Preference Ambiguity Averse Decision Making Using Robust Optimization and Sensitivity Analysis

by

Gevorg Stepanyan

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Systems Engineering)
in the University of Michigan-Dearborn
2021**

Doctoral Commitee:

**Associate Professor Jian Hu, Chair
Assistant Professor Xi Chen
Assistant Professor Ruiwei Jiang, University of Michigan
Assistant Professor David Kaufman
Professor Armen Zakarian**

Table of Contents

List of Figures

List of Tables

Abstract

In this work, we study decision making with personalized stochastic optimization models. The methods, we propose, develop custom-tailored stochastic optimization models for a specific decision maker, while preserving the robustness of an optimal decision as expressions of the decision maker's attitude towards ambiguity.

We present an optimization model using a novel robust preference relationship — reference-based almost stochastic dominance (RSD). We use decision maker's utility function as a reference to individualize constraints of stochastic dominance. The concept of RSD addresses the two problems in utility-based decision making: (i) ambiguity and inaccuracy in characterizing the decision maker's individual risk attitude, (ii) over-conservativeness of stochastic dominance representing general properties of risk aversion. The RSD rule reveals the maximum dominance level quantifying the robustness of the decision maker's preference between alternative choices. We develop an approximation model using Bernstein polynomials, show the asymptotic convergence of its optimal value and set of optimal solutions to the true counterparts as the degree of Bernstein polynomials increases, and analyze the convergence rate of its feasible region. We next develop a cut-generation algorithm to solve the approximation model. Finally, we further adapt this cut-generation algorithm to seek a valid option most robustly preferable to a random benchmark. The effectiveness and computational complexity of the model are illustrated using a portfolio optimization problem.

We study the sensitivity of the personalized stochastic optimization models with regards to risk entangled with the decision maker's ambiguous preference itself. We present a bi-objective stochastic optimization model —expected utility and sensitivity-averse maximization (ESM), incorporating classical risk-aversion and sensitivity analysis with regards to decision maker's preference. Unlike classical sensitivity analysis approaches which are

post-analyses after optimization, ESM incorporates sensitivity analysis in the optimization procedure in terms of the second objective function. It thus allows to produce solutions which are both risk-averse in the classical sense and sensitivity-averse with regards to ambiguity in the decision maker's preference.

ESM adapts the sensitivity measure (SMU) from the general Bayesian sensitivity analysis to build connection between classical expected utility maximization and the sensitivity-aversion. We develop two solution methods ofESM. A mixed-integer reformulation is given for a preference maximizer decision maker, while a linear programming reformulation for a risk-averse decision maker. The effect of ESM is illustrated using a homeland security budget allocation problem.

Chapter 1

Introduction

Optimization models are common tools used in normative decision theory, which allow to find the option at least as ”good” as all the others. Applications of such models can be found in almost any area of industry, including finance (Benidis et al., 2018; Bruni et al., 2014; Cornuejols & Tutuncu, 2006; Hu et al., 2011; Roman et al., 2013), transportation (Li et al., 2014; Nie et al., 2011), response systems (Noyan, 2010), energy management (Gollmer et al., 2008) and others. Classically the ”goodness” of the decision is measured by its utility from the perspective of given decision maker. This kind of measurement of the goodness of a decision is commonly refereed as utility theory. It was proposed by (Bernoulli, 1954) and rigorously defined by (von Neumann & Morgenstern, 2007) (see Aase, 2001, for a comprehencive historical review).In (von Neumann & Morgenstern, 2007) does provide the axiomatic foundations of utility theory, proving the existence of a unique function (named *utility function*) mapping outcomes to decision maker's utilities under the following axioms-

(3:A) Axioms of rationality of the decision maker

   (a) Completeness axiom - for any two outcomes X and Y decision maker either prefers X to Y, or Y to X, or he is indifferent between X and Y.

   (b) Transitivity axiom - if a decision maker prefers outcome X to Y, and Y to Z, then he must prefer the outcome X to Z.

(3:B) Continuity axiom - if a decision maker prefers outcome X to Y, and Y to Z, then there exists a probability $p \in [0, 1]$, for which the decision maker becomes indifferent between $pX + (1 - p)Z$ and Y.

(3:C) Independence axiom - if a decision maker prefers outcome X to Y, then for any probability $p \in [0,1]$ and outcome Z, the decision maker still prefers $pX + (1-p)Z$ to $pY + (1-p)Z$.

Thus, if the decision maker does satisfy axioms (3:A)-(3:C), in order to solve classical utility based decision making problem, one only need to assess the decision maker's utility function and use it in the optimization model. A large body of research has addressed utility assessment methods, ranging from preference and/or probability comparison to statistical regression (Farquhar, 1984; Wakker & Deneffe, 1996, see). Initially it was thought that it is not possible to precisely identify decision makers utility function, because of technical difficulties related to assessment methods. However, over the years a large body of studies from psychology and descriptive decision making indicate that (3:A)-(3:C) axioms rarely describe real decision makers' preferences (see Agranov & Ortoleva, 2017; Samson, 1988; Thurstone, 1994, and others in references). Based on the results of these studies we identify following challenges in decision making when working with decision maker's utility functions (These challenges were also indicated in Samson, 1988):

(Chl 1) Assessed utility function may be imprecise, because

- (a) all assessment methods can construct it using only *a finite amount* of questionnaire answers of decision maker (Farquhar, 1984; Wakker & Deneffe, 1996),
- (b) decision makers tend to simplify complex questions, resulting in answers uncharacteristic to them (H. Simon, 1955; Tversky, 1969; Tversky & Kahneman, 1981).

(Chl 2) It may not be possible to construct one function to describe decision maker's preference at all, because decision maker's preferences may be mutable and ambivalent (Agranov & Ortoleva, 2017; Dean & Martin, 2016; Echenique et al., 2011).

Ambiguities in the decision maker's utility function arising from the problem (Chl 1.a) are well known, but are considered relatively easy to solve when detected. Usually by increasing the number (to solve problem (Chl 1.a)) and reducing complexity (to solve problem (Chl 1.b)) of questions in the questionnaires designed for the utility function assessment.

