Robust Concave Utility Maximization over Chance Constraints

Shanshan Wang
Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Shatin, Hong Kong
shshwangbit@gmail.com

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

Chun Peng*
School of Economics and Management, Beijing Jiaotong University, Beijing, China
chunpeng@bjtu.edu.cn

This paper first studies an expected utility problem with chance constraints and incomplete information on a decision maker's utility function. The model maximizes the worst-case expected utility of random outcome over a set of concave functions within a novel ambiguity set, while the underlying probability distribution is known. To obtain computationally tractable formulations, we employ a discretization approach to derive a max-min chance-constrained approximation of this problem. This approximation is further reformulated as a mixed-integer program. We show that the discrete approximation converges to the true counterpart under mild assumptions. We also present a row generation algorithm for optimizing the max-min program. A computational study for a bin-packing problem and a multi-item newsvendor problem is conducted to demonstrate the benefit of the proposed framework and the computational efficiency of our algorithm. We find that the row generation algorithm can significantly reduce the computational time, and our robust policy could achieve a better out-of-sample performance when compared with the non-robust policy and the one without the chance constraints.

Key words: Discrete optimization, robust expected utility, chance constraint, bin packing, multi-item newsvendor

1. Introduction

Decision-making under uncertainty frequently involves balancing the value of under- and over-utilization of a resource, while ensuring certain service performance. For example, in healthcare operations, random demand results in underage and overage costs (Kim and Mehrotra 2015). In an operating room (OR) planning problem, underage and overage costs are regarded as undertime and overtime costs in the opened ORs, respectively. In this context, the underage and overage can be modeled using a nonlinear function, which is important since the costs may not be linear

* corresponding author

(Davis et al. 2014). As additional examples, for a finite duration work shift, in the case of assigning patients to a finite number of timeslots, the random duration required to fully serve a patient has implications on work-life balance (Guest 2002, Azeem and Akhtar 2014). Here to achieve the work-life balance, we want to ensure that the assigned patients finish within a certain time with a desirable probability. A similar situation arises in scheduling ORs with multiple surgeries (e.g., Wang et al. 2021, 2022). In a newsvendor model, especially for the current business trends of shortening product life-cycles, it is crucial to find a solution that optimally trades off between under- and over-stocking (Natarajan et al. 2018, Hu et al. 2019, Wang and Delage 2024). To ensure inventory capacity, we also want a bound on the over-stocking probability. In this paper, we use utility functions to address the value of under- and over-utilization of a resource. In the meantime, a chance constraint paradigm can be employed to ensure certain service performance.

In practice, it is difficult to estimate an exact utility function $u(\cdot)$ due to the absence of precise and comprehensive explanations regarding human behavior, while crude estimated approaches might be available (see Chajewska et al. 2000, Hu and Mehrotra 2015, Hu et al. 2018, Armbruster and Delage 2015, and references therein). Therefore, for such problems it is prudent to assume that the utility function u is unknown, and specify a model over an ambiguity set $\mathscr{U}$ that allows a family of utility functions based on their shape and properties such that the decision-making under uncertainty could balance the implications of under- and over-utilization. In the following, we will propose a general modeling framework to address such an important class of decision-making problems with incomplete information on the utility function.

1.1. Modeling Framework

The decision-making framework maximizes the worst-case expected utility of random outcomes over a set of concave functions, with chance constraints. We consider the functions that first increase and then decrease, which suitably model the situations described above. This framework is novel since, for the first time, it combines the concept of robust decision making and chance constraint optimization as a complementary synergistic mechanism for decision modeling under risk and uncertainty. Specifically, we consider the robust utility maximization (RUM) problem with chance constraints, represented as

$$\text{(RUM)} \quad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{maximize}}\ \underset{u\in\mathscr{U}}{\text{minimize}} \sum_{j\in\mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))], \tag{1a}$$

$$\text{subject to } \mathbb{P}_{\mathbb{Q}}\left(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}) \le t_j\right) \ge 1-\varepsilon, \qquad \forall j\in\mathcal{J}, \tag{1b}$$

where $\mathcal{X} := \left\{\boldsymbol{x} := (\boldsymbol{x}_1,\boldsymbol{x}_2)\in\mathbb{R}^{n_1}\times\mathbb{N}^{n_2} \;\middle|\; \boldsymbol{A}\boldsymbol{x}\le\boldsymbol{d}\right\}$, $n := n_1+n_2$, and $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $\boldsymbol{d}\in\mathbb{R}^m$. We assume that $\mathcal{X}$ is a non-empty compact set. $u(\cdot):\mathbb{R}\mapsto[0,1]$ is a first increasing then decreasing concave utility function and lies in an ambiguity set $\mathscr{U}$. We define the set $\mathscr{U}$ by using functional bounds on the utility and an additional condition that is specified by using a reference utility function. $\mathbb{Q}$ denotes a joint distribution of random vector $\tilde{\boldsymbol{\xi}} := (\tilde{\xi}_1, \cdots, \tilde{\xi}_n)^\top \in \Xi$. $f_j(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) : \mathcal{X} \times \Xi \mapsto \mathbb{R}$ is a measurable random function for $j \in \mathcal{J} := \{1, \cdots, |\mathcal{J}|\}$, and $|\cdot|$ is the cardinality of a set. $\boldsymbol{t} := \{t_1, \cdots, t_{|\mathcal{J}|}\} \in \mathbb{R}^{|\mathcal{J}|}$, and $\varepsilon \in [0,1]$ is a given upper bound on the allowed probability of violation. The objective of (RUM) is to maximize the worst-case total expected utility. Chance constraints (1b) require that the constraints are satisfied with a given probability $1-\varepsilon$. In this paper, we assume that the probability distribution of $\tilde{\boldsymbol{\xi}}$ has finite support $(\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^{|\Omega|})$ such that $\mathbb{P}(\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}^\omega) = p_\omega$ for $\omega \in \Omega := \{1, \cdots, |\Omega|\}$, where $\sum_{\omega \in \Omega} p_\omega = 1$ and $p_\omega \geq 0$ for $\omega \in \Omega$.

1.2. Mathematical Formulations of Illustrative Examples

As discussed before, the model studied in this paper is motivated by many applications that involve balancing costs resulting from random overage and underage. The number of patients assigned to a clinician for service, or the number of surgeries assigned to an OR can be thought of as a bin-packing problem with each item having a random size. We formally describe this model below, followed by a description of a multi-item newsvendor problem.

Bin Packing with Chance and Utility. Let $\mathcal{I} := \{1, \cdots, |\mathcal{I}|\}$ denote a set of items and $\mathcal{J} := \{1, \cdots, |\mathcal{J}|\}$ denote a set of homogeneous bins. We assign $|\mathcal{I}|$ items with random size $\tilde{\boldsymbol{\xi}} := (\tilde{\xi}_1, \cdots, \tilde{\xi}_{|\mathcal{I}|})^\top$ to $|\mathcal{J}|$ bins. We use ξ_i^ω to denote the size of item $i \in \mathcal{I}$ under scenario $\omega \in \Omega$, and c_j to represent the capacity of bin $j \in \mathcal{J}$.

We define a binary variable x_{ij} such that $x_{ij} = 1$ if item $i \in \mathcal{I}$ is assigned to bin $j \in \mathcal{J}$, and $x_{ij} = 0$ otherwise. The bin packing robust utility maximization (BP-RUM) problem is formulated as follows:

$$\text{(BP-RUM)} \quad \underset{\boldsymbol{x} \in \{0,1\}^{|\mathcal{I}||\mathcal{J}|}}{\text{maximize}} \ \underset{u \in \mathscr{U}}{\text{minimize}} \sum_{j \in \mathcal{J}} \mathbb{E}[u(\sum_{i \in \mathcal{I}} \tilde{\xi}_i x_{ij} - c_j)] \tag{2a}$$

$$\text{subject to } \mathbb{P}_{\mathbb{Q}} \left\{ \sum_{i \in \mathcal{I}} \tilde{\xi}_i x_{ij} - c_j \leq 0 \right\} \geq 1 - \varepsilon, \qquad \forall j \in \mathcal{J}, \tag{2b}$$

$$\sum_{j \in \mathcal{J}} x_{ij} = 1, \qquad \forall i \in \mathcal{I}. \tag{2c}$$

For the (BP-RUM), the objective function (2a) is to maximize the worst-case total expected utility of over- and under-utilization. Constraints (2b) require that the sum of item sizes assigned to bin j is less than the capacity of bin j with a probability $1-\varepsilon$. Constraints (2c) ensure that each item is assigned to exactly one bin.

Multi-Item Newsvendor with Chance and Utility. The multi-item newsvendor robust expected utility problem is to decide the order quantities for each item with a random demand, so as to maximize the worst-case expected utility of the under- and over-stocking. More specifically, let $\mathcal{J} := \{1, \cdots, |\mathcal{J}|\}$ denote the set of items. We let the inventory capacity $\boldsymbol{t} := \{t_1, \cdots, t_{|\mathcal{J}|}\}$, and the random demand of items $\tilde{\boldsymbol{\xi}} := \{\tilde{\xi}_1, \cdots, \tilde{\xi}_{|\mathcal{J}|}\}$. Under the scenarios $\omega \in \Omega$, the demand $\boldsymbol{\xi}^\omega := \{\xi_1^\omega, \cdots, \xi_{|\mathcal{J}|}^\omega\}$. Each item $j \in \mathcal{J}$ has an ordering cost o_j, and d is the total budget. Decision variable $\boldsymbol{x} := \{x_1, \cdots, x_{|\mathcal{J}|}\}$ denotes the ordering quantities. The following formulates the multi-item newsvendor robust expected utility problem (MN-RUM):

$$\text{(MN-RUM)} \quad \underset{\boldsymbol{x}\in\mathbb{R}_+^{|\mathcal{J}|}}{\text{maximize}} \ \underset{u\in\mathscr{U}}{\text{minimize}} \sum_{j\in\mathcal{J}} \mathbb{E}\left[u\left(x_j - \tilde{\xi}_j\right)\right] \tag{3a}$$

$$\text{subject to } \mathbb{P}_{\mathbb{Q}}\left\{x_j - \tilde{\xi}_j \le t_j\right\} \ge 1 - \varepsilon, \qquad \forall j \in \mathcal{J}, \tag{3b}$$

$$\boldsymbol{o}^\top \boldsymbol{x} \le d. \tag{3c}$$

For the (MN-RUM), the objective function (3a) is to maximize the worst-case total expected utility of under- and over-stocking. Constraints (3b) ensure that the over-stocking is less than the inventory capacity with the probability $1-\varepsilon$. Constraint (3c) ensures that the total ordering cost is no more than the budget.

1.3. Literature Review

In this section, we provide a review of the existing studies that are relevant to our work from both the methodology and application aspects. More specifically, we mainly focus on the literature about robust expected utility frameworks and two applications that are mentioned above.

1.3.1. Literature Review on Robust Expected Utility In the robust optimization (RO) framework, the uncertain data lies in an uncertainty set (e.g., Bertsimas and Sim 2004), and the decision-makers aim to identify the solutions that perform best under the worst-case realizations within an uncertainty set and are robust to estimation errors. RO has been extensively developed in terms of new methodologies and its practical applications (e.g., see recent reviews by Bertsimas et al. 2011, Gabrel et al. 2014, Gorissen et al. 2015, Rahimian and Mehrotra 2019).

In terms of robust expected utility, Hansen and Sargent (2001) described a connection between the max-min expected utility theory and robust-control theory. Schied (2005) defined the robust utility function by using a set of probability measures and reformulated the terminal wealth problem as a standard utility-maximization problem associated with a "subjective" probability measure. Natarajan et al. (2010) studied a robust expected utility model for portfolio optimization, where only the mean, covariance, and support information are available and the investor's utility is a piecewise-linear concave function. Armbruster and Delage (2015) considered the problem of maximizing the worst-case expected utility of random outcome over a set of utility functions that are assumed to be risk-averse, S-shaped, or prudent, and finally derived a linear program (LP) reformulation. Haskell et al. (2016) further extended this work by considering ambiguity about both the decision maker's risk preferences and the underlying distribution. They obtained an LP reformulation under the assumption of a polyhedral distributional ambiguity set with a finite number of vertices. For more general ambiguity sets, they proposed conservative approximations that are based on reformulation-linearization techniques. Delage et al. (2022) considered a utility-based shortfall risk measure where the true loss function is unavailable and proposed a preference robust model by constructing a set of utility-based loss functions from empirical data or subjective judgments. Luo and Mehrotra (2024) studied a service center location problem with ambiguous utility gains upon receiving service under a distributionally robust optimization (DRO) framework, where the elicited location-dependent utilities are assumed to be described by an expected value and variance constraint.

Perhaps, the most relevant studies to ours are Hu and Mehrotra (2015) and Hu et al. (2018). Hu and Mehrotra (2015) assumed that the utility function is increasing and concave. They specified the uncertainty set by using upper bound and lower bound on the utility and marginal utility functions, as well as auxiliary equality and inequality constraints on the utility. They used a partitioning-based approach to formulate the problem as an LP. More recently, Hu et al. (2018) assumed that the uncertainty set of the utility function is non-decreasing and satisfies additional boundary and auxiliary conditions. They developed a sample average approximation (SAA) based approach (Kleywegt et al. 2002) to solve the problem. Unlike the aforementioned two studies, our work considers the utility-dependent decisions within a chance-constrained framework and constructs a novel ambiguity set in the space of risk-averse utility centered at a reference utility function using a distance metric. Moreover, our problem is to balance the value of under- and over-utilization of a resource. This allows us to model a more general set of utility functions that are first increasing then decreasing. In addition, the different ambiguity sets lead to different convergence analyses, Hu and Mehrotra (2015) used probability theory to demonstrate the convergence, while our convergence analysis relies on the convergence theory of optimization problems. Finally, we propose a row generation algorithm with the strengthened strategy proposed by Bodur and Luedtke (2017) to efficiently solve our problem with two practical applications.

