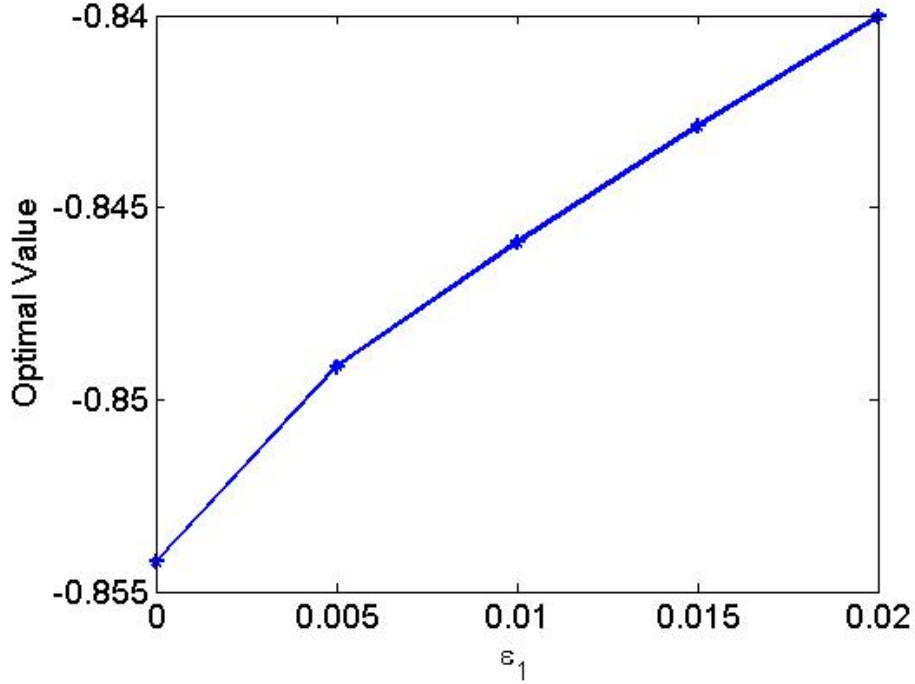
Figure 2: Price function of perturbation $c(\epsilon_1)$

Table 2: Convergence and computational complexity

Sample				Optimal Solution					Optimal	Running
Size	A1	A2	A3	A4	A5	A6	A7	A8	Value	Time (Sec)
500	0	0	0	18.40%	0	0	77.34%	4.26%	0.8413	0.236
1000	0	0	0	15.00%	0	0	79.89%	5.12%	0.8413	0.698
2000	0	0	0	16.98%	0	0	79.51%	3.51%	0.8405	1.669
3000	0	0	0	17.48%	0	0	78.94%	3.58%	0.8406	2.768
4000	0	0	0	17.83%	0	0	78.95%	3.22%	0.8407	4.030
5000	0	0	0	17.98%	0	0	78.70%	3.33%	0.8406	4.923

is well know that the ARA indicates the level of risk aversion (Arrow (1965) and Pratt (1964)). As we mentioned, the CARA reference utility function u_r gives the constant ARA 1.5. In comparison, u^* is more risk averse in [0, 1] and less risk averse in [1, 2]. By adjusting ϵ_1 from 0 to 0.02, we test the impact of ϵ_1 on the optimal values and solutions of model (22). Table 1 gives the optimal investment portfolios for different ϵ_1's. At the growth of ϵ_1, Asset A7 is given more investment, while Assets 4 and 8 are weaken in the portfolios. The solutions at $\epsilon_1 = 0$ and 0.005 are very unstable. Increasing ϵ_1 results in a large change in the investment. In comparison, the solutions are stable for $\epsilon_1 \geq 0.01$. Theorem 2.5 shows that the price function of perturbation $c(\epsilon_1)$ is an increasing concave function of ϵ_1. This conclusion is observed in Figure 2, and also indicates that increasing ϵ_1 from 0 to 0.005 results in the largest change.

Table 2 gives the convergence analysis and computational performance of Cplex 12.4.0 for solving the approximation problem (20) for $\epsilon_1 = 0.02$ using different sample sizes. The test runs on HP ProBook 6470b with Intel® Core™ i7-3520M Processor. After the sample size is above 2000, the optimal values and solutions have very small fluctuation.