On the other hand, problem (Chl 2) (and in some sense (Chl 1.b)) is risen because of decision maker himself. These problems usually are detected when the decision maker is tested to satisfy the (3:A:a) completeness and (3:A:b) transitivity axioms. These axioms require for the decision maker to present consistent choice behavior. Unfortunately this assumptions have a lot of contradictions with reality. Starting from mid 1950s it was illustrated that under complex, uncertain environment decision makers violate this choice consistency assumption (H. A. Simon, 1956; Tversky, 1969; Tversky & Kahneman, 1981). This studies lead to assumption that decision makers represent such inconsistent (heterogeneous) behavior due to limitations in their mental capacity and that in general a rational decision makers still shall have consistent preferences.

However, starting from mid 2000s, this assumption was shown to be wrong as well. The subjects were shown to change their minds in simple choice problems (Agranov & Ortoleva, 2017; Dean & Martin, 2016; Echenique et al., 2011). In their study of preference consistency Agranov & Ortoleva (2017) indicate that when given simple questions 3 times in a row 23% of participants had switching answers. Additionally they show that 94% of all participants have preference inconsistencies somewhere in their choices, indicating that the inconsistencies in decision maker's preference cannot be disregarded as white noise and must be considered when using utility functions to describe a real decision makers.

Thus, we conclude that in majority of utilitarian decision making frameworks challenges (Chl 1) and (Chl 2) can induce ambiguity around the assessed utility function of the decision maker. In this work we define this ambiguity as follows-

Preference ambiguity is the inevitable imprecision in decision maker's utility function, caused by challenges (Chl 1) and (Chl 2).

It is obvious that the results of optimal decision making model can be sensitive to the preference ambiguity. Studies identifying it can be found in market analysis (Dean & Martin, 2016; Echenique et al., 2011), in health care (Chapman & Elstein, 2013; Stalmeier et al., 2001, 1997), in portfolio optimization (Agranov & Ortoleva, 2017; Echenique et al., 2011; Hu et al., 2011), and in others. In the chapter 3 we present detailed survey of the decision making studies, where the solutions of theoretical and real world problems expressed sensitivity with regards to the preference ambiguity. This dissertation is dedicated to study the preference ambiguity in the utility theory of decision making, and to present methods to solve the sensitivity problems it may cause. The rest of the dissertation is organized as follows.

In the chapter 2 we review one of the most used techniques in the utility theory- robust optimization with stochastic optimization constraints. Introduced by (Dentcheva & Ruszczynski, 2003), this models exhibit preference ambiguity averse properties for the decision making problems under uncertainty in the outcomes. Due to its construction, the classical optimization model with stochastic dominance constraints requires the solution to be preferable for all hypothetically possible decision makers sharing a specific risk attitude. Unfortunately, it also includes some extreme utility functions rarely characterizing a human decision maker's preference.

In this section we construct a new type of stochastic relationship reference-based almost stochastic dominance (RSD). The concept of RSD addresses the two problems in utility-based decision making:

(i) ambiguity and inaccuracy in characterizing the decision maker's individual risk attitude,

(ii) over-conservativeness of stochastic dominance representing general properties of risk aversion.

The RSD rule reveals the maximum dominance level quantifying the robustness of the decision maker's preference between alternative choices. We first develop an approximation model using Bernstein polynomials, show the asymptotic convergence of its optimal value and set of optimal solutions to the true counterparts as the degree of Bernstein polynomials increases, and analyze the convergence rate of its feasible region. We next develop a cut-generation algorithm to solve the approximation model. Finally, we further adapt this cut-generation algorithm to seek a valid option most robustly preferable to a random benchmark. The effectiveness and computational complexity of the model have been illustrated using a portfolio optimization problem.

In the chapter 3 we study the sensitivity of the solutions of decision making models with regards to preference ambiguity. We propose a novel bi-objective optimization model incorporating preference maximization and sensitivity analysis on decision maker's preference ambiguity (ESM). The concept of (ESM) addresses the preference ambiguity and inaccuracy in characterizing the decision maker's individual risk attitude. Unlike classical approaches, the (ESM) does sensitivity analysis in the optimization step. Which allows optimal solutions to be both risk-averse with regards to outcome uncertainty and sensitivity-averse with regards to decision maker's preference ambiguity. We first study special (SMU) sensitivity measure properties. Based on which, construct two linearizion reformulations of (ESM). The first reformulation is for preference maximizer decision makers. It is a mixed-integer, bi-objective optimization model. The second reformulation is for risk-averse decision makers. In this case the bi-objective optimization model is linear. Finally, we apply the (ESM) model for the budget allocation to ten major urban areas in the United States under the Urban Areas Security Initiative (UASI). We study the various solutions of the problem for different decision makers. Compare their allocation policies to the government's average allocation and to the allocation suggested by RAND corporation.

Chapter 2

Optimization with Reference-Based Robust Preference Constraints

2.1 Introduction

Extensive research has addressed utility assessment methods using preference comparison, probability equivalence, value equivalence, certainty equivalence, and statistical regression (Farquhar, 1984; Wakker & Deneffe, 1996). However, typically it is not possible to specify the decision maker's preference precisely and this preference ambiguity due to a lack of accurate description of human behavior is a basic assumption and concern in the random utility theory (Thurstone, 1994). In the literature, this ambiguity is caused by two major challenges in realistic applications: (Chl 1) technical and (Chl 2) inherit (Chajewska & Koller, 2000; Karmarkar, 1978; Savage, 2012; Thurstone, 1994; Weber & Borcherding, 1993).

Stochastic dominance is an alternative stochastic ordering approach (Bawa, 1975; Hadar & Russell, 1969; Hanoch & Levy, 1975; Müller & Stoyan, 2002; Shaked & Shanthikumar, 1994). Consider two random variables $X, Y \in (\Omega, \mathcal{F}, P; \Theta)$ with the support $\Theta := [\underline{\theta},\ \overline{\theta}]$. X stochastically dominates Y in the mth order if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for any utility function u satisfying $(-1)^{k-1}u^{(k)}(x) \geq 0$, $k = 1, \ldots, m$, for all $x \in \Theta$ (Brockett & Golden, 1987; Levy, 2016). This utility class represents basic properties of risk aversion. For example, the decision maker prefers more to less if $u'(x) \geq 0$, is risk averse if $u''(x) \leq 0$, and becomes prudent if $u'''(x) \geq 0$ on Θ. The use of stochastic dominance to compare alternatives avoids the ambiguity in assessing the decision maker's specific utility. However, stochastic dominance based preference is often unnecessarily over-conservative (Hu et al., 2014; Leshno & Levy, 2002; Lizyayev & Ruszczyński, 2012). We consider the following example analogous to one given in (Levy, 2016). Suppose that Hannah wants to invest in one of the following lottery tickets priced at \$1:

- $\bar{X}_1$: yielding $\$0$ with the probability of 1% and \$2 with the probability of 99%,
- $\bar{Y}_1$: yielding \$0.01 (1 penny) with certainty.