1.3.2. Literature Review on Chance-Constrained Bin Packing Problem Chance-constrained programs (CCPs) were first introduced by Charnes and Cooper (1959) to address optimization problems under uncertainty, which have been widely used for various decision-making contexts. CCPs are generally difficult to solve (e.g., Song et al. 2014), especially when the coefficients matrix is random or the chance constraints contain integer decision variables (as is shown in our (BP-RUM) in Section 1.2). For the study of more general CCPs under different optimization settings, we refer the interested reader to a recent review by Küçükyavuz and Jiang (2022).

In recent years, the chance-constrained bin packing problem (CCBP) has been extensively studied, especially under the context of healthcare resources allocation (e.g., ORs and surgeries) and cloud computing management (e.g., Hoogervorst et al. 2019, Cohen et al. 2019, Martinovic et al. 2021). For healthcare resource allocation, Deng and Shen (2016), Deng et al. (2019) and Zhang et al. (2020) investigated a surgery scheduling problem with chance constraints to determine ORs allocation and surgery scheduling using the stochastic programming and DRO paradigms. More recently, Wang et al. (2021) studied a chance-constrained multiple bin packing problem with application to ORs planning. Wang et al. (2022) further extended this work to a DRO model with joint chance constraints with partial distribution information. Instead of minimizing the total cost or the number of bins as in the above studies, we consider utility-dependent decisions within a chance-constrained framework to ensure certain service performance, while balancing the implications of under- and over-utilization of resources. Although CCBP is widely studied, to our best knowledge, such applications under a robust expected utility framework are very rare.

1.3.3. Literature on Multi-Item Newsvendor Problem The newsvendor problem is a fundamental operations management problem with various applications (see a recent review by Qin et al. 2011). To determine the order quantities for multiple products, the retailers assume a specifically known distribution of the random demand (e.g., Erlebacher 2000). However, in reality, the true demand distribution is hardly ever known to the retailers. Leveraging recent advances in RO, robust multi-item newsvendor problems aim to maximize the worst-case expected operating revenue over all possible demand realizations within an uncertainty set (e.g., Ardestani-Jaafari and Delage 2016, Hu et al. 2019, Zhang et al. 2021). For most real-world applications, the solutions of RO models are generally over-conservative, thus DRO multi-item newsvendor problems have been extensively explored in recent years, where one seeks a more robust solution that performs best under the worst-case demand distribution within an ambiguity set of distributions (e.g., Hanasusanto et al. 2015, Natarajan et al. 2018, Rahimian et al. 2019, Chen et al. 2020, Wang and Delage 2024). Very few studies attempt to use alternative risk preferences within the expected utility framework (Wang et al. 2012, Choi and Ruszczyński 2011), which assume that the distributions of product demands and the utility function are exactly known in advance. However, it is very difficult to derive the exact representation of the utility function in practice. This also further motivates us to model the problem under a robust expected utility framework using a novel ambiguity set of utility functions.

1.4. Contributions of This Paper

This paper addresses the modeling framework and resolution method of a general chance-constrained robust expected utility problem over a set of concave utility functions that lie in a novel ambiguity set. Under mild conditions, we derive a mixed-integer program (MIP), and conduct the convergence analysis for (RUM) that relies on the convergence theory of optimization problems, and also develop a row generation-based solution scheme to solve the resulting problem efficiently. The contributions of this paper are summarized as follows:

- We construct a novel ambiguity set in the space of risk-averse utility centered at a reference utility function using a distance metric. To model the novel utility ambiguity set $\mathscr{U}$, we employ a discretization scheme where piecewise-linear approximations with N partitions are used. In doing so, we are able to reformulate (RUM) as a tractable MIP with the help of a big-M technique. Then a convergence analysis is provided to show that the discrete approximation converges to the true counterpart under some mild assumptions. To the best of our knowledge, this is the first attempt to study a general robust expected utility problem with chance constraints when the information of the utility function is incomplete.
- We propose an efficient row generation-based solution scheme to solve the robust expected utility model. More specifically, we represent (RUM) as a max-min formulation and investigate the row generation approach for solving the problem. The algorithm considers a master problem as a chance-constrained problem and a subproblem. We also strengthened the row generation algorithm by using the strategy proposed by Bodur and Luedtke (2017).
- We perform an extensive numerical study for the bin-packing problem using real data from surgery planning and the multi-item newsvendor problem to analyze the general structure of the decisions from the decision-making framework and show the benefits of the techniques developed in this paper for computational improvement. We find that the row generation algorithm significantly outperforms a commercial solver. Furthermore, the average objective value is mainly unchanged after the number of partitions reaches 10 for (BP-RUM) and (MN-RUM). Therefore, $N = 10$ can be used to obtain approximate solutions with reasonable solution time. We also evaluate the out-of-sample performance of the solutions generated from (BP-RUM) and (MN-RUM). The results show the advantages of (BP-RUM) and (MN-RUM) when compared with the ones without the chance constraints and their determinant counterparts.

1.5. Organization

The remainder of this paper is organized as follows. Section 2 gives a definition of ambiguity set $\mathscr{U}$ and formulates (RUM) as a MIP using a discrete approximation of $\mathscr{U}$ and big-M techniques. We then present convergence analysis to show that the discrete approximation converges to the true counterpart under some mild assumptions in Section 3. In Section 4, we present a row generation algorithm to solve (RUM). Section 5 reports the computational results of the robust expected utility model for the bin-packing problem and the multi-item newsvendor problem. Section 6 concludes the paper with a summary of the important findings. The Online Appendix includes all the proofs of lemmas and theorems stated in the paper.

2. Model Formulation

Section 2.1 introduces the definition of the ambiguity set $\mathscr{U}$ and examples for the risk-averse utility functions. Using a discretization scheme, we then reformulate the robust expected chance-constrained problem as an MIP in Section 2.2.

2.1. The Ambiguity Set Definition

We assume the function $f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}) \in \Theta := [-\theta_1, \theta_2]$ for all $\boldsymbol{x} \in \mathcal{X}, \tilde{\boldsymbol{\xi}} \in \Xi$, and $j \in \mathcal{J}$, where $\theta_1, \theta_2 \in \mathbb{R}^+$. We construct the ambiguity set $\mathscr{U}$ such that every function u in $\mathscr{U}$ satisfies these (and later) properties.

$$u(-\theta_1) = 0, \quad u(0) = 1, \quad u(\theta_2) = 0. \tag{4}$$

We use function $\bar{u}$ and $\underline{u}$ as the bounds of u, that is to say,

$$\underline{u}(a) \leq u(a) \leq \bar{u}(a), \qquad\qquad a \in \Theta. \tag{5}$$

We construct the following ball in the space of risk-averse utility centered at a reference utility function u_0:

$$d(u, u_0) \leq b^{\frac{1}{p}}, \tag{6}$$

where radius b is a positive constant and $d(u, u_0)$ is the distance between two functions u and u_0 which is defined as the L_p-norm of $u - u_0$, i.e.,

$$d(u, u_0) = \left(\int_{-\theta_1}^{\theta_2} |u(a) - u_0(a)|^p \, da \right)^{\frac{1}{p}}.$$

Constraint (6) ensures that the utility functions are real-valued integrable in the domain using u_0 as a reference, with a pre-specified bound. Let $\mathscr{U}'$ be the set of all first increasing then decreasing concave utility functions defined on Θ. We have the following ambiguity set $\mathscr{U}$,

$$\mathscr{U} := \left\{ u \in \mathscr{U}' \;\middle|\; u \text{ satisfies the conditions in the constraints } (4) - (6) \right\}.$$

This utility set is different from the one used in Hu and Mehrotra (2015), in the use of constraint (6). The set used in Hu and Mehrotra (2015) is specified using constraints linear in u, and u is non-decreasing. Here we are allowing u to increase as well as decrease. In addition, Hu and Mehrotra (2015) used limited preference queries to define an ambiguity set $\mathscr{U}$. In Section 2.1.1, we provide how to assess the reference utility function with the help of multiple methods that are typically used in the literature. This allows us to model a more flexible ambiguity set. Moreover, our distance-based ambiguity set contains all utility functions that are close to a most likely function (i.e., the reference utility function). The decision-maker could adjust the radius of the ambiguity set to control the degree of conservatism of the expected utility problem. While the confidence of the uncertainty set is also of high theoretical relevance, the main goal of our paper is practical relevance. Thus, we primarily focus on providing reformulations for our model with the use of a discretization scheme. We leave the statistical performance guarantee open for further research.

2.1.1. Approaches for Assessing the Ambiguity Set In the following, we provide some approaches that could be used to assess the parameters in $\mathscr{U}$. For the bounds on u, we may use information from parametric approaches to specify them. Let the bounds on u be derived from the power utility function (see, e.g., Brunello 2002, Holt and Laury 2002), which are given by

$$u_r(a) = \begin{cases} \left(\dfrac{a+\theta_1}{\theta_1}\right)^{1-r}, & a \in [-\theta_1, 0], \\ \left(\dfrac{\theta_2 - a}{\theta_2}\right)^{1-r}, & a \in (0, \theta_2], \end{cases} \tag{7}$$

where parameter r is the constant coefficient of relative risk aversion. Holt and Laury (2002) suggested that one could set $r \in [0.41, 0.68]$ for a risk-averse decision-makers. Thus, we use $u_{0.41}(a) \leq u(a) \leq u_{0.68}(a)$ to specify a risk-averse decision maker's utility set. Figure 1 gives the upper bound and lower bound of the utility function when $\theta_1 = \theta_2 = 1$:

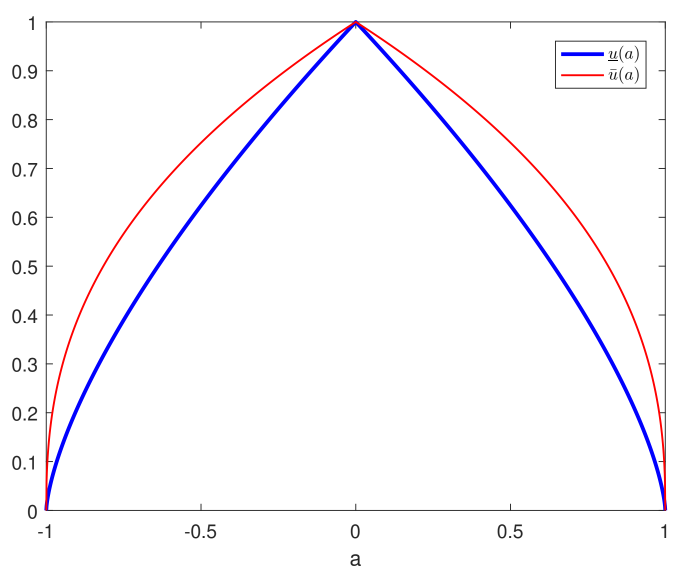

Figure 1 **The upper and lower bounds of utility functions when $\theta_1 = \theta_2 = 1$.**

Several approaches have been developed in the literature that can be used to assess the reference utility function, such as discrete choice models (Train and Weeks 2005), the value equivalence method, and the certainty equivalence method (Farquhar 1984). Nowadays, some studies have used the new proposals based on machine learning (ML) methods to evaluate the classical utility function, which showed a high predictive capacity of ML methods (see, e.g., Wang et al. 2020, Martín-Baos et al. 2021). In practice, one can use more than one assessment procedure to evaluate the reference utility function, which is helpful in choosing appropriate utility assessment procedures for particular decision situations (Farquhar 1984).

2.2. Reformulation of (RUM)

We present a reformulation for (RUM) associated with a set $\mathscr{U}_N$, which is defined using a discretization of the continuous problem. Let N be the number of partitions, $\mathcal{A}(N)=\{a_0,\cdots,a_l,\cdots,a_N\}$ be a set of break points such that $a_0<\cdots<a_l<\cdots<a_N$; $a_0=-\theta_1$, $a_l=0$; and $a_N=\theta_2$. We assume that if $N_1<N_2$, then $\mathcal{A}(N_1)\subset\mathcal{A}(N_2)$. We define the following piecewise linear approximation functions of $\underline{u}$ and $\bar{u}$:

$$\underline{u}_N(a)=\sum_{k=0}^{N-1}\left(\frac{\underline{u}(a_{k+1})-\underline{u}(a_k)}{a_{k+1}-a_k}a+\frac{a_{k+1}\underline{u}(a_k)-a_k\underline{u}(a_{k+1})}{a_{k+1}-a_k}\right)\mathbf{1}(a_k\le a<a_{k+1}),\tag{8}$$

$$\bar{u}_N(a)=\sum_{k=0}^{N-1}\left(\frac{\bar{u}(a_{k+1})-\bar{u}(a_k)}{a_{k+1}-a_k}a+\frac{a_{k+1}\bar{u}(a_k)-a_k\bar{u}(a_{k+1})}{a_{k+1}-a_k}\right)\mathbf{1}(a_k\le a<a_{k+1}),\tag{9}$$

where $\mathbf{1}(\cdot)$ represents the indicator function, which returns 1 if the clause inside is correct, and otherwise 0.