References

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

D. Bertsimas, D. B. Brown, and C. Caramnis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research. Springer-Verlag, New York, NY, 1997.

C. Boutilier, R. Patrascu, P. Poupart, and D. Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9):686–713, 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.

D. Dentcheva and A. Ruszczyński. Optimization with stochastic dominance constraints. *SIAM J. Optim.*, 14(2):548–566, 2003.

N. Dunford and J. T. Schwartz. *Linear operators*, volume 1. Wiley-Interscience, 1958.

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

D. G. Fromberg 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. Mehrotra. Robust decision making using a risk averse utility set with an application to portfolio optimization. Submitted to IIE Transaction `http://www.optimization-online.org/DB_FILE/2012/03/3412.pdf`, 2012a.

J. Hu and S. Mehrotra. Robust decision making using a general utility set. under second review for SIAM Journal on Optimization `http://www.optimization-online.org/DB_FILE/2012/07/3527.pdf`, 2012b.

S. Jacobson and R. Petrie. Inconsistent choices in lottery experiments: Evidence from rwanda. Technique Report, Georgia State University, Experimental Economics Center Working Paper Series 01/2007, `http://www.researchgate.net/publication/5021404_Inconsistent_Choices_in_Lottery_Experiments_Evidence_from_Rwanda`, 2007.

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.

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.

H. L. Royden. *Real Analysis*. Macmillan, 1988.

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. Monte Carlo sampling methods. In A. Ruszczynski and A. Shapiro., editors, *Handbook of Stochastic Optimization*. Elsevier Science Publishers B.V., Amsterdam, Netherlands, 2003.

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.

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

M. Weber. Decision making with incomplete information. *European Journal of Operational Research*, 28(1):44–57, 1987.

A   Asset Returns given in Dentcheva and Ruszczyński (2003)

Table 3: Asset Returns (in %) in Dentcheva and Ruszczyński (2003)

Year	Asset 1	Asset 2	Asset 3	Asset 4	Asset 5	Asset 6	Asset 7	Asset 8
1	7.5	-5.8	-14.8	-18.5	-30.2	2.3	-14.9	67.7
2	8.4	2.0	-26.5	-28.4	-33.8	0.2	-23.2	72.2
3	6.1	5.6	37.1	38.5	31.8	12.3	35.4	-24.0
4	5.2	17.5	23.6	26.6	28.0	15.6	2.5	-4.0
5	5.5	0.2	-7.4	-2.6	9.3	3.0	18.1	20.0
6	7.7	-1.8	6.4	9.3	14.6	1.2	32.6	29.5
7	10.9	-2.2	18.4	25.6	30.7	2.3	4.8	21.2
8	12.7	-5.3	32.3	33.7	36.7	3.1	22.6	29.6
9	15.6	0.3	-5.1	-3.7	-1.0	7.3	-2.3	-31.2
10	11.7	46.5	21.5	18.7	21.3	31.1	-1.9	8.4
11	9.2	-1.5	22.4	23.5	21.7	8.0	23.7	-12.8
12	10.3	15.9	6.1	3.0	-9.7	15.0	7.4	-17.5
13	8.0	36.6	31.6	32.6	33.3	21.3	56.2	0.6
14	6.3	30.9	18.6	16.1	8.6	15.6	69.4	21.6
15	6.1	-7.5	5.2	2.3	-4.1	2.3	24.6	24.4
16	7.1	8.6	16.5	17.9	16.5	7.6	28.3	-13.9
17	8.7	21.2	31.6	29.2	20.4	14.2	10.5	-2.3
18	8.0	5.4	-3.2	-6.2	-17.0	8.3	-23.4	-7.8
19	5.7	19.3	30.4	34.2	59.4	16.1	12.1	-4.2
20	3.6	7.9	7.6	9.0	17.4	7.6	-12.2	-7.4
21	3.1	21.7	10.0	11.3	16.2	11.0	32.6	14.6
22	4.5	-11.1	1.2	-0.1	-3.2	-3.5	7.8	-1.0