Although $\bar{X}_1$ is much more preferred by the rational decision maker to $\bar{Y}_1$, stochastic dominance does not support the preference of $\bar{X}_1$ over $\bar{Y}_1$. In this case, the support $\Theta = [0, 2]$, where $\underline{\theta} = 0$ means that Hannah loses all investment and $\overline{\theta} = 2$ shows that Hannah gains 100% profit. Let Hannah's utilities at \$0 and \$2 be 0 and 1, respectively. It can be seen that, for the utility function $\tilde{u}(x) := \sqrt[1000]{x/2}$, we have that $\mathbb{E}[\tilde{u}(\bar{X}_1)] \leq \mathbb{E}[\tilde{u}(\bar{Y}_1)]$. Since $\tilde{u}$ is infinitely differentiable in Θ ($\tilde{u}'(0) = \infty$ is allowed) and has the derivatives with alternating signs regarding the degrees of the derivatives, $\tilde{u}$ belongs to the utility class of any order stochastic dominance. The risk characterization specified by $\tilde{u}$ overvalues very small gains, but completely neglects the possible difference in large gains ($\tilde{u}(x)$ has a very stiff increase for a small x, while being flat for a large x). Such behavior is quite unnatural; however, stochastic dominance requires that the preference should hold for all suitable utility functions including $\tilde{u}$.

This study proposes a novel robust preference relationship — reference-based almost stochastic dominance (RSD) — to resolve the problems of both the ambiguity in characterizing the decision maker's individual risk preference and the over-conservativeness of stochastic dominance representing the generality of risk-aversion. These two problems have been studied separately in the literature. Leshno and Levy (Leshno & Levy, 2002), Lizyayev and Ruszczyǹski (Lizyayev & Ruszczyński, 2012), and Hu et al. (Hu et al., 2014) relaxed the distributionally defined forms of the first and second order stochastic dominance. Leshno and Levy (Leshno & Levy, 2002) imposed relatively boundary restrictions on the derivatives of utility functions for eliminating extreme cases (e.g. $\tilde{u}$). The relaxations proposed in (Hu et al., 2014; Lizyayev & Ruszczyński, 2012) use error tolerance in comparison with the expected utility values of lotteries rather than specify restrictions on utility functions. Working on the CVaR based interpretation of the second order stochastic dominance, Noyan and Rudolf (Noyan & Rudolf, 2013) proposed the relaxed stochastic ordering which requires the CVaR of the preferred random variable to be larger over a shrunk set of confidence levels. These approaches do not comply with the decision maker's individual risk preference, while resolving the problem of the over-conservativeness of stochastic dominance. Armbruster and Delage (Armbruster & Delage, 2015), and Hu and Mehrotra (Hu & Mehrotra, 2014) developed robust expected utility maximization models which consider individualizing the set of utility functions to meet the decision maker's risk attitude. Armbruster and Delage (Armbruster & Delage, 2015) used the paired game method where the decision maker's risk attitude is partially characterized by her pairwise preference of designed lotteries. Hu and Mehrotra (Hu & Mehrotra, 2014) specified boundary conditions on utility and marginal utility functions using parametric utility assessments, and construct auxiliary conditions using both standard and paired game methods. Our approach, differently, is to construct a perturbation neighborhood around the reference utility function assessed to characterize the decision maker's risk attitude. The size of the perturbation neighborhood quantifies the level of the ambiguity in the assessment, which are not addressed by Armbruster and Delage (Armbruster & Delage, 2015) and Hu and Mehrotra (Hu & Mehrotra, 2014).

We address optimization problems with the RSD constraint. Dentcheva and Ruszczyński (Dentcheva & Ruszczynski, 2003, 2004) first introduced stochastic dominance constrained optimization problems, which pursue expected profit while hedging risk by choosing options preferable to a random benchmark. Since the last decade, optimization models using stochastic dominance have been the subject of theoretical considerations and practical applications in areas such as business, finance, energy and transportation, e.g. (Dentcheva et al., 2007; Dentcheva & Ruszczyński, 2006; Drapkin & Schultz, 2010; Haskell & Jain, 2015; Hu et al., 2010; Karoui & Meziou, 2006; Lean et al., 2010; Luedtke, 2008; Nie et al., 2011; Roman et al., 2006; Sun & Xu, 2013).

Our model uses the concept of functional robustness. We specify a nonparametric shape-preserving neighborhood of the decision maker's reference utility function. This specification is suitable for classical nonparametric standard gamble methods and paired gamble methods such as preference comparison, probability equivalence, value equivalence, and certainty equivalence ((Farquhar, 1984; Wakker & Deneffe, 1996), and reference therein). The concept of functionally robust optimization was first proposed by Hu et al. (Hu et al., 2019) in a news vendor problem for the unknown mathematical form of a price-demand function, which is different from traditional robust approaches requiring the knowledge of the functional form, e.g. (Ben-Tal & Nemirovski, 1998; Bertsimas et al., 2010, 2004; Delage & Ye, 2010; el Ghaoui & Lebret, 1997; Scarf, 1958; Shapiro & Ahmed, 2004). To specify the uncertainty set of the price-demand function, Hu et al. (Hu et al., 2019) considered an error allowance for the least-squares fitting at discrete data points. We generalize their approach to introduce an $\mathcal{L}_2$-norm based tolerance of perturbation around the reference utility function. In the context of stochastic dominance, this tolerance is interpreted as the decision maker's desired dominance level.