Following Figure 1, Figure 2 gives the approximation of bounds of utility functions when the number of partitions $N=20$.

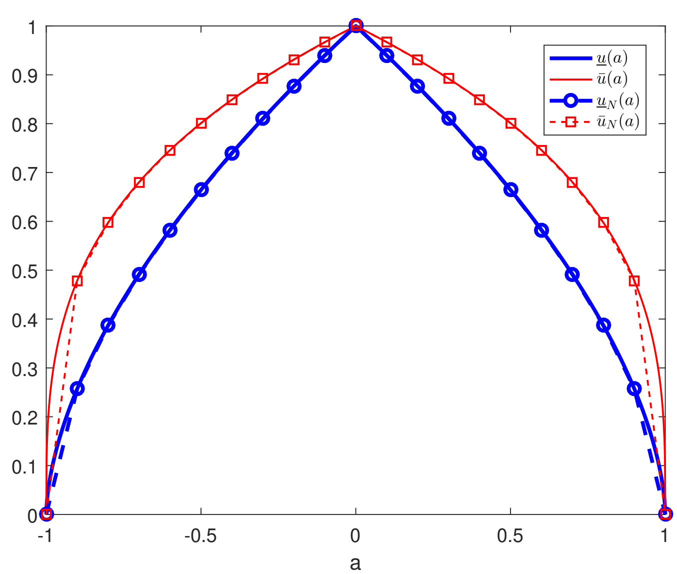

Figure 2 **Approximations of bounds of utility functions when** $N=20$**.**

Let the ambiguity set

$$\mathscr{U}_N := \left\{ u \in \mathscr{U}' \;\middle|\; \begin{array}{l} u \text{ satisfies the conditions in constraint (4)}, \; \underline{u}_N \le u \le \bar{u}_N, \text{ and} \\ \sum_{k=0}^{N-1} \left(u_0(a_k) - u(a_k)\right)^p (a_{k+1} - a_k) \le b, \end{array} \right\},$$

and

$$\pi_N(\boldsymbol{x}) = \underset{u \in \mathscr{U}_N}{\text{minimize}} \sum_{j \in \mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))], \tag{10}$$

then Lemma 1 below gives a reformulation of $\pi_N(\boldsymbol{x})$ for any given $\boldsymbol{x} \in \mathcal{X}$.

Lemma 1. *For a given $\boldsymbol{x} \in \mathcal{X}$, problem (10) is equivalent to the following reformulation:*

$$\underset{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}}{\text{minimize}} \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} p_\omega (f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \alpha^{j\omega} + \beta^{j\omega}) \tag{11a}$$

$$\text{subject to } (a_{k+1} - a_{k-1})\theta_k \ge (a_{k+1} - a_k)\theta_{k-1} + (a_k - a_{k-1})\theta_{k+1}, \quad \forall k \in \{1, \cdots, N-1\}, \tag{11b}$$

$$\theta_k \le \bar{u}_N(a_k), \quad \forall k \in \{0, \cdots, N\}, \tag{11c}$$

$$\theta_k \ge \underline{u}_N(a_k), \quad \forall k \in \{0, \cdots, N\}, \tag{11d}$$

$$||\boldsymbol{Q}(\boldsymbol{\theta} - \boldsymbol{u}_0)||_p \le 1, \tag{11e}$$

$$a_k \alpha^{j\omega} + \beta^{j\omega} - \theta_k \ge 0, \quad \forall k \in \{0, \cdots, N\}, j \in \mathcal{J}, \omega \in \Omega, \tag{11f}$$

$$\beta^{j\omega} \ge 0, \quad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{11g}$$

where $\theta_0 = 0$, $\theta_l = 1$ and $\theta_N = 0$, $\boldsymbol{Q} = diag\left((\frac{a_1 - a_0}{b})^{\frac{1}{p}}, \cdots, (\frac{a_N - a_{N-1}}{b})^{\frac{1}{p}} \right)$ is a diagonal matrix, and $\boldsymbol{\theta} - \boldsymbol{u}_0 := (\theta_0 - u_0(a_0), \cdots, \theta_{N-1} - u_0(a_{N-1}))^\top$.

Note that the chance constraints can be rewritten as

$$f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) + (M_{j\omega} - t_j) z_{j\omega} \le M_{j\omega}, \quad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{12a}$$

$$\sum_{\omega \in \Omega} p_\omega z_{j\omega} \ge 1 - \varepsilon, \quad \forall j \in \mathcal{J}, \tag{12b}$$

$$z_{j\omega} \in \{0, 1\}, \quad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{12c}$$

where $M_{j\omega}$ is a large constant such that constraint (12a) still holds when $z_{j\omega} = 0$ for $j \in \mathcal{J}, \omega \in \Omega$. We use the coefficient strengthening procedure inspired from Song et al. (2014) to obtain a tight value of $M_{j\omega}$. For all $j \in \mathcal{J}, \omega \in \Omega$,

$$M_{j\omega} \ge \bar{M}_{j\omega} := \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{maximize}} \left\{ f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \middle| \mathbb{P}\left\{ f_j(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) - t_j \le 0 \right\} \ge 1 - \varepsilon \right\}.$$

For any $j \in \mathcal{J}, \omega, k \in \Omega$, let

$$m_{j\omega}(k) := \underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{maximize}} \left\{ f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \middle| f_j(\boldsymbol{x}, \boldsymbol{\xi}^k) \le t_j \right\}.$$

We sort $m_{j\omega}(k)$ in a non-decreasing order such that $m_{j\omega}(k_1) \le \ldots \le m_{j\omega}(k_N)$. Then $m_{j\omega}(k_{q+1})$ is an upper bound for $\bar{M}_{j\omega}$, if $q = \max\left\{l : \sum_{i=1}^{l} p_{k_i} \le \varepsilon\right\}$.

Using the strong duality theory for the minimization problem (11), we obtain a reformulation of (RUM) with the ambiguity set $\mathscr{U}_N$ in Theorem 1.

THEOREM 1. *The problem*

$$\begin{aligned} &\underset{\boldsymbol{x}\in\mathcal{X}}{\text{maximize}}\ \underset{u\in\mathscr{U}_N}{\text{minimize}} \sum_{j\in\mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))] && (13a)\\ &\text{subject to (1b)}, \end{aligned}$$

is equivalent to the following problem:

$$\begin{aligned}
&\underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{z},\boldsymbol{\mu},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\delta},\boldsymbol{\eta},\lambda}{\text{maximize}} && \sum_{k=0}^{N}(\underline{u}_N(a_k)\gamma_{2k} - \bar{u}_N(a_k)\gamma_{1k}) + \sum_{k=0}^{N-1}\eta_k u_0(a_k) - \lambda && && (14a)\\
&\text{subject to} && p_\omega f_j(\boldsymbol{x},\boldsymbol{\xi}^\omega) - \sum_{k=0}^{N} a_k\delta_{kj\omega} = 0, && \forall j\in\mathcal{J},\omega\in\Omega, && (14b)\\
& && p_\omega - \sum_{k=0}^{N}\delta_{kj\omega} \ge 0, && \forall j\in\mathcal{J},\omega\in\Omega, && (14c)\\
& && \eta_k = -\mu_k(a_{k+1}-a_{k-1}) + \mu_{k+1}(a_{k+2}-a_{k+1}) && &&\\
& && +\mu_{k-1}(a_{k-1}-a_{k-2}) + \gamma_{1k} - \gamma_{2k} + \sum_{j\in\mathcal{J}}\sum_{\omega\in\Omega}\delta_{kj\omega}, && \forall k\in\{1,\cdots,N-1\}, && (14d)\\
& && ||\boldsymbol{Q}^{-1}\boldsymbol{\eta}||_* \le \lambda, && && (14e)\\
& && f_j(\boldsymbol{x},\boldsymbol{\xi}^\omega) + (m_{j\omega}(k_{q+1}) - t_j)z_{j\omega} \le m_{j\omega}(k_{q+1}), && \forall j\in\mathcal{J},\omega\in\Omega, && (14f)\\
& && \sum_{\omega\in\Omega} p_\omega z_{j\omega} \ge 1-\varepsilon, && \forall j\in\mathcal{J}, && (14g)\\
& && \boldsymbol{\mu},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\delta} \ge 0, && && (14h)\\
& && z_{j\omega}\in\{0,1\}, && \forall j\in\mathcal{J},\omega\in\Omega, && (14i)
\end{aligned}$$

where $\mu_N = \mu_0 = 0$, $a_{-1}, a_{N+1} \in \mathbb{R}$, *and* $||\cdot||_*$ *is the dual norm of* $||\cdot||_p$.

In the following, Corollary 1 and Corollary 2 give the final reformulations of (BP-RUM) and (MN-RUM), respectively.

COROLLARY 1. *Based on Theorem 1, the final reformulation of (BP-RUM) under the ambiguity set* $\mathscr{U}_N$ *can be represented as follows:*

$$\begin{aligned}
&\underset{\boldsymbol{x},\boldsymbol{z},\boldsymbol{\mu},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\delta},\lambda}{\text{maximize}} \sum_{k=0}^{N}(\underline{u}_N(a_k)\gamma_{2k} - \bar{u}_N(a_k)\gamma_{1k}) + \sum_{k=0}^{N-1}\eta_k u_0(a_k) - \lambda && (15a)\\
&\text{subject to (2c)}, && (15b)
\end{aligned}$$

$$p_\omega \left(\sum_{i \in \mathcal{I}} \xi_i^\omega x_{ij} - c_j \right) - \sum_{k=0}^{N} a_k \delta_{kj\omega} = 0, \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{15c}$$

$$p_\omega - \sum_{k=0}^{N} \delta_{kj\omega} \geq 0, \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{15d}$$

$$\begin{aligned} \eta_k &= -\mu_k(a_{k+1} - a_{k-1}) + \mu_{k+1}(a_{k+2} - a_{k+1}) \\ &+ \mu_{k-1}(a_{k-1} - a_{k-2}) + \gamma_{1k} - \gamma_{2k} + \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} \delta_{kj\omega}, \qquad \forall k \in \{1, \cdots, N-1\}, \end{aligned} \tag{15e}$$

$$||\boldsymbol{Q}^{-1} \boldsymbol{\eta}||_* \leq \lambda, \tag{15f}$$

$$\sum_{i \in \mathcal{I}} \xi_i^\omega x_{ij} - c_j + m_{j\omega}(k_{q+1}) z_{j\omega} \leq m_{j\omega}(k_{q+1}), \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{15g}$$

$$\sum_{\omega \in \Omega} p_\omega z_{j\omega} \geq 1 - \varepsilon, \qquad \forall j \in \mathcal{J}, \tag{15h}$$

$$\boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta} \geq 0, \tag{15i}$$

$$x_{ij}, z_{j\omega} \in \{0,1\}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, \omega \in \Omega, \tag{15j}$$

where $\mu_N = \mu_0 = 0$, $a_{-1}, a_{N+1} \in \mathbb{R}$, *and* $m_{j\omega}(k) = \underset{x_{ij} \in \{0,1\}}{\text{maximize}} \left\{ \sum_{i \in \mathcal{I}} \xi_i^\omega x_{ij} - c_j \middle| \sum_{i \in \mathcal{I}} \xi_i^k x_{ij} - c_j \leq 0 \right\}$, *for* $j \in \mathcal{J}$, $k, \omega \in \Omega$.

Corollary 2. *Based on Theorem 1, we can reformulate (MN-RUM) with the ambiguity set* $\mathscr{U}_N$ *as the following problem:*

$$\underset{\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta}, \lambda}{\text{maximize}} \sum_{k=0}^{N} (\underline{u}_N(a_k)\gamma_{2k} - \bar{u}_N(a_k)\gamma_{1k}) + \sum_{k=0}^{N-1} \eta_k u_0(a_k) - \lambda \tag{16a}$$

$$\text{subject to (3c)}, \tag{16b}$$

$$p_\omega(x_j - \xi_j^\omega) - \sum_{k=0}^{N} a_k \delta_{kj\omega} = 0, \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{16c}$$

$$p_\omega - \sum_{k=0}^{N} \delta_{kj\omega} \geq 0, \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{16d}$$

$$\begin{aligned} \eta_k &= -\mu_k(a_{k+1} - a_{k-1}) + \mu_{k+1}(a_{k+2} - a_{k+1}) \\ &+ \mu_{k-1}(a_{k-1} - a_{k-2}) + \gamma_{1k} - \gamma_{2k} + \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} \delta_{kj\omega}, \qquad \forall k \in \{1, \cdots, N-1\}, \end{aligned} \tag{16e}$$

$$||\boldsymbol{Q}^{-1} \boldsymbol{\eta}||_* \leq \lambda, \tag{16f}$$

$$x_j - \xi_j^\omega + (m_{j\omega}(k_{q+1}) - t_j) z_{j\omega} \leq m_{j\omega}(k_{q+1}), \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{16g}$$

$$\sum_{\omega \in \Omega} p_\omega z_{j\omega} \geq 1 - \varepsilon, \qquad \forall j \in \mathcal{J}, \tag{16h}$$

$$\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta} \geq 0, \tag{16i}$$

$$z_{j\omega} \in \{0,1\}, \qquad \forall j \in \mathcal{J}, \omega \in \Omega, \tag{16j}$$

where $\mu_N = \mu_0 = 0$, $a_{-1}, a_{N+1} \in \mathbb{R}$, *and* $m_{j\omega}(k) = \underset{x_j \in \mathbb{R}_+}{\text{maximize}} \left\{ x_j - \xi_j^\omega \middle| x_j - \xi_j^k \leq t_j \right\}$, *for* $j \in \mathcal{J}$, $k, \omega \in \Omega$.

3. Convergence Analysis

In this section, we show that the optimal solutions obtained by using the discrete approximation of the set $\mathscr{U}$ converge to the true optimal solutions as N increases to ∞. Throughout this section, we make the following assumptions, which are also commonly used in the literature.

ASSUMPTION 1. *(i) $f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}})$ is linear in an open neighborhood of $\mathcal{X}$ for all $j \in \mathcal{J}$; (ii) $\mathscr{U}$ is a non-empty set, and function $u \in \mathscr{U}$ has bounded derivative almost everywhere; (iii) $\bar{u}$, $\underline{u}$ and u_0 are continuous concave functions.*

Note that $\underline{u}_N$ and $\bar{u}_N$ are the approximation functions of $\underline{u}$ and $\bar{u}$ based on the $\mathcal{A}(N)$. Also, under the above assumption, $\underline{u}_N$ and $\bar{u}_N$ uniformly converge to $\underline{u}$ and $\bar{u}$, respectively. In the following, under Assumption 1, we first show that $\pi(\boldsymbol{x}) := \underset{u \in \mathscr{U}}{\text{minimize}} \sum_{j \in \mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))]$ and $\pi_N(\boldsymbol{x})$ are continuous concave functions.

LEMMA 2. *$\pi(\boldsymbol{x})$ and $\pi_N(\boldsymbol{x})$ are continuous concave functions on $\mathcal{X}$.*

The following Lemma 3 to Lemma 7 give some preliminary results, which are needed to prove the convergence of $\pi_N(\boldsymbol{x})$ to $\pi(\boldsymbol{x})$.

LEMMA 3. *All $u \in \mathscr{U}_N$ are equicontinuous.*

LEMMA 4. *(Arzelà–Ascoli Theorem in Green and Valentine (1961)) Let $\mathcal{K}$ be a compact metric space, with metric $d_K(p,p')$, and let $C(\mathcal{K})$ denote the space of real (or complex) valued continuous functions on $\mathcal{K}$. If $\{f_n\}_{n\in N}$ is a sequence in $C(\mathcal{K})$ obeying:*

- *$\{f_n\}_{n\in N}$ is pointwise bounded, and*
- *$\{f_n\}_{n\in N}$ is equicontinuous,*

then, the sequence $\{f_n\}_{n\in N}$ contains a uniformly convergent subsequence.

LEMMA 5. *For any sequence $\{u_N \in \mathscr{U}_N\}$, there exists a subsequence $\{u_{N_K}\}$ that uniformly converges to $\hat{u} \in \mathscr{U}$.*

LEMMA 6. *For any $u \in \mathscr{U}$, there exists a sequence $u_N \in \mathscr{U}_N$ such that $u = \lim_{N\to\infty} u_N$.*

We state Theorem 2.3 in Alvarez-Mena and Hernández-Lerma (2005) in the following lemma for completeness, which gives some convergence conditions of $\pi_N(\boldsymbol{x})$.

LEMMA 7. *For $\boldsymbol{x} \in \mathcal{X}$, let $\{u_N\}$ be a sequence of minimizers of $\{\pi_N(\boldsymbol{x})\}$. Suppose (i) a subsequence $\{u_{N_k}\}$ of $\{u_N\}$ converges to $u \in \mathscr{U}$; (ii) $\liminf_{k\to\infty} \pi_{N_k}(\boldsymbol{x}) \geq \pi(\boldsymbol{x})$; (iii) for any $u \in \mathscr{U}$, there exists a sequence $u_N \in \mathscr{U}_N$ such that $u = \lim_{N\to\infty} u_N$ and $\sum_{j\in\mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))] = \lim_{N\to\infty} \sum_{j\in\mathcal{J}} \mathbb{E}[u_N(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))]$. Then u is optimal for $\{\pi(\boldsymbol{x})\}$. Furthermore, $\pi_{N_k}(\boldsymbol{x})$ converges to $\{\pi(\boldsymbol{x})\}$.*

Based on Lemma 5 and 7, we shows that $\pi_N(\boldsymbol{x})$ converges to $\pi(\boldsymbol{x})$ for all $\boldsymbol{x}\in\mathcal{X}$ in Lemma 8 .

Lemma 8. *$\pi_N(\boldsymbol{x})\to\pi(\boldsymbol{x})$ as $N\to\infty$ for all $\boldsymbol{x}\in\mathcal{X}$.*

Referring to theorem 5.3 in Shapiro et al. (2009), the following Lemma gives some convergence conditions of problem (13).

Lemma 9. *Let y_N and Z_N be the optimal objective value and the set of optimal solutions of problem (13) and y^* and Z^* be those of (RUM). Denote the deviation of sets Z_N and Z^* as $D(Z_N,Z^*):=\max_{x_1\in Z_N}\min_{x_2\in Z^*}||x_1-x_2||$. Suppose (i) X is a non-empty compact set; (ii) the function $\pi(\cdot)$ is continuous on X; and (iii) $\pi_N(\cdot)$ uniformly converges to $\pi(\cdot)$ on X as $N\to\infty$. Then, $y_N\to y^*$ and $D(Z_N,Z^*)\to 0$ as $N\to\infty$.*

Theorem 2 shows that the optimal solutions of problem (13) converge to the true optimal solution of (RUM) in the limit.

Theorem 2. *Let y_N and Z_N be the optimal objective value and the set of solutions of problem (13), and y^* and Z^* be the optimal objective value and the set of solutions of (RUM). Then $y_N\to y^*$ and $D(Z_N,Z^*):=\underset{b_1\in Z_N}{\text{maximize}}\,\underset{b_2\in Z^*}{\text{minimize}}\,||b_1-b_2||\to 0$ as $N\to\infty$.*

4. A Row Generation Solution Scheme

Note that solving the reformulation of (RUM) by an off-the-shelf commercial solver (e.g., CPLEX) directly might be time-consuming, which will be further confirmed by our numerical study in Section 5. Instead, in this section we propose a row generation algorithm as a solution method for our (RUM).

Based on Lemma 1, (RUM) can be further approximated by the following max-min problem:

$$\underset{\boldsymbol{x}\in\mathcal{X},\rho\in\mathcal{P}}{\text{maximize}}\ \rho \tag{17a}$$

$$\text{subject to } (\rho,\boldsymbol{x})\in\mathcal{Z}, \tag{17b}$$

where set $\mathcal{P}:=[0,|\mathcal{J}|]$, $\mathcal{Z}=\{(\rho,\boldsymbol{x})\in\mathbb{R}\times\mathbb{R}^n:\rho\le Z(\boldsymbol{x})\}$, and $Z(\boldsymbol{x})$ is given as follows:

$$\text{(SP)}\qquad Z(\boldsymbol{x})=\underset{\boldsymbol{\theta},\boldsymbol{\alpha},\boldsymbol{\beta}}{\text{minimize}}\sum_{j\in\mathcal{J}}\sum_{\omega\in\Omega}p_\omega\left(f_j(\boldsymbol{x},\boldsymbol{\xi}^\omega)\alpha^{j\omega}+\beta^{j\omega}\right)$$
$$\text{subject to } (11b)-(11g).$$

Give $\boldsymbol{x}\in\mathcal{X}$, since $Z(\boldsymbol{x})$ is the equivalent formulation of problem $\pi_N(\boldsymbol{x})$, $Z(\boldsymbol{x})$ is based on problem (11). Moreover, we rewrite the chance constraints as constraints (14f) and (14g) to obtain the master problem. We then define the master problem as follows:

$$\text{(MP)}\qquad \underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{z},\rho\in\mathcal{P}}{\text{maximize}}\ \rho$$

$$\text{subject to } (14f), (14g), (14i),$$
$$(\rho, \boldsymbol{x}) \in \hat{\mathcal{Z}},$$

where $\hat{\mathcal{Z}}$ is a polyhedral relaxation of $\mathcal{Z}$. If we use $\hat{\mathcal{Z}} = \mathcal{Z}$, then (MP) is an equivalent reformulation of (17). In the row generation algorithm, $\hat{\mathcal{Z}}$ is gradually improved by adding cuts defining the set $\mathcal{Z}$.

An outline of the row generation algorithm is given in Algorithm 1. The row generation solves (MP) and (SP) iteratively until the stopping criteria are met. We let UB and LB denote the upper and lower bound, respectively. We initialize the number of iterations ℓ to 0, UB to positive infinity, and LB to negative infinity. In each iteration, we update $\ell := \ell + 1$ and solve the linear relaxation of (MP) to obtain an optimal solution $(\boldsymbol{x}^\ell, \boldsymbol{z}^\ell, \rho^\ell)$ and optimal objective value $uobj^\ell$. Recall that $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathbb{R}^{n_1} \times \mathbb{N}^{n_2}$, thus, $\boldsymbol{x}^\ell = (\boldsymbol{x}_1^\ell, \boldsymbol{x}_2^\ell)$. If the objective value $uobj^\ell$ is larger than the current lower bound: when $(\boldsymbol{x}_2^\ell, \boldsymbol{z}^\ell)$ is integer, we solve (SP) with $\boldsymbol{x}$ fixed to be $\boldsymbol{x}^\ell$ to attain an optimal solution $(\boldsymbol{\theta}^\ell, \boldsymbol{\alpha}^\ell, \boldsymbol{\beta}^\ell)$ and optimal objective value $lobj^\ell$. If $uobj^\ell$ is larger than $lobj^\ell$, we add the cut $\rho \leq \sum_{j\in\mathcal{J}} \sum_{\omega\in\Omega} p_\omega (f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \alpha^{j\omega\ell} + \beta^{j\omega\ell})$ to the (MP), else update LB if necessary. When $(\boldsymbol{x}_2^\ell, \boldsymbol{z}^\ell)$ is fractional, we update UB if necessary. We terminate the algorithm when the stopping criteria are satisfied, and return the optimal value LB and the optimal solution $(\boldsymbol{x}^*, \boldsymbol{z}^*, \rho^*)$. The following result from Kelley (1960) shows that Algorithm 1 can find a solution of (RUM) in a finite number of iterations under certain conditions.

THEOREM 3. *Let* $\boldsymbol{v} := (\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ *and* $\mathcal{V} := \{\boldsymbol{v} \mid (11b) - (11g)\}$. *We also let* $\boldsymbol{t} := (\boldsymbol{x}, \rho)$, $\mathcal{T} := \mathcal{X} \times \mathcal{P}$, *and* $g(\boldsymbol{t}, \boldsymbol{v}) := \rho - \sum_{j\in\mathcal{J}} \sum_{\omega\in\Omega} p_\omega (f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \alpha^{j\omega} + \beta^{j\omega})$. *We assume that* $\mathcal{T} \times \mathcal{V}$ *is compact and* $g(\boldsymbol{t}, \boldsymbol{v})$ *is continuous on* $\mathcal{T} \times \mathcal{V}$. *If there exists an oracle that solves (MP) and (SP) to optimality at each iteration, then Algorithm 1 terminates within finitely many iterations. If* $UB < +\infty$, *then Algorithm 1 obtains a solution* $(\boldsymbol{x}^*, \boldsymbol{z}^*, \rho^*)$ *of a desired accuracy at termination.*

The row generation algorithm could be further improved by a strategy proposed by Bodur and Luedtke (2017). Specifically, the linear relaxation of (MP) is solved using row generation. At each iteration, the linear relaxation is solved, and the cut (18) violated by the current linear relaxation solution is added to the linear relaxation. This loop is repeated until no more violated cuts are found or some stopping criteria are met. Then, all the cuts found by far are added to the initial (MP) formulation of the row generation algorithm. Our numerical study indicates that this strategy can improve the performance of the row generation method for most types of instances.

5. Numerical Study

We numerically evaluate the performance of our robust utility maximization model and the proposed row generation solution scheme with the help of the bin packing problem in Section 5.1 and the multi-item newsvendor problem in Section 5.2.

Algorithm 1: Row Generation Algorithm

Initialize The number of iteration $\ell = 0$, UB $= +\infty$, LB $= -\infty$, a tolerance $\epsilon' > 0$, and a small number $\tau > 0$.

Initialize $\mathcal{N} = \{o\}$, where o is the root node with the LP relaxation of (MP).

while *($\mathcal{N}$ is nonempty and $UB - LB > \epsilon'$)* **do**

- Select a node $o \in \mathcal{N}$, $\mathcal{N} \leftarrow \mathcal{N}/\{o\}$.
- Solve the linear relaxation of (MP) at the node o to obtain the optimal solution $(\boldsymbol{x}^\ell, \boldsymbol{z}^\ell, \rho^\ell)$ and objective value $uobj^\ell$. $\ell = \ell + 1$.
- **if** $uobj^\ell > LB$ **then**
  - **if** $(\boldsymbol{x}_2^\ell, \boldsymbol{z}^\ell)$ *is integer* **then**
    - Fix $\boldsymbol{x}$ to be $\boldsymbol{x}^\ell$, and solve (SP) and obtain an optimal solution $(\boldsymbol{\theta}^\ell, \boldsymbol{\alpha}^\ell, \boldsymbol{\beta}^\ell)$ and objective value $lobj^\ell$.
    - **if** $uobj^\ell - lobj^\ell > \tau$ **then**
      - Add the following cut

$$\rho \leq \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} p_\omega (f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) \alpha^{j\omega\ell} + \beta^{j\omega\ell}) \tag{18}$$

        to (MP).
    - **end**
  - **else**
    - Update LB $:= \max\{\text{LB}, lobj^\ell\}$.
  - **end**
  - **end**
- **else**
  - Update UB $:= \min\{\text{UB}, uobj^\ell\}$.
  - Branch, resulting in nodes o^* and o^{**}, $\mathcal{N} \leftarrow \mathcal{N} \cup \{o^*, o^{**}\}$.
- **end**
- **end**

end

return LB and its corresponding optimal solution $(\boldsymbol{x}^*, \boldsymbol{z}^*, \rho^*)$.