We develop an approximate solution method using Bernstein polynomials. Bernstein polynomial based approximation is one of broadly-applied interpolation and curve fitting techniques in engineering computation (Ditzian, 1989; Lorentz, 2012; Ray, 2018; Rivlin, 2010). The optimal value and the set of optimal solutions of the approximation model asymptotically converge to the true counterparts as the degree n of Bernstein polynomials increases. It is also shown that the feasible region of the approximation model converges with a rate $n^{-1/2}$. Finally, a cut-generation algorithm is developed to solve the approximation model. In addition, we further adapt this cut-generation algorithm to seek a valid option most robustly preferable to a random benchmark.

This chapter is organized as follows. In Section 2.2, we define RSD, and use the examples of Hannah's comparing lotteries to illustrate the maximum and desired dominance levels. Section 2.3 develops an optimization model with the RSD constrain, and discusses its approximation using Bernstein polynomials. We study the convergence rate of the approximate feasible region, and analyze the asymptotic convergence of the optimal value and the set of optimal solutions in the approximation approach. In Section 2.4 we develop a cut-generation algorithm to solve the approximation model. Section 2.5 adapts this cut-generation algorithm to seek a valid option most robustly preferable to a random benchmark. The effect of the RSD constraint and the complexity of the algorithm are illustrated in Section 2.6 by using the financial portfolio optimization problem given in (Dentcheva & Ruszczynski, 2003). Section 2.7 concludes.

2.2 Reference-Based Almost Stochastic Dominance (RSD)

In this section we first give the definition of RSD, and next discuss the robust level of preference based on the concept of RSD.

2.2.1 Definition of RSD

The concept of RSD is specified as a preference relationship based on the neighborhood of the decision maker's reference utility function. For the risk averse decision maker, utility functions should be increasing and concave. Without loss of generality, assume that the reference utility function, denote by u_{ref}, is increasing and concave on Θ, and satisfies $u_{ref}(\underline{\theta}) = 0$ and $u_{ref}(\overline{\theta}) = 1$. In practice, u_{ref} could be assessed in any classical parametric or non-parametric approach surveyed in (Farquhar, 1984; Wakker & Deneffe, 1996), such as preference comparison, probability equivalence, value equivalence, certainty equivalence, statistical regression, etc.

DEFINITION 1. For the reference utility function u_{ref} which is increasing and concave on Θ and satisfies $u_{ref}(\underline{\theta}) = 0$ and $u_{ref}(\overline{\theta}) = 1$, a random variable $X \in (\Omega, \mathcal{F}, P; \Theta)$ is preferred to another random variable $Y \in (\Omega, \mathcal{F}, P; \Theta)$ in the mth ($m \geq 2$) order reference-based almost stochastic dominance (RSD) for a given $\epsilon \in [0, 1]$ (written as $X \succeq^{\epsilon}_{(m)} Y$ w.r.t. u_{ref}), if

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)],$$

for any utility function u satisfying

(A1). for any $x \in \Theta$, $(-1)^{i-1}u^{(i)}(x) \geq 0$, $i = 1, \ldots, m$,

(A2). $u(\underline{\theta}) = u_{ref}(\underline{\theta}) = 0$, and $u(\overline{\theta}) = u_{ref}(\overline{\theta}) = 1$,

(A3). $\|u - u_{ref}\|_{\mathcal{L}_2} \leq \epsilon$, where $\|f\|_{\mathcal{L}_2} := \left(\int_\Theta f(t)^2 \mu(dt)\right)^{1/2}$ for a given nonnegative measure μ with $\mu(\Theta) = 1$,

(A4). for any $x \in \Theta$, $u(x) \leq \frac{M}{1-\epsilon} u_{ref}(x)$, for a fixed $M > 1$ (if $\epsilon = 1$, then $u(x) < \infty$).

Condition (A1) represents the utility class required by the mth order stochastic dominance to describe the basic properties of risk aversion. Condition (A2) is the normalization of utility functions which preserves preference rankings (Theorem 4.1 in (Keeney & Raiffa, 1993)). Condition (A3) generates a neighborhood around the reference u_{ref}. This neighborhood is a closed ball, with the radius ϵ, on the $\mathcal{L}_2$-normed space with the measure μ on Θ. The $\mathcal{L}_2$ norm, as the fundamental metric in least-squares methods, is the widely used to measure estimation errors and perturbation tolerance, particularly in the fields of statistical regression, model fitting, and perturbation analysis. Condition (A4) provides a uniform bound on the moduli of continuity of utility functions. For $\delta > 0$, the modulus of continuity of u is denoted as

$$\omega(\delta) := \sup_{\substack{x_1, x_2 \in \Theta \\ |x_1 - x_2| \leq \delta}} |u(x_1) - u(x_2)|. \tag{2.1}$$

By conditions (A1), (A2), and (A4), we have that

$$\begin{aligned} \omega(\delta) = u(\underline{\theta} + \delta) - u(\underline{\theta}) = u(\underline{\theta} + \delta) &\leq \frac{M}{1-\epsilon} u_{ref}(\underline{\theta} + \delta) \\ &= \frac{M}{1-\epsilon}(u_{ref}(\underline{\theta} + \delta) - u_{ref}(\underline{\theta})) = \frac{M}{1-\epsilon}\omega_{ref}(\delta), \end{aligned} \tag{2.2}$$

where ω_{ref} is the modulus of continuity of u_{ref}. It requires that the changes of utility values over any small interval vary in a reasonable range decided by ω_{ref} and the constant $M/(1-\epsilon)$.

Note that under conditions (A1) and (A2), if $\epsilon = 1$, conditions (A3) and (A4) become redundant, and RSD represents stochastic dominance. On the other hand, if u_{ref} satisfies condition (A1), it is a unique utility function satisfying conditions (A3) and (A4) when $\epsilon = 0$. RSD in this case is the preference under the characterization of the decision maker's individual risk attitude. In comparison, an ϵ in $(0, 1)$ balances the individual characterization and general representation of risk aversion. This balance excludes utility functions very different from u_{ref} from the general risk averse utility classes to weaken the over-conservativeness of stochastic dominance, while it uses the perturbation allowance to resolve the problem of the ambiguity of the individual characterization. In Sections 2.2.2 and 2.2.3, we further interpret ϵ by introducing a concept of maximum dominance level and discussing the decision maker's desired dominance level.