5.1. Bin Packing Robust Expected Utility Problem

We use the real data from a surgery planning problem to show the performance of the algorithm proposed in Section 4 and the general structure of the decision made from (BP-RUM). In the context of the surgery planning problem, the bins are ORs, items are surgeries, and capacity denotes the OR time limit. We describe implementation details in Section 5.1.1 and present the computational results in Section 5.1.2. Section 5.1.3 shows the out-of-sample performance.

5.1.1. Implementation Details We use historical surgery duration data from a large public hospital in Beijing, China from January 2015 to October 2015, in which 5,721 historical observations of surgery duration are employed to generate our samples (see Wang et al. (2021) for a more detailed description). More specifically, we use log-normal distribution with the mean and the standard deviation of the surgery duration to generate surgery duration samples and round the samples to the nearest 15 minutes. Equal probabilities are used as in the SAA method. The daily time limit c_j is set to 10 hours for $j \in \mathcal{J}$. Nine major surgery types are performed in a day and the percentage for each surgery type are used to calculate the number of surgeries for each surgery type. We assign 12 surgeries ($|\mathcal{I}| = 12$) to four ORs ($|\mathcal{J}| = 4$) a day. The number of scenarios is set to be 30. The bound support $\Theta = [-10, 14]$, and the power utility functions (7) are used as the bounds. We consider the number of partitions $N \in \{4, 10, 20\}$, $b \in \{1, 2, 5\}$ and $\varepsilon \in \{0.1, 0.2\}$. $u_{0.41}(a)$ and $u_{0.68}(a)$ are used as the lower bound and upper bound of the utility function, respectively. We also let the reference utility function be the power utility function with $r = 0.55$. For each sample size, five instances were generated. Therefore, all the performance is reported over five instances on average.

All experiments are conducted on a laptop with Intel(R) 2.80 GHz processor and 16 GB RAM. We implement the algorithm and models in the C programming language using IBM CPLEX solver, version 12.71 callable libraries. We set the runtime limit as two hours and the relative optimality gap tolerance as 0.5%. For instances that could not be solved to optimality, we give the average relative optimality gap, where the gap is calculated as $\frac{(\text{UB}-\text{LB})}{\text{UB}} * 100$, and UB and LB are the upper and lower bound, respectively. We report the average CPU solution time (in seconds) for the instances that are solved to optimality within the runtime limit.

5.1.2. Computational Results We first provide the computational performance of the row generation algorithm (i.e., Algorithm 1) for solving (BP-RUM), then present the objective values of (BP-RUM) associated with the set $\mathscr{U}_N$.

Performance of row generation algorithm. The performances of the following three variants are compared:

- CPX: refers to using CPLEX to directly solve MIP reformulation of (BP-RUM).
- RG: refers to using the row generation algorithm (Algorithm 1) to solve (BP-RUM).
- RG_stren: refers to using the row generation algorithm (Algorithm 1) with the strategy proposed by Bodur and Luedtke (2017), which is described in Section 4. When solving the linear relaxation of (MP), we set the maximum number of iterations to be 5.

Table 1 reports the average CPU solution time for solving (BP-RUM), the subproblem (SP), and the linear relaxation of (MP), the average number of cuts (18), and the number of instances that are solved to optimality over the five generated instances.

We observe from Table 1 that, RG significantly outperforms CPX in terms of the average solution time for all types of instances, and RG_stren has a better performance than RG for most types of instances. Specifically, RG saves more than 90% solution time when compared with CPX, and RG_stren saves more than 28% solution time when compared with RG. We also observe that CPX

Table 1 The average CPU (in seconds) solution time for solving (BP-RUM) (AvT), the subproblem (SP) (AvT-Z), and the linear relaxation of (MP) (AvT-L), the average number of cuts (cuts), and the number of solved instances from the five instances (solved) are reported.

b	ε	N	CPX		RG				RG_stren				
			AvT	solved	AvT	AvT-Z	cuts	solved	AvT	AvT-Z	AvT-L	cuts	solved
1	0.1	4	6,555[0.6]	1/5	71	22	950	5/5	43	12	0.4	959	5/5
		10	[3.1]	0/5	208	97	1,639	5/5	99	48	0.3	1,636	5/5
		20	[3.2]	0/5	1,039	917	1,640	5/5	528	458	2.7	1,615	5/5
	0.2	4	[5.4]	0/5	98	31	2,077	5/5	98	25	0.3	2,081	5/5
		10	[5.5]	0/5	290	111	3,753	5/5	278	104	0.4	3,517	5/5
		20	[5.5]	0/5	2,506[0.6]	2,304	3,400	4/5	1,818	1,621	3.4	3,329	5/5
2	0.1	4	5,285[0.8]	1/5	83	24	939	5/5	51	13	0.3	1,041	5/5
		10	[2.6]	0/5	237	108	1,715	5/5	115	59	0.5	1,620	5/5
		20	[3.3]	0/5	816	709	1,679	5/5	454	391	2.3	1,674	5/5
	0.2	4	[5.5]	0/5	73	23	1,985	5/5	129	26	0.3	2,191	5/5
		10	[5.4]	0/5	266	111	3,629	5/5	265	105	0.4	3,460	5/5
		20	[5.5]	0/5	2,065	1,833	3,511	5/5	2,415	2,149	4.4	3,611	5/5
5	0.1	4	[3.6]	0/5	70	20	927	5/5	43	11	0.3	955	5/5
		10	[3.7]	0/5	138	71	1,653	5/5	101	46	0.3	1,653	5/5
		20	[4.2]	0/5	1,164	1,029	1,798	5/5	470	399	2.1	1,692	5/5
	0.2	4	[5.5]	0/5	380	99	2,355	5/5	92	25	0.3	2,029	5/5
		10	[5.6]	0/5	1,202	470	3,680	5/5	415	123	0.4	3,876	5/5
		20	[5.7]	0/5	4,318[0.7]	3,897	2,745	2/5	3,039	2,726	5.1	3,521	5/5

"[·]" in column of AvT means the average relative optimality gap (%) for instances that cannot be solved to optimality within the time limit.

can only solve 2 out of the 90 instances to optimality, RG can solve 86 instances to optimality, whereas, RG_stren solves all of these instances to optimality. We see from Table 1 that the average solution time for the row generation algorithm increases dramatically as the number of partitions N increases. The required solution times did not change significantly when we consider different sizes of the ambiguity set or different ε values.

The average objective values. Figure 3 presents the average objective values of (BP-RUM) as the number of partitions N varies.

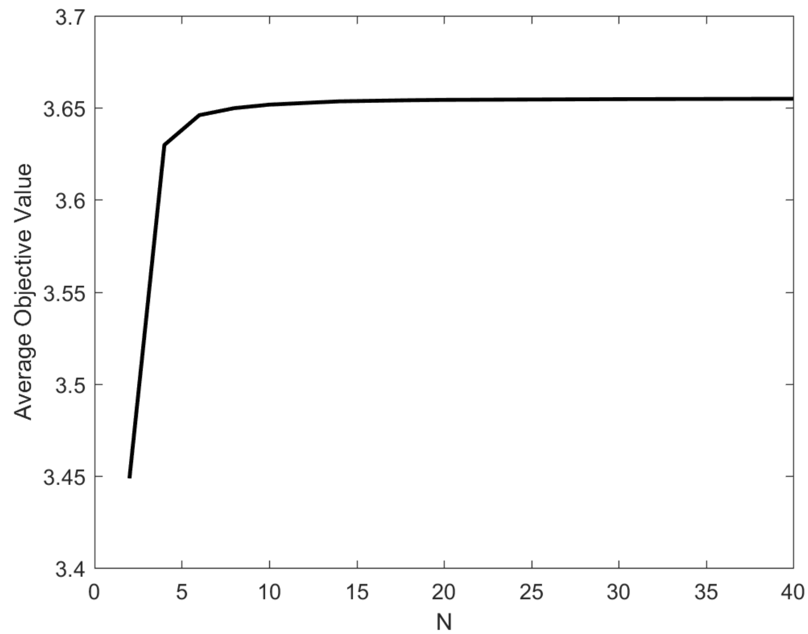

Figure 3 The average objective values of (BP-RUM) as the number of partitions N varies.

From Figure 3 we can see that the average objective value increases significantly when the number of partitions N varies from 2 to 4, and the improvement of the average objective value is not significant when the number of partitions N is larger than 4. Moreover, the average objective value is mainly unchanged after the number of partitions reaches 10. Therefore, in this example, we can use the number of partitions $N = 10$ to obtain approximate solutions of (BP-RUM).

5.1.3. Out-of-Sample Performance We now discuss the solutions in the context of data from the OR planning problem and the out-of-sample performance of the solutions generated from (BP-RUM). For this purpose, we generated 10,000 scenarios from the log-normal distribution. We set the number of partitions $N = 10$, and b is selected using a three-fold cross-validation scheme. Specifically, the training dataset is randomly split into three equal-sized groups. One group is used as the validation set, and the remaining groups are used as the training set. The cross-validation process is repeated three times. At the end of the process, the radius with the maximum average objective value is used. In order to evaluate the influence of the chance constraints, we remove the chance constraints from (BP-RUM) and solve the resulting problem referred to as (BP-RUM-NC). We also share our experience in solving a determinant problem, i.e. $b = 0$, which is referred to as (BP-DUM). Figure 4 presents the average number of opened ORs, the out-of-sample over-utilization probability, over-utilization (hour), and under-utilization (hour) for the calibrated (BP-RUM), (BP-DUM), and (BP-RUM-NC).

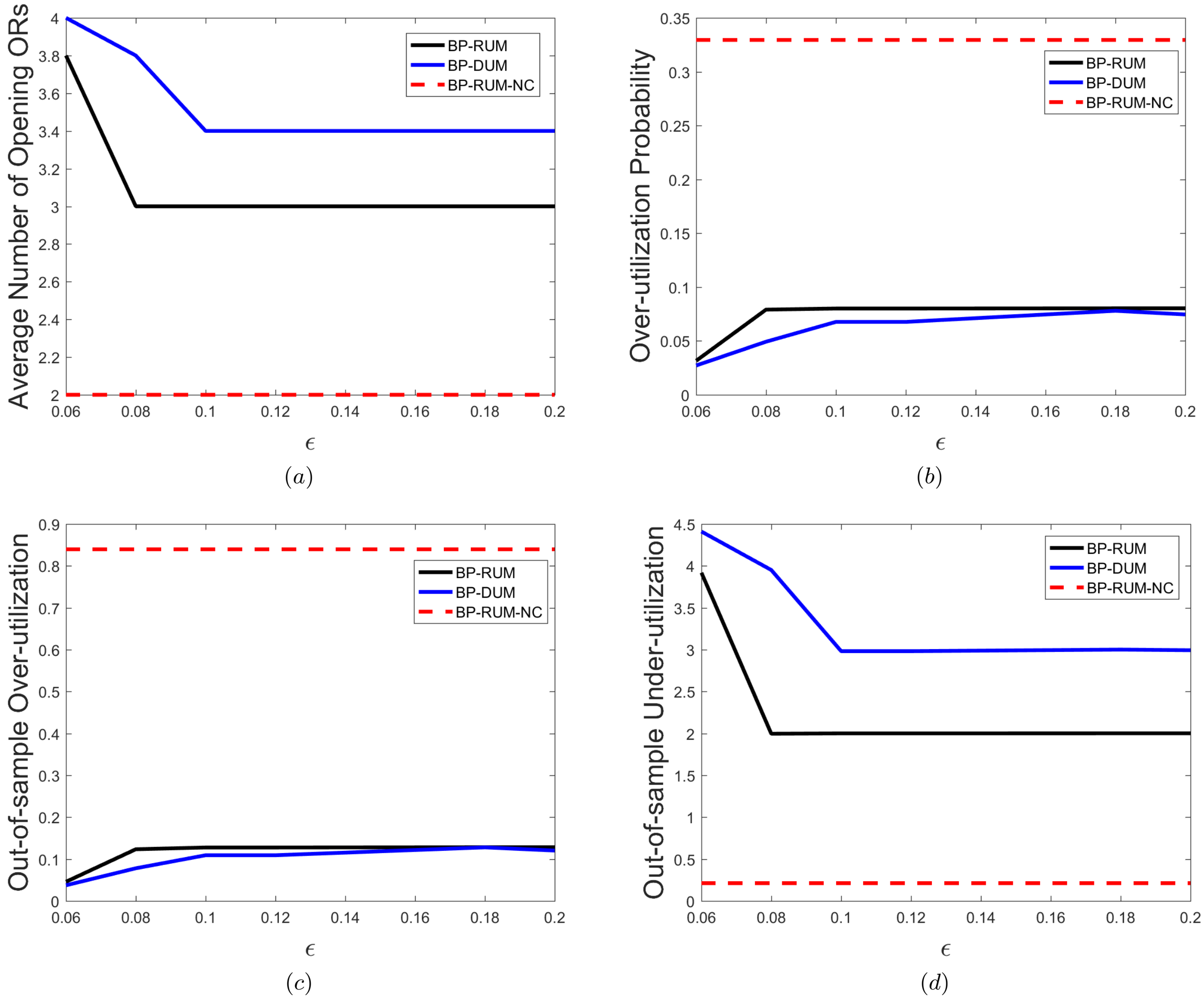

Figure 4 **The average number of opening ORs, the out-of-sample over-utilization probability, over-utilization (hour), and under-utilization (hour) for the calibrated (BP-RUM), (BP-DUM), and (BP-RUM-NC) as ε varies.**

From Figure 4, we observe that for (BP-RUM) and (BP-DUM), the average out-of-sample over-utilization probability is smaller than the predefined ε value, which suggests that these solutions deliver reasonable quality-of-service performance. Moreover, the average under-utilization is significantly larger than the over-utilization. This can be explained by the fact that in order to obtain a desirable over-utilization probability, one needs to open more ORs and as a result, it will have a larger under-utilization. And when ε further increases, the optimal number of opening ORs remains unchanged, thus under and over-utilization are unchanged. The results from Figure 4 also show that compared with (BP-RUM), (BP-DUM) has slightly smaller over-utilization probability values and over-utilization. However, (BP-DUM) has a larger average number of opening ORs, which results in a significantly larger under-utilization. In the meantime, the under-utilization and over-utilization are less than an hour when using the solution obtained from (BP-RUM-NC). Nevertheless, the average over-utilization probability is larger than 0.3. Overall, (BP-RUM) can meet the desired out-of-sample chance satisfaction ε with reasonable over- and under-utilization.