The measure μ in condition (A3) is used to quantify the relative importance of different subintervals of Θ. In a special case μ is a discrete measure, and condition (A3) is the weighted least squares fitting criterion. Actually, nonparametric utility assessments can only generate finitely many utility value points, and then use a piecewise linear curve to link all these points. We may specify a perturbation set based on those discrete points instead of the piecewise linear curve. Let $(x_i, u_{ref}(x_i))$, $i = 1, \ldots, I$, be the reference utility points. The measure μ should be assigned on x_i and condition (A3) is thus represented as

$$\left(\sum_{i=1}^{I} \mu(x_i)(u(x_i) - u_{ref}(x_i))^2 \right)^{1/2} \leq \epsilon. \tag{2.3}$$

Let

$$\mathfrak{U}^m(\epsilon) := \{u : u(x) \text{ satisfies conditions (A1)-(A4) for given } \epsilon\}. \tag{2.4}$$

By Definition 1, $X \succeq_{(m)}^{\epsilon} Y,\ w.r.t.\ u_{ref}$ is equivalent to $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)], \forall u \in \mathfrak{U}^m(\epsilon)$. The following propositions show the relation between different orders of RSD, and claim the non emptiness and convexity of the set $\mathfrak{U}(\epsilon)$.

PROPOSITION 1. If $X \succeq_{(m)}^{\epsilon} Y,\ w.r.t.\ u_{ref}$, then $X \succeq_{(m+1)}^{\epsilon} Y,\ w.r.t.\ u_{ref}$.

PROPOSITION 2. $\mathfrak{U}^m(\epsilon)$ is a convex set. If u_{ref} satisfies condition (A1), then $\mathfrak{U}^m(\epsilon)$ is nonempty.

2.2.2 Maximum Dominance Level

RSD balances the individual characterization and general representation of risk aversion for an appropriate ϵ. To further interpret ϵ, we need to define the maximum dominance level as follows

DEFINITION 2. For two random variables $X, Y \in (\Omega, \mathcal{F}, P; \Theta)$ satisfying $\mathbb{E}[u_{ref}(X)] \geq \mathbb{E}[u_{ref}(Y)]$, the maximum level of X almost dominating Y in the mth order w.r.t u_{ref} is

$$\mathcal{E}^{(m)}(X, Y; u_{ref}) := \sup \{\epsilon \in [0, 1] : X \succeq_{(m)}^{\epsilon} Y \text{ w.r.t. } u_{ref}\}. \tag{2.5}$$

The maximum dominance level $\mathcal{E}^{(m)}(X, Y; u_{ref})$ quantifies how robustly X is preferred to Y under u_{ref}. In conditions (A3) and (A4), ϵ is the allowed tolerance of the perturbation around u_{ref} defined using the $\mathcal{L}_2$-norm metric and the bound on the modulus of continuity. The maximum dominance level is the largest tolerance of perturbation under which the preference of X over Y holds, and this tolerance indicates the robust level of the preference. Note that X stochastically dominates Y in the mth order if and only if $\mathcal{E}^{(m)}(X, Y; u_{ref}) = 1$. Moreover, this measure of preference robustness only relies on u_{ref}, and is consistent in comparing different lotteries. For example, for two pairs (X, Y) and (R, S), if $\mathcal{E}^{(m)}(X, Y; u_{ref}) > \mathcal{E}^{(m)}(R, S; u_{ref})$, it means that the preference of X over Y holds under a larger tolerance of perturbation around u_{ref}. That is, for the given u_{ref}, the preference of X over Y is more robust than that of R over S. In the following example, we discuss the maximum dominance level of Hannah's purchasing lottery.

EXAMPLE 1. *Maximum dominance level of Hannah's purchasing lottery.*

We now illustrate the maximum dominance level using the case of Hannah's purchasing lottery tickets $\bar{X}_1$ and $\bar{Y}_1$. It has been shown that stochastic dominance is unable to reveal the preference of $\bar{X}_1$ over $\bar{Y}_1$, for conservatively taking unreasonable utility functions (e.g. $\tilde{u}$ given in the introduction) into consideration. Now suppose that Hannah's risk preference is approximately characterized as $\bar{u}_{ref}(x) = \sqrt{x/2}$, which we use as the reference utility function. The maximum dominance level $\mathcal{E}^{(2)}(\bar{X}_1, \bar{Y}_1; \bar{u}_{ref}) = 0.398$. Hence, with the given $\epsilon \le 0.398$, we have that $\bar{X}_1 \succeq^{\epsilon}_{(2)} \bar{Y}_1,\ w.r.t.\ \bar{u}_{ref}$, and condition (A3) excludes utility functions that Hannah is unwilling to choose (e.g. $\tilde{u}$). To understand the statement that the maximum dominance level quantifies the preference level in the sense of robustness, we denote by $\bar{Y}_{\$1}$ the non-purchase option that Hannah does not purchase any lottery ticket. This option is equivalent to purchasing the lottery of which the ticket is priced \$1 and the return is also \$1 certainly. Hence, Hannah has no gain but never takes a risk. The rational decision maker does not purchase $\bar{Y}_1$ on which the investment is lost for sure. This undoubted fact is supported by $\mathcal{E}^{(2)}(\bar{Y}_{\$1}, \bar{Y}_1; \bar{u}_{ref}) = 1$, which indicates $\bar{Y}_{\$1}$ stochastically dominates $\bar{Y}_1$. So we can claim that Hannah absolutely prefers $\bar{Y}_{\$1}$ to $\bar{Y}_1$. In contrast, although $\bar{X}_1$ could be better than $\bar{Y}_{\$1}$ ($\mathbb{E}[\bar{u}_{ref}(\bar{X}_1)] = 0.99 > \bar{u}_{ref}(\bar{Y}_{\$1}) = 0.707$), we cannot say that $\bar{X}_1$ is absolutely preferred to $\bar{Y}_1$ according to stochastic dominance rules. Indeed, there is still a theoretical possibility that $\bar{Y}_1$ is preferred to $\bar{X}_1$. It can only be said that the preference of $\bar{X}_1$ over $\bar{Y}_1$ reaches the level $\mathcal{E}^{(2)}(\bar{X}_1, \bar{Y}_1; \bar{u}_{ref})$ of robustness. We also consider an additional \$1 lottery ticket as

- $\bar{X}_2$: yielding $\$0$ with the probability of 10% and \$2 with the probability of 90%,

While $\bar{X}_1$ stochastically dominates $\bar{X}_2$, $\bar{X}_2$ is still more attractive then $\bar{Y}_1$. This preference is visible also from $\mathcal{E}^{(2)}(\bar{X}_2, \bar{Y}_1; \bar{u}_{ref}) = 0.069$. Comparing the two pairs $(\bar{X}_1, \bar{Y}_1)$ and $(\bar{X}_2, \bar{Y}_1)$, we have that $\mathcal{E}^{(2)}(\bar{X}_1, \bar{Y}_1; \bar{u}_{ref}) > \mathcal{E}^{(2)}(\bar{X}_2, \bar{Y}_1; \bar{u}_{ref})$. These results show that $\bar{X}_1$ is more robustly preferred to $\bar{Y}_1$ than $\bar{X}_2$. □

2.2.3 The Decision Maker's Desired Dominance Level

The concept of the maximum dominance level quantifies the largest tolerance of perturbation, under which a preference holds, around the reference utility function u_{ref}. We interpret ϵ in conditions (A3) and (A4) as the decision maker's desired dominance level which is her accepted largest tolerance of perturbation around u_{ref}. This desired dominance level is also interpreted as the unique robust level of the decision maker's preference with which she can assert a lottery is sufficiently preferred to another in the sense that the ambiguity and inconsistency in the elicitation of u_{ref} is not very sensitive.

A principle in utility theory is that the decision maker's risk attitude does not vary in lottery comparisons. Utility functions, excluded in a lottery comparison for not representing the decision maker's risk attitude, should be also excluded in any other lottery comparison. Hence, her desired dominance level is solely related to her reference utility function. This property allows us to adapt the probability equivalence method, one of classical utility assessment approaches (Farquhar, 1984; Wakker & Deneffe, 1996), to evaluate the decision maker's desired dominance level. Denote by $X(p)$ the lottery which yields $\underline{\theta}$ with the probability $1-p$ and $\overline{\theta}$ with the probability p, and let $\vartheta_1, \ldots, \vartheta_K$ be selected deterministic points in Θ. The traditional probability equivalence method asks the decision maker to decide an exact p_k such that $X(p_k)$ is equivalent to ϑ_k for each $k = 1, \ldots, K$, i.e., the expected utility of $X(p_k)$ equals the utility of ϑ_k. However, it could be difficult to give this exact p_k.

Our approach requires the decision maker to suggest a range of p_k, denoted by $\left[\underline{p}_k, \overline{p}_k\right]$. It means that, in her belief, $X(p)$ is sufficiently better than ϑ_k if $p \geq \overline{p}_k$, and worse if $p \leq \underline{p}_k$. Note that

$$\mathbb{E}[u_{ref}(X(\underline{p}_k))] \leq u_{ref}(\vartheta_k) \leq \mathbb{E}[u_{ref}(X(\overline{p}_k))]. \tag{2.6}$$

for u_{ref} characterizes her individual risk attitude. The decision maker's desired dominance level is estimated by both

$$\underline{\epsilon}_k := \mathcal{E}^{(m)}(\vartheta_k, X(\underline{p}_k); u_{ref}), \text{ and } \overline{\epsilon}_k := \mathcal{E}^{(m)}(X(\overline{p}_k), \vartheta_k; u_{ref}). \tag{2.7}$$

Therefore, we hypothetically have that

$$\underline{\epsilon}_1 = \overline{\epsilon}_1 = \cdots = \underline{\epsilon}_K = \overline{\epsilon}_K. \tag{2.8}$$

However in practice, the decision maker could give inconsistent estimates. In this case, we choose the average, $\frac{1}{2K}\sum_{k=1}^{K}(\underline{\epsilon}_k + \overline{\epsilon}_k)$, as the overall estimate of the desired dominance level.

Further relieving the difficulty in assessing the probability range $\left[\underline{p}_k, \overline{p}_k\right]$ in the adapted probability equivalence method, we design a simple "best-to-worst" lottery table (e.g. Table 2.1) to assist the decision maker in her choice. This table lists $X(p)$ for some discrete points of p and their maximum dominance levels with respect to ϑ_k. Example 2 addresses the estimation of Hannah's desired dominance level.

EXAMPLE 2. *Estimation of Hannah's desired dominance level.*

We illustrate the estimation of Hannah's desired dominance level. Choose $\vartheta_1 = \$0.5$, $\vartheta_2 = \$1$, and $\vartheta_3 = \$1.5$ to equally partition her interested region $\Theta = [0, 2]$, and compare these points and the lottery ticket $X(p)$ which in this case is priced at \$1, and yields $\$0$ with the probability $1 - p$ and \$2 with the probability p. Note that ϑ_2 is the non-purchase option $\bar{Y}_{\$1}$ of which the ticket is priced \$1 and the return is also \$1 certainly. In what follows we discuss the comparison between $X(p)$ and ϑ_2 as an example.

Table 2.1: Dominance Levels of Lottery Tickets

Lottery Ticket	Probability of Yield $0	$2	Maximum Dominance Level $\mathcal{E}^{(2)}(\bar{X}_i, \vartheta_2; \bar{u}_{ref})$	$\mathcal{E}^{(2)}(\vartheta_2, \bar{X}_i; \bar{u}_{ref})$
$\bar{X}_1$	1%	99%	0.122	—
$\bar{X}_2$	10%	90%	0.069	—
$\bar{X}_3$	15%	85%	0.045	—
$\bar{X}_4$	20%	80%	0.024	—
$\bar{X}_5$	25%	75%	0.007	—
$\bar{X}_6$	30%	70%	—	0.001
$\bar{X}_7$	35%	65%	—	0.014
$\bar{X}_8$	40%	60%	—	0.034
$\bar{X}_9$	45%	55%	—	0.057