5.2. Multi-Item Newsvendor Robust Expected Utility Problem

We present computational results of (MN-RUM). We follow a similar parameter setting for the multi-item newsvendor problem in recent work by Chen et al. (2020). For each scenario ω and item j, we randomly generate ξ_j^ω from a uniform distribution on $[0, \bar{u}_j]$, where $\bar{u}_j$ is randomly generated from a uniform distribution on $[0, 100]$. The inventory capacity t_j is randomly generated from a uniform distribution on $[10, 20]$. We let $d = 50|\mathcal{J}|$, and $o_j = 1$, for $j \in \mathcal{J}$. The number of items $|\mathcal{J}| = 100$ and the number of scenarios $|\Omega| = 1000$. For other implementation details, we use the same settings as in Section 5.1.1.

5.2.1. Computational Results We first show the performance of the row generation algorithm (Algorithm 1) for solving (MN-RUM), then we discuss the out-of-sample performance for (MN-RUM).

Performance of row generation algorithm. We present the performance of the methods described in Section 5.1.2. Table 2 reports the average CPU solution time for solving (MN-RUM), the subproblem (SP), and the linear relaxation of (MP), and the average number of cuts (18). Since all five generated instances could be solved within the time limit, we ignore the number of instances solved to optimality over the five generated instances in Table 2.

The results from Table 2 further demonstrate the efficiency of the proposed row generation algorithm. RG_stren has a significantly better performance than CPX for the harder instances ($N \in \{10, 20\}$), and slightly outperforms RG for most types of instances. More specifically, for the harder instances, in comparison to CPX, the average solution time is decreased by more than 54% by RG_stren, and about 50% by RG. Similar to Table 1, we also observe from Table 2 that as the number of partitions N increases, the average solution time increases significantly, while when ε varies from 0.1 to 0.2 or b from 1 to 5, the average solution time is not significantly different. In the meantime, the average solution time for CPX increases more dramatically than the time increase in the row generation algorithm as the number of partitions increases. This suggests that compared with CPX, the row generation algorithm has a more stable computational performance.

The average objective values. Figure 5 presents the average objective values of (MN-RUM) as the number of partitions N varies.

Table 2 The average CPU (in seconds) solution time for solving (MN-RUM) (AvT), the subproblem (SP) (AvT-Z), and the linear relaxation of (MP) (AvT-L), and the average number of cuts (cuts) are reported.

b	ε	N	CPX	RG			RG_stren			
			AvT	AvT	AvT-Z	cuts	AvT	AvT-Z	AvT-L	cuts
1	0.1	4	110	145	144	11.4	131	73	56	5.4
		10	453	277	276	8.4	253	125	127	3.4
		20	3,248	547	545	6	574	288	286	3
	0.2	4	103	257	250	20	258	178	67	14.4
		10	799	614	605	19.6	539	379	148	12
		20	3,412	1,344	1,337	17.2	1,068	746	312	10
2	0.1	4	78	124	123	9.2	111	53	56	3.6
		10	851	317	316	9.4	317	173	142	4.4
		20	2,788	489	488	5.8	620	301	318	3.4
	0.2	4	84	296	287	22.2	252	173	68	13.8
		10	598	586	578	19.4	535	382	143	13.2
		20	2,643	1,414	1,406	18.8	1,046	715	321	10
5	0.1	4	89	134	133	10.6	143	80	61	5.4
		10	537	266	265	8.2	219	82	136	2
		20	2,956	624	623	7.4	607	310	296	3.4
	0.2	4	84	278	269	21	299	215	72	16.4
		10	616	572	564	18.4	488	322	156	10.2
		20	3,050	1,179	1,173	16.4	1,321	928	383	11.6

Similar to Figure 5, the average objective value for the instances with $N = 4$ is significantly better than the ones with $N = 2$, and the improvement of the average objective value is not significant when we increase the number of partitions to $N = 10$. Therefore, in this example, $N = 10$ can also be used to obtain approximate solutions with reasonable solution time.

5.2.2. Out-of-Sample Performance. We now discuss the out-of-sample performance of the solutions generated from (MN-RUM) and compare the performance with (MN-RUM) but without chance constraints, which is referred to (MN-RUM-NC), and the determinant problem, which is

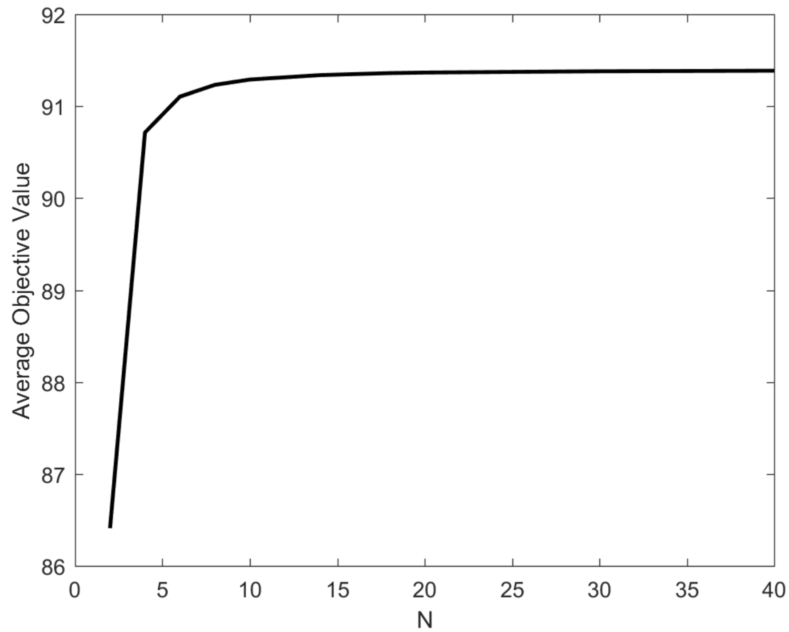

Figure 5 **The average objective value of (MN-RUM) as the number of partitions N varies.**

referred to as (MN-DUM). We generated 10,000 scenarios from the uniform distribution $(0, \bar{u})$. Figure 6 presents the improvement (in %) of the out-of-sample over-stocking probability, over-stocking, and under-stocking for the calibrated (MN-RUM) over (MN-RUM-NC) and (MN-DUM). Since (MN-RUM) has a similar out-of-sample under-stocking to (MN-DUM), thus we did not present it in this section.

Conclusions from the results in Figure 6 are similar to those observed from Figure 4 with some differences. For example, using the solutions obtained from (MN-RUM-NC), the out-of-sample over-stocking probability and over-stocking are significantly larger than the one obtained from (MN-RUM), although the out-of-sample over-stocking probability for (MN-RUM) is slightly larger than the predefined ε value. Whereas, compared with (MN-RUM), the out-of-sample under-stocking is smaller when using the solutions obtained from (MN-RUM-NC). Different from Figure 4, compared with (MN-RUM), (MN-DUM) has slightly larger out-of-sample over-stocking probability and over-stocking. In the meantime, (MN-RUM) and (MN-DUM) have a comparable number of the out-of-sample under-stocking. Hence, the out-of-sample superiority of robust solutions indicates that (MN-RUM) also performs well in the multi-item newsvendor problem.

6. Concluding Remarks

In this work, we study a general robust expected utility maximization problem with chance constraints over a set of concave utility functions that lie in an ambiguity set. Methodologically speaking, we first apply a discrete approximation approach to formulate the ambiguity set $\mathscr{U}$

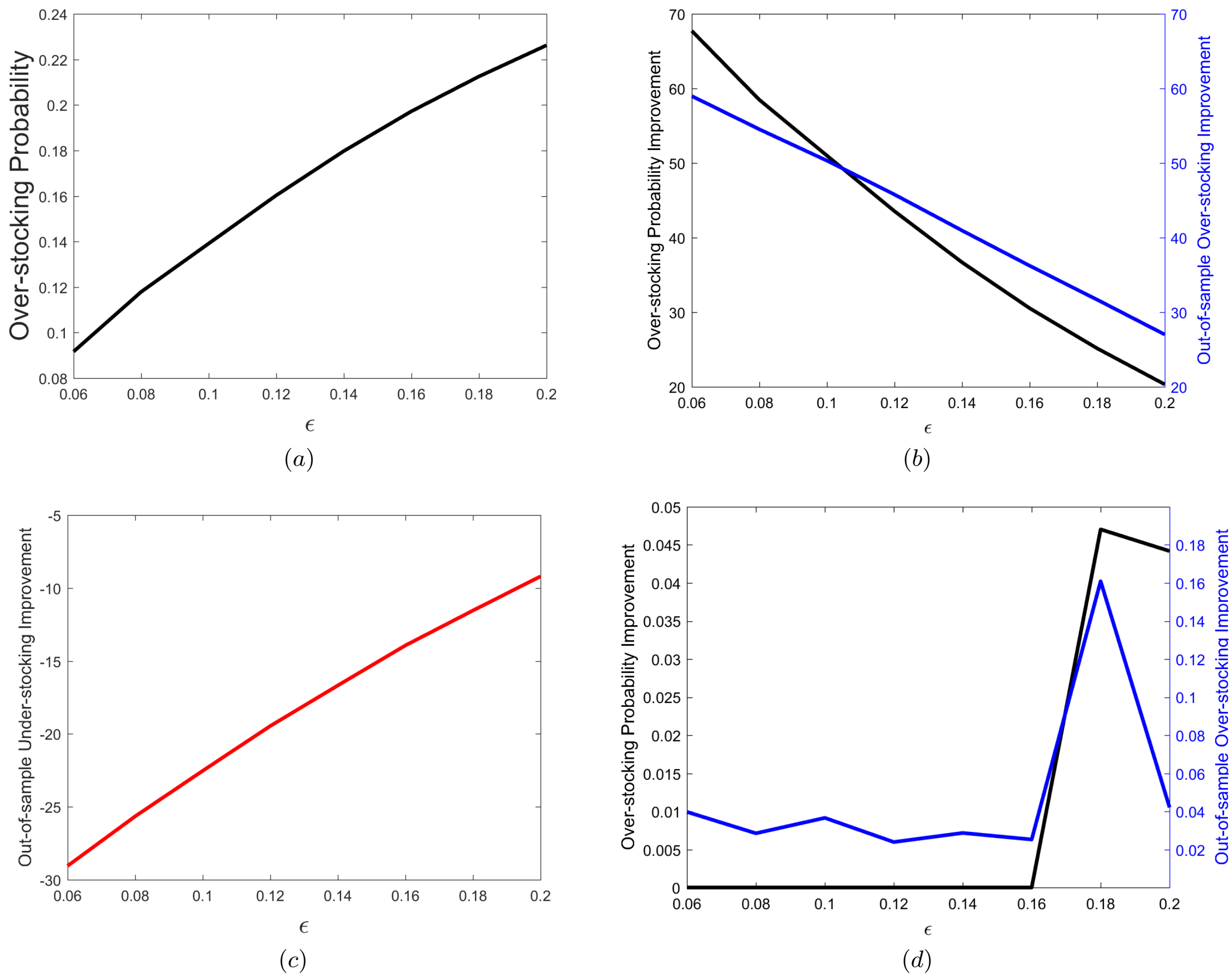

Figure 6 **The out-of-sample over-stocking probability for the calibrated (MN-RUM) (a), the improvement of the out-of-sample over-stocking probability and over-stocking for (MN-RUM-NC) (b), and under-stocking for (MN-RUM-NC) (c), and the improvement of the out-of-sample over-stocking probability and over-stocking for (MN-DUM) (d) as ε varies.**

and reformulate (RUM) as a mixed-integer program. We then conduct a detailed convergence analysis to show that the discrete approximation converges to the true counterpart under some mild assumptions. On the computational side, we propose a row generation-based solution scheme to solve our chance-constrained robust expected utility model efficiently. Finally, from a practical application viewpoint, we perform an extensive numerical study for the bin packing problem and the multi-item newsvendor problem to analyze the general structure of the decisions from the decision-making framework and show the benefits of the techniques developed in this paper for computational improvement. The numerical results show that the row generation algorithm can significantly reduce the computational time for a certain problem size when compared with CPLEX solver, and the solutions that are obtained from (RUM) achieve a desirable out-of-sample performance.