Table 2.1 gives lottery tickets $\bar{X}_i$, for $i = 1, \ldots, 9$, which are $X(p)$ at selected discrete points. Since $\mathbb{E}[\bar{u}_{ref}(\bar{X}_i)] > \bar{u}_{ref}(\vartheta_2)$, $i = 1, \ldots, 5$, $\mathcal{E}^{(2)}(\vartheta_2, \bar{X}_i; \bar{u}_{ref})$ do not exist. Analogously, it is the same for $\mathcal{E}^{(2)}(\bar{X}_i, \vartheta_2; \bar{u}_{ref})$, $i = 6, \ldots, 9$. Since $\bar{X}_i$ stochastically dominates $\bar{X}_{i+1}$, the preference of $\bar{X}_i$ is monotonously weakened such that $\mathcal{E}^{(2)}(\bar{X}_i, \vartheta_2; \bar{u}_{ref})$ decreases, and $\mathcal{E}^{(2)}(\vartheta_2, \bar{X}_i; \bar{u}_{ref})$ increases. Hannah is requested to select the lottery tickets from the list which she is not reluctant to purchase or must not purchase. Since ϑ_2 is a non-yielding but risk-free option, Hannah's choice indicates the level of her insistence on risky investment. Suppose that she picks first three lotteries unhesitatingly, and decides not to purchase the last two. But she feels it difficult to make a decision on $\bar{X}_4$ and $\bar{X}_7$. Then her choices on the upside and downside yield two estimates of her desired dominance level, which are $\bar{\epsilon}_2 = \mathcal{E}^{(2)}(\bar{X}_3, \vartheta_2; \bar{u}_{ref}) = 0.045$ and $\underline{\epsilon}_2 = \mathcal{E}^{(2)}(\vartheta_2, \bar{X}_8; \bar{u}_{ref}) = 0.034$. In other words, she would like to invest in the lottery ticket that almost dominates the non-purchase option with the maximum level no less than 0.045, and not to purchase ones that are almost dominated with the maximum level no less than 0.034.

To simplify the discussion, we assume that $\bar{\epsilon}_1 = \bar{\epsilon}_2 = \bar{\epsilon}_3$ and $\underline{\epsilon}_1 = \underline{\epsilon}_2 = \underline{\epsilon}_3$ without loss of generality. Hannah's desired dominance is estimated to be $(0.045 + 0.034)/2 = 0.040$. Hannah's decision is that $\bar{X}_1, \bar{X}_2, \bar{X}_3$ are sufficiently preferred to ϑ_2 in the 2nd order RSD, while $\bar{X}_4$ and $\bar{X}_5$ may be indifferent. She also thinks that $\bar{X}_9$ is obviously worse than ϑ_2, and $\bar{X}_6, \bar{X}_7$ are worth to carefully consider. A special case is $\bar{X}_8$. Although Hannah prefers ϑ_2 to $\bar{X}_8$, ϑ_2 and $\bar{X}_8$ are indifferent under RSD with the estimated dominance level. This disagreement is ascribed to not only the inconsistency of the decision maker's describing her preference in different lottery comparisons, but also the estimation error coming from the discretization of the probability in Table 2.1.

Conditions (A3) and (A4) with Hannah's desired dominance level $\epsilon = 0.040$ exclude utility functions which are too different from her reference utility function $\bar{u}_{ref}$ to represent her risk attitude possibly. For example, $\tilde{u}(x) = \sqrt[1000]{x/2}$ in the introduction is an irrational utility function for Hannah since $\|\tilde{u} - \bar{u}_{ref}\|_{\mathcal{L}_2} = 0.56$ is larger than her desired dominance level 0.040. It is worth to mention that stochastic dominance and Leshno and Levy's almost stochastic dominance in (Leshno & Levy, 2002) do not exclude a risk neutral (linear) utility function. It means that X cannot dominate Y if $\mathbb{E}[X] < \mathbb{E}[Y]$. However, for the risk averse decision maker, the risk neutral utility function may be an extreme, and unable to represent her risk averse risk attitude. Hence, by choosing a suitable ϵ, condition (A3) is able to exclude the risk neutral utility function. In the Hannah example, $\bar{u}(x) := x/2$ is the risk neutral utility function, and its $\mathcal{L}_2$-norm based distance from $\bar{u}_{ref}$ is $\|\bar{u} - \bar{u}_{ref}\|_{\mathcal{L}_2} = 0.167$. Since Hannah's desire dominance level is 0.040, $\bar{u}$ is not included in the utility set specified by conditions (A1)-(A4) of RSD. In Table 2.1 $\mathbb{E}[\bar{X}_i] > \mathbb{E}[\vartheta_2]$ for $i = 6, \ldots, 9$. Hannah's decision on not purchasing $\bar{X}_9$ implies that the risk neutral utility function conflicts with her risk attitude. □

2.3 Optimization Model Using the Reference-Based Almost Stochastic Dominance

In this section we first represent a RSD constrained stochastic optimization model, and next study an approximation approach to the RSD constrained model using Bernstein polynomials. Finally, we discuss the connection of the RSD constrained model with its approximation, and show the asymptotic convergence of the approximation.

2.3.1 RSD Constrained Optimization Model

A stochastic optimization model using the mth ($m \geq 2$) order RSD as a risk averse constraint is specified as follows:

$$\begin{aligned} \max_{z} \quad & f(z) \\ \text{subject to} \quad & X(z) \succeq^{\epsilon}_{(m)} Y, \; w.r.t. \; u_{ref}, \end{aligned} \tag{RSD-P}$$

$$z \in \mathfrak{Z},$$

where $\mathfrak{Z} \subseteq \mathbb{R}^d$ is a decision region, $f : \mathbb{R}^d \mapsto \mathbb{R}$ is an objective function, $X : \mathfrak{Z} \mapsto (\Omega, \mathcal{F}, P; \Theta)$ represents a random outcome function of the decision, and $Y \in (\Omega, \mathcal{F}, P; \Theta)$ is a benchmark. In RSD-P, the RSD constraint requires that the random outcome $X(z)$ at a valid decision z should almost dominate the random benchmark Y in the mth order for the given dominance level ϵ with respect to the reference utility function u_{ref}.