To the best of our knowledge, this is the first attempt to combine the concept of robust decision-making with utility-dependent decisions and chance constraint optimization as a complementary synergistic mechanism for decision modeling under risk and uncertainty, especially when the information of the utility function is incomplete. For future research, we suspect that the modeling framework and resolution methods that are presented in this paper should also benefit several other practical applications of interest, e.g., those in facility location (e.g., Luo and Mehrotra 2024), and cloud computing (e.g., Cohen et al. 2019, Martinovic et al. 2021), etc.

Acknowledgments

This research of the first author and third author was partially supported by the National Natural Science Foundation of China (NSFC) grants 72272014, 72371022.

References

Alvarez-Mena J, Hernández-Lerma O (2005) Convergence and approximation of optimization problems. *SIAM Journal on Optimization* 15(2):527–539.

Ardestani-Jaafari A, Delage E (2016) Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research* 64(2):474–494.

Armbruster B, Delage E (2015) Decision making under uncertainty when preference information is incomplete. *Management Science* 61(1):111–128.

Azeem SM, Akhtar N (2014) The influence of work life balance and job satisfaction on organizational commitment of healthcare employees. *International Journal of Human Resource Studies* 4(2):18.

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

Bertsimas D, Sim M (2004) The price of robustness. *Operations Research* 52(1):35–53.

Bodur M, Luedtke JR (2017) Mixed-integer rounding enhanced benders decomposition for multiclass service-system staffing and scheduling with arrival rate uncertainty. *Management Science* 63(7):2073–2091.

Boyd S, Boyd SP, Vandenberghe L (2004) *Convex optimization* (Cambridge university press).

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

Buck RC (1943) A note on subsequences. *Bulletin of the American Mathematical Society* 49(12):898–899.

Chajewska U, Koller D, Parr R (2000) Making rational decisions using adaptive utility elicitation. *Aaai/Iaai*, 363–369.

Charnes A, Cooper WW (1959) Chance-constrained programming. *Management Science* 6(1):73–79.

Chen Z, Sim M, Xiong P (2020) Robust stochastic optimization made easy with rsome. *Management Science* 66(8):3329–3339.

Choi S, Ruszczyński A (2011) A multi-product risk-averse newsvendor with exponential utility function. *European Journal of Operational Research* 214(1):78–84.

Cohen MC, Keller PW, Mirrokni V, Zadimoghaddam M (2019) Overcommitment in cloud services: Bin packing with chance constraints. *Management Science* 65(7):3255–3271.

Davis A, Mehrotra S, Holl J, Daskin MS (2014) Nurse staffing under demand uncertainty to reduce costs and enhance patient safety. *Asia-Pacific Journal of Operational Research* 31(01):1450005.

Delage E, Guo S, Xu H (2022) Shortfall risk models when information on loss function is incomplete. *Operations Research* 70(6):3511–3518.

Deng Y, Shen S (2016) Decomposition algorithms for optimizing multi-server appointment scheduling with chance constraints. *Mathematical Programming* 157(1):245–276.

Deng Y, Shen S, Denton B (2019) Chance-constrained surgery planning under conditions of limited and ambiguous data. *INFORMS Journal on Computing* 31(3):559–575.

Erlebacher SJ (2000) Optimal and heuristic solutions for the multi-item newsvendor problem with a single capacity constraint. *Production and Operations Management* 9(3):303–318.

Farquhar PH (1984) State of the art—utility assessment methods. *Management Science* 30(11):1283–1300.

Gabrel V, Murat C, Thiele A (2014) Recent advances in robust optimization: An overview. *European Journal of Operational Research* 235(3):471–483.

Gorissen BL, Yanıkoğlu İ, den Hertog D (2015) A practical guide to robust optimization. *Omega* 53:124–137.

Green J, Valentine F (1961) On the arzela-ascoli theorem. *Mathematics Magazine* 34(4):199–202.

Guest DE (2002) Perspectives on the study of work-life balance. *Social Science Information* 41(2):255–279.

Hanasusanto GA, Kuhn D, Wallace SW, Zymler S (2015) Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming* 152(1):1–32.

Hansen L, Sargent TJ (2001) Robust control and model uncertainty. *American Economic Review* 91(2):60–66.

Haskell WB, Fu L, Dessouky M (2016) Ambiguity in risk preferences in robust stochastic optimization. *European Journal of Operational Research* 254(1):214–225.

Holt CA, Laury SK (2002) Risk aversion and incentive effects. *American Economic Review* 92(5):1644–1655.

Hoogervorst R, Zhang Y, Tillem G, Erkin Z, Verwer S (2019) Solving bin-packing problems under privacy preservation: Possibilities and trade-offs. *Information Sciences* 500:203–216.

Hu J, Bansal M, Mehrotra S (2018) Robust decision making using a general utility set. *European Journal of Operational Research* 269(2):699–714.

Hu J, Li J, Mehrotra S (2019) A data-driven functionally robust approach for simultaneous pricing and order quantity decisions with unknown demand function. *Operations Research* 67(6):1564–1585.

Hu J, Mehrotra S (2015) Robust decision making over a set of random targets or risk-averse utilities with an application to portfolio optimization. *IIE Transactions* 47(4):358–372.

Kelley JE Jr (1960) The cutting-plane method for solving convex programs. *Journal of the society for Industrial and Applied Mathematics* 8(4):703–712.

Kim K, Mehrotra S (2015) A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research* 63(6):1431–1451.

Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12(2):479–502.

Küçükyavuz S, Jiang R (2022) Chance-constrained optimization under limited distributional information: A review of reformulations based on sampling and distributional robustness. *EURO Journal on Computational Optimization* 10:100030.

Luo F, Mehrotra S (2024) Service center location problems with decision dependent utilities and a pandemic case study. *Naval Research Logistics (NRL)* 71(1):3–26.

Lytle B (2015) Introduction to the convergence of sequences. *University of Chicago* .

Martinovic J, Hähnel M, Scheithauer G, Dargie W (2021) An introduction to stochastic bin packing-based server consolidation with conflicts. *TOP* 1–36.

Martín-Baos J, García-Ródenas R, Rodriguez-Benitez L (2021) Revisiting kernel logistic regression under the random utility models perspective. an interpretable machine-learning approach. *Transportation Letters The International Journal of Transportation Research* 13(3):1–12.

Natarajan K, Sim M, Uichanco J (2010) Tractable robust expected utility and risk models for portfolio optimization. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 20(4):695–731.

Natarajan K, Sim M, Uichanco J (2018) Asymmetry and ambiguity in newsvendor models. *Management Science* 64(7):3146–3167.

Qin Y, Wang R, Vakharia AJ, Chen Y, Seref MM (2011) The newsvendor problem: Review and directions for future research. *European Journal of Operational Research* 213(2):361–374.

Rahimian H, Bayraksan G, Homem-de Mello T (2019) Controlling risk and demand ambiguity in newsvendor models. *European Journal of Operational Research* 279(3):854–868.

Rahimian H, Mehrotra S (2019) Distributionally robust optimization: A review. *working paper, http://www.optimization-online.org/DB_FILE/2019/08/7332.pdf* .

Schied A (2005) Optimal investments for robust utility functionals in complete market models. *Mathematics of Operations Research* 30(3):750–764.

Shapiro A, Dentcheva D, Ruszczynski A (2009) Lectures on stochastic programming: modeling and theory. *(SIAM) MPS, Philadelphia* .

Song Y, Luedtke JR, Küçükyavuz S (2014) Chance-constrained binary packing problems. *INFORMS Journal on Computing* 26(4):735–747.

Train K, Weeks M (2005) *Discrete Choice Models in Preference Space and Willingness-to-Pay Space*, 1–16 (Springer Netherlands).

Wang CX, Webster S, Zhang S (2012) Newsvendor models with alternative risk preferences within expected utility theory and prospect theory frameworks. *Handbook of Newsvendor Problems*, 177–196 (Springer).

Wang S, Delage E (2024) A column generation scheme for distributionally robust multi-item newsvendor problems. *INFORMS Journal on Computing*. https://doi.org/10.1287/ijoc.2022.0010.

Wang S, Li J, Mehrotra S (2021) Chance-constrained multiple bin packing problem with an application to operating room planning. *INFORMS Journal on Computing* 33(4):1661–1677.

Wang S, Li J, Mehrotra S (2022) A solution approach to distributionally robust joint-chance-constrained assignment problems. *INFORMS Journal on Optimization* 4(2):125–147.

Wang S, Wang Q, Zhao J (2020) Deep neural networks for choice analysis: Extracting complete economic information for interpretation. *Transportation Research Part C Emerging Technologies* 118:102701.

Zhang J, Xie W, Sarin SC (2021) Robust multi-product newsvendor model with uncertain demand and substitution. *European Journal of Operational Research* 293(1):190–202.

Zhang Z, Denton BT, Xie X (2020) Branch and price for chance-constrained bin packing. *INFORMS Journal on Computing* 32(3):547–564.

Appendix A: Proof of Lemmas and Theorems

A.1. Proof of Lemma 1

Let $\theta_k = u(a_k)$, for $k = 0, \cdots, N$, then $\theta_0 = u(a_0) = 0$, $\theta_l = u(a_l) = 1$ and $\theta_N = u(a_N) = 0$. Given the concavity property of $u \in \mathscr{U}_N$, we have

$$\frac{\theta_1 - \theta_0}{a_1 - a_0} \geq \cdots \geq \frac{\theta_l - \theta_{l-1}}{a_l - a_{l-1}} \geq 0 \geq \frac{\theta_{l+1} - \theta_l}{a_{l+1} - a_l} \geq \cdots \geq \frac{\theta_N - \theta_{N-1}}{a_N - a_{N-1}},$$

which implies that constraint (11b) holds. In addition, constraints (11c), (11d) and (11e) represent the constraints of $u(a)$.

Let $\mathscr{U}'_N$ be a subset of $\mathscr{U}_N$ which consists of all the piecewise linear functions with break points $\{a_0, \cdots, a_N\}$, and

$$\pi'_N(\boldsymbol{x}) = \underset{u \in \mathscr{U}'_N}{\text{minimize}} \sum_{j \in \mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))]. \tag{19}$$

If u^* is an optimal solution of problem (10), we can define a piecewise linear function belonging to $\mathscr{U}'_N$ that bounds u^* from below. Hence, we can rewrite problem (10) as problem (19). Since $u \in \mathscr{U}'_N$ is piecewise linear with the break point a_k and corresponding value θ_k, thus given a $v \in \Theta$, $u(v)$ is equivalent to

$$\begin{aligned} \underset{\alpha,\beta}{\text{minimize}} \;& v\alpha + \beta \\ \text{subject to } \;& a_k\alpha + \beta - \theta_k \geq 0 && \forall k \in \{0, \cdots, N\}, \\ & \beta \geq 0. \end{aligned}$$

Therefore, when $v = f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega)$ for $j \in \mathcal{J}, \omega \in \Omega$, problem (10) is equivalent to problem (11). □

A.2. Proof of Theorem 1

Let $\boldsymbol{y} = \boldsymbol{Q}(\boldsymbol{\theta} - \boldsymbol{u}_0)$, then $||\boldsymbol{y}||_p \leq 1$ based on constraint (11e) and $\boldsymbol{\theta} = \boldsymbol{Q}^{-1}\boldsymbol{y} + \boldsymbol{u}_0$. Let $\boldsymbol{\mu}$, $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, $\boldsymbol{\delta}$ be the dual variables of constraints (11b) to (11d), and (11f) respectively. The dual function can be formulated as

$$\begin{aligned} g(\boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta}) \; = \; & \inf_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{y}} \; L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{y}, \boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta}) \\ & \text{subject to } ||\boldsymbol{y}||_p \leq 1, \\ & \qquad \beta^{j\omega} \geq 0, && \forall j \in \mathcal{J}, \omega \in \Omega, \end{aligned}$$

where

$$\begin{aligned} & L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{y}, \boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta}) \\ & = \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} [(p_\omega f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) - \sum_{k=0}^{N} a_k \delta_{kj\omega})\alpha^{j\omega} + (p_\omega - \sum_{k=0}^{N} \delta_{kj\omega})\beta^{j\omega}] + \sum_{k=0}^{N} (\underline{u}(a_k)\gamma_{2k} - \bar{u}(a_k)\gamma_{1k}) + \boldsymbol{\eta}^\top(\boldsymbol{Q}^{-1}\boldsymbol{y} + \boldsymbol{u}_0). \end{aligned}$$

and $\eta_k = -\mu_k(a_{k+1} - a_{k-1}) + \mu_{k+1}(a_{k+2} - a_{k+1}) + \mu_{k-1}(a_{k-1} - a_{k-2}) + \gamma_{1k} - \gamma_{2k} + \sum_{j \in \mathcal{J}} \sum_{\omega \in \Omega} \delta_{kj\omega}$, for all $k = 1, \cdots, N-1$.

Based on the domain of variables $\boldsymbol{y}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, and the definition of dual norm, we have,

$$g(\boldsymbol{\mu}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\delta}_2) = \sum_{k=0}^{N} (\underline{u}(a_k)\gamma_{2k} - \bar{u}(a_k)\gamma_{1k}) + \sum_{k=0}^{N-1} \eta_k u_0(a_k) - ||\boldsymbol{Q}^{-1}\boldsymbol{\eta}||_*,$$

if $p_\omega f_j(\boldsymbol{x}, \boldsymbol{\xi}^\omega) - \sum_{k=0}^{N} a_k \delta_{kj\omega} = 0, \;\; p_\omega - \sum_{k=0}^{N} \delta_{kj\omega} \geq 0$, for $\omega \in \Omega$. Thus, constraints (14b)-(14d) are the dual formulation of problem (11). Note that when $\theta_k = u_0(a_k)$ for all $k = 0, \cdots, N$, constraint (11e) can be reformulated as $b \geq 0$. Since b is a positive constant, $u_0 \in$ **relint** $\mathscr{U}_N$ such that constraint (11e) hold with strict inequality, and problem (11) satisfies Slater's condition (Boyd et al. 2004). Strong duality holds under Slater's conditions. □

A.3. Proof of Lemma 2

Based on Assumption 1, we know that u is concave and $f_j(\cdot)$ is linear in an open neighborhood of $\mathcal{X}$ which can be denoted by $\mathcal{N}(\mathcal{X})$. Therefore, $u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))$ is concave in $\mathcal{N}(\mathcal{X})$ and thus $\mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))]$ is concave in $\mathcal{N}(\mathcal{X})$. Moreover, for $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathcal{X}$, we have

$$\begin{aligned}\pi(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2) &= \underset{u\in\mathscr{U}}{\text{minimize}}\sum_{j\in\mathcal{J}}\mathbb{E}[u(f_j(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2,\tilde{\boldsymbol{\xi}}))]\\ &\geq \underset{u\in\mathscr{U}}{\text{minimize}}\,\lambda\sum_{j\in\mathcal{J}}\mathbb{E}[u(f_j(\boldsymbol{x}_1,\tilde{\boldsymbol{\xi}}))]+(1-\lambda)\sum_{j\in\mathcal{J}}\mathbb{E}[u(f_j(\boldsymbol{x}_2,\tilde{\boldsymbol{\xi}}))]\\ &\geq \lambda\,\underset{u\in\mathscr{U}}{\text{minimize}}\sum_{j\in\mathcal{J}}\mathbb{E}[u(f_j(\boldsymbol{x}_1,\tilde{\boldsymbol{\xi}}))]+(1-\lambda)\,\underset{u\in\mathscr{U}}{\text{minimize}}\sum_{j\in\mathcal{J}}\mathbb{E}[u(f_j(\boldsymbol{x}_2,\tilde{\boldsymbol{\xi}}))]\\ &= \lambda\pi(\boldsymbol{x}_1)+(1-\lambda)\pi(\boldsymbol{x}_2).\end{aligned}$$

$\pi(\boldsymbol{x})$ is concave function in $\mathcal{N}(\mathcal{X})$, thus $\pi(\boldsymbol{x})$ is continuous in $\mathcal{X}$. Similarly, $\pi_N(\boldsymbol{x})$ is continuous concave functions in $\mathcal{X}$. This completes our proof. □

A.4. Proof of Lemma 3

For $u\in\mathscr{U}_N$, since the derivative u' of u satisfies $|u'|<M$ for almost everywhere, then u is Lipschitz with Lipschitz constant M. Since a set of functions with bounded Lipschitz constant forms an equicontinuous set, u is equicontinuous. □

A.5. Proof of Lemma 5

Based on Lemma 3 and Lemma 4, we have that, for any sequence $\{u_N\in\mathscr{U}_N\}$, there exists a subsequence $\{u_{N_K}\}$ that uniformly converges to $\hat{u}$. Now we show that $\hat{u}\in\mathscr{U}$. Since $\{u_{N_K}\}$ is first increasing then decreasing concave function, for any $-\theta_1\leq a_1\leq a_2\leq 0$, we have

$$\hat{u}(a_1)=\lim_{K\to\infty}u_{N_K}(a_1)\leq\lim_{K\to\infty}u_{N_K}(a_2)=\hat{u}(a_2).$$

Therefore, $\hat{u}$ is increasing on $[-\theta_1,0]$. Similarly, $\hat{u}$ is decreasing on $[0,\theta_2]$. For $\lambda>0$ and $-\theta_1\leq a_1\leq a_2\leq\theta_2$,

$$\begin{aligned}\hat{u}(\lambda a_1+(1-\lambda)a_2) &= \lim_{K\to\infty}u_{N_K}(\lambda a_1+(1-\lambda)a_2)\\ &\geq \lim_{K\to\infty}\lambda u_{N_K}(a_1)+(1-\lambda)u_{N_K}(a_2)\\ &= \lambda\hat{u}(a_1)+(1-\lambda)\hat{u}(a_2).\end{aligned}$$

Hence, $\hat{u}$ is first increasing then decreasing concave function. We then consider the bound constraints. Since $\underline{u}_{N_K}$ and $\bar{u}_{N_K}$ uniformly converge to $\underline{u}$ and $\bar{u}$, respectively; we have

$$\hat{u}=\lim_{K\to\infty}u_{N_K}\geq\lim_{K\to\infty}\underline{u}_{N_K}=\underline{u},\quad \hat{u}=\lim_{K\to\infty}u_{N_K}\leq\lim_{K\to\infty}\bar{u}_{N_K}=\bar{u}.$$

We now claim the auxiliary constraint. Let

$$h(u_{N_K})=\sum_{k=0}^{N_K-1}\left(u_0(a_k)-u_{N_K}(a_k)\right)^p(a_{k+1}-a_k)-b,$$

and

$$f(u_{N_K})=\int_{-\theta_1}^{\theta_2}\left(u_0(a)-u_{N_k}(a)\right)^p da-b, \tag{20}$$

hence

$$\lim_{K\to\infty} h(u_{N_K}) = \lim_{K\to\infty} f(u_{N_K}).$$

Given the uniform convergence of $\{u_{N_K}\}$ to $\hat{u}$, for any $\delta > 0$, there exists $\hat{K}$ such that for all $K \geq \hat{K}$:

$$|u_{N_K}(a) - \hat{u}(a)| \leq \delta.$$

Thus we have

$$\begin{aligned}
\big|f(u_{N_K}) - f(\hat{u})\big| &= \Big|\int_{-\theta_1}^{\theta_2} (u_0(a) - u_{N_k}(a))^p - (u_0(a) - \hat{u}(a))^p \, da\Big| \\
&= \Big|\int_{-\theta_1}^{\theta_2} (\hat{u}(a) - u_{N_k}(a)) \left((u_0(a) - u_{N_k}(a))^{p-1} + \cdots + (u_0(a) - \hat{u}(a))^{p-1}\right) da\Big| \\
&\leq \delta \int_{-\theta}^{\theta_2} \big|(u_0(a) - u_{N_k}(a))^{p-1} + \cdots + (u_0(a) - \hat{u}(a))^{p-1}\big| da
\end{aligned}$$

Since $(u_0(a) - u_{N_k}(a))^{p-1} + \cdots + (u_0(a) - \hat{u}(a))^{p-1}$ is bounded, $f(\hat{u}) = \lim_{k\to\infty} f(u_{N_k})$. It follows that,

$$f(\hat{u}) = \lim_{k\to\infty} f(u_{N_k}) = \lim_{k\to\infty} h_{N_k}(u_{N_k}) \leq 0.$$

Therefore, we have $\hat{u} \in \mathscr{U}$. □

A.6. Proof of Lemma 6

For any $u \in \mathscr{U}$, since $(u_0(a) - u(a))^p$ is uniformly continuous on $[-\theta_1, \theta_2]$, $h(u)$ converges uniformly to $f(u)$. Hence, for $\delta > 0$ and $\delta < \tau/2$, there exists $\hat{N}$ such that for all $N \geq \hat{N}$:

$$|h(u) - f(u)| \leq \delta,$$

then we have $h(u) \leq f(u) + \delta \leq \delta$, which gives us that

$$\sum_{k=0}^{N-1} (u_0(a_k) - u(a_k))^p \, (a_{k+1} - a_k) \leq b + \delta. \tag{21}$$

For any $\lambda \in [0,1]$, by constraint (21), it follows that

$$(1-\lambda)^p \sum_{k=0}^{N-1} (u_0(a_k) - u(a_k))^p \, (a_{k+1} - a_k) \leq (1-\lambda)^p (b+\delta) \leq (1-\lambda)(b+\delta). \tag{22}$$

Similarly, we have

$$0 \leq \lambda(b + \delta - \tau), \tag{23}$$

for all $\tau \in \mathbb{R}^+$ and $\tau \leq b$. From constraints (22) and (23), we can obtain

$$\sum_{k=0}^{N-1} [u_0(a_k) - ((1-\lambda)u(a_k) + \lambda u_0(a_k))]^p \, (a_{k+1} - a_k) \leq b + \delta - \lambda\tau$$

We let $v_\lambda = (1-\lambda)u + \lambda u_0$, thus,

$$\sum_{k=0}^{N} (u_0(a_k) - v_\lambda(a_k))^p \leq b, \quad \forall \lambda \in [\delta/\tau, 1].$$

Hence, we have for each $\lambda \in [\delta/\tau, 1]$, $v_\lambda \in \mathscr{U}_N$ for all $N \geq \hat{N}$.

To construct the sequence u_N, we define a positive sequence $\{\delta_i\}$ such that $\delta_i \to 0$ as $i \to \infty$. Let $v_i = (1 - \frac{\delta_i}{\tau})u + \frac{\delta_i}{\tau}u_0$. Based on the above discussion, for each δ_i, there exists a positive integer number N_i such that $v_i \in \mathscr{U}_N$ for $N \geq N_i$, and $v_i \to u$ as $i \to \infty$. Let $u_N = v_i$ for $N_i \leq N < N_{i+1}$. Therefore, we have $u_N \in \mathscr{U}_N$ for all $N \geq N_1$ and $u_N \to u$ as $N \to \infty$. □

A.7. Proof of Lemma 8

We let sequence $\{u_N\}$ be the optimal solution of $\pi_N(\boldsymbol{x})$ for given $\boldsymbol{x} \in \mathcal{X}$. By Lemma 5 we know that there exists a subsequence $\{u_{N_k}\}$ of $\{u_N\}$ converges to $u \in \mathscr{U}$. Since the limit of sequence $\{u_{N_k}\}$ in $\mathscr{U}$, hence, $\liminf_{k\to\infty} \pi_{N_k}(\boldsymbol{x}) \geq \pi(\boldsymbol{x})$. Moreover, Lemma 6 shows that for any $u \in \mathscr{U}$, there exists a sequence $u_N \in \mathscr{U}_N$ such that $u = \lim_{N\to\infty} u_N$. Since $\sum_{j\in\mathcal{J}} \mathbb{E}[u_N(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))] = \sum_{j\in\mathcal{J}}\sum_{\omega\in\Omega} p_\omega u_N(f_j(\boldsymbol{x},\boldsymbol{\xi}^\omega))$ and $\sum_{j\in\mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))] = \sum_{j\in\mathcal{J}}\sum_{\omega\in\Omega} p_\omega u(f_j(\boldsymbol{x},\boldsymbol{\xi}^\omega))$, we have $\sum_{j\in\mathcal{J}} \mathbb{E}[u(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))] = \lim_{N\to\infty}\sum_{j\in\mathcal{J}} \mathbb{E}[u_N(f_j(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))]$ (Lytle 2015). Therefore, $\pi_{N_k}(\boldsymbol{x}) \to \pi(\boldsymbol{x})$ as $K \to \infty$ by Lemma 7. To prove $\pi_N(\boldsymbol{x}) \to \pi(\boldsymbol{x})$, let $\{\pi_m(\boldsymbol{x})\}$ be a subsequence of $\pi_N(\boldsymbol{x})$. By Lemma 5, there exists a subsequence $\{u_{m_i}\}$ of $\{u_m\}$ such that u_{m_i} converges to u, which with Lemma 7 implies that $\pi_{m_i}(\boldsymbol{x})$ converges to $\pi(\boldsymbol{x})$. Since $\{\pi_m(\boldsymbol{x})\}$ is an arbitrary subsequence of $\pi_N(\boldsymbol{x})$, thus, $\pi_N(\boldsymbol{x}) \to \pi(\boldsymbol{x})$ as $N \to \infty$ (Buck 1943). □

A.8. Proof of Theorem 2

By Lemma 2, we have that $\mathcal{X}$ is compact and $\pi_N(\boldsymbol{x})$ and $\pi(\boldsymbol{x})$ are continuous, then $\{\pi_N\}$ uniformly converges to π (see Hu and Mehrotra 2015). Since $\mathscr{X}$ and $\mathscr{T}$ are non-empty compact sets, and the function $\pi(\boldsymbol{x})$ is continuous on $\mathscr{X}$ and $\mathscr{T}$, then $y_N \to y^*$ and $D(Z_N, Z^*) := \underset{b_1\in Z_N}{\text{maximize}}\ \underset{b_2\in Z^*}{\text{minimize}}\ ||b_1 - b_2|| \to 0$ as $N \to \infty$, based on Lemma 9. □