We now state notations needed in the discussion. Denote

$$\pi^m(z) := \min_{u \in \mathfrak{U}^m(\epsilon)} \left\{ \Pi(z, u) := \mathbb{E}[u(X(z)) - u(Y)] \right\}, \tag{2.9}$$

and

$$\Psi^m(\delta) := \{z \in \mathfrak{Z} : \pi^m(z) \geq \delta\}. \tag{2.10}$$

By Definition 1, the RSD constraint in RSD-P equals to $\pi^m(z) \geq 0$. Hence, $\Psi^m(0)$ is the feasible region of RSD-P. By Proposition 1, we have the following relationship of the set $\Psi^m(\delta)$ for different m's.

PROPOSITION 3. For any $\delta \in \mathbb{R}$, $\Psi^m(\delta) \subseteq \Psi^{m+1}(\delta)$.

2.3.2 Approximation Using Bernstein Polynomials

Model RSD-P is a functionally robust optimization problem, where the RSD constraint is specified using the set $\mathfrak{U}^m(\epsilon)$ of nonparametric utility functions. We now discuss an approximate solution method of RSD-P using Bernstein polynomials. Bernstein polynomial based approximation is one of broadly-applied interpolation and curve fitting techniques in engineering computation (Ditzian, 1989; Lorentz, 2012; Ray, 2018; Rivlin, 2010). In our approach, Bernstein polynomials preserve the geometry of risk aversion (condition (A1)) that utility functions in $\mathfrak{U}^m(\epsilon)$ are m times differentiable, and have the derivatives with alternating signs, regarding the degrees of the derivatives. Meanwhile, the behavior of Bernstein polynomials at boundary points tallies with the normalization of utility function (condition (A2)), and their uniform convergence property satisfies the requirements on the perturbation level (condition (A3)) and the uniformly bounded modulus of continuity (condition (A4)).

Let the vector $\phi(x) := (\phi_0(x), \ldots, \phi_n(x))^T$, where

$$\phi_j(x) := \begin{pmatrix} n \\ j \end{pmatrix} \left(\frac{x-\underline{\theta}}{\overline{\theta}-\underline{\theta}}\right)^j \left(1-\frac{x-\underline{\theta}}{\overline{\theta}-\underline{\theta}}\right)^{n-j}, \quad j = 0, \ldots, n, \tag{2.11}$$

are the bases of Bernstein polynomial on Θ. The nth degree Bernstein polynomial is given as

$$B_n(x; c) := c^T \phi(x),\ x \in \Theta, \tag{2.12}$$

where $c := (c_0, \ldots, c_n)^T$ is a vector of coefficients. Note that condition (A1) requires utility functions to be m times differentiable. In order to avoid trivial solutions, we require that $n \geq m$. Consider the following conditions on coefficients $c = (c_0, \ldots, c_n)^T$:

(B1). $(-1)^{i-1}\Delta^i c_j \geq 0$, $i = 1 \ldots m,\ j = 0 \ldots n-i$, where

$$\Delta^i c_j := \sum_{k=0}^{i} (-1)^k \begin{pmatrix} k \\ i \end{pmatrix} c_{j+i-k},$$

(B2). $c_0 = u_{ref}(\underline{\theta}) = 0$, and $c_n = u_{ref}(\overline{\theta}) = 1$,

(B3). $c^T A c + g^T c + r \leq \epsilon^2$, where

$$A := \int_{\Theta} \phi(x)\phi^T(x)\mu(dx),\ g := -2\int_{\Theta} u_{ref}(x)\phi(x)\mu(dx),$$
$$\text{and } r := \int_{\Theta} u_{ref}^2(x)\mu(dx).$$

(B4). $c_j \leq \min\left\{1,\ \frac{M}{1-\epsilon} u_{ref}(\underline{\theta} + j\frac{\overline{\theta}-\underline{\theta}}{n})\right\},\ j = 0 \ldots n.$

Denote the set of coefficient by

$$\mathfrak{C}_n^m(\epsilon) := \{c \in \mathbb{R}^{n+1} : c \text{ satisfies conditions (B1)-(B4)}\}. \tag{2.13}$$

The theorem below states that $\mathfrak{U}^m(\epsilon)$ contains the set of Bernstein polynomials with the coefficients belonging to $\mathfrak{C}_n^m(\epsilon)$.

THEOREM 1. Let $\mathfrak{B}_n^m(\epsilon) := \{B_n(x;c) : c \in \mathfrak{C}_n^m(\epsilon)\}$. $\mathfrak{B}_n^m(\epsilon) \subseteq \mathfrak{U}^m(\epsilon)$.

Theorem 1 shows that a Bernstein polynomial $B_n(x;c)$ belongs to $\mathfrak{U}_n^m(\epsilon)$ if its coefficient c satisfies conditions (B1) - (B4). On the other hand, if $B_n(x;c) \in \mathfrak{U}_n^m(\epsilon)$, c must obey by conditions (B1) - (B3), but may not satisfy condition (B4). Letting

$$\pi_n^m(z) := \min_{c \in \mathfrak{C}_n^m(\epsilon)} \{\Pi_n(z,c) := \mathbb{E}[B_n(X(z),c) - B_n(Y,c)]\}, \tag{2.14}$$

we now present an approximation of the RSD constraint using Bernstein polynomials as

$$\pi_n^m(z) \geq 0. \tag{BSD}$$

Similarly, the set $\Psi^m(\delta)$ and RSD-P are approximated as

$$\Psi_n^m(\delta) := \{z \in \mathfrak{Z} : \pi_n^m(z) \geq \delta\}, \tag{2.15}$$

and

$$\max_{z \in \Psi_n^m(0)} f(z). \tag{BSD-P}$$

The degree n of Bernstein polynomials is an important parameter of the approximation BSD-P. The next section will discuss the asymptotic convergence of BSD-P to RSD-P as n increases to infinity and its convergence rate. Hence, we call n the degree of BSD-P.

2.3.3 Relationship Between Models RSD-P and BSD-P

We now discuss the connection between RSD-P and its approximation BSD-P. Two theorems are given below to analyze the asymptotic convergence of BSD-P and its convergence rate. Theorem 2 describes that the feasible region of BSD-P converges to that of RSD-P with a rate $n^{-1/2}$ (n is the degree of BSD-P). Theorem 3 shows the asymptotic convergence of the optimal value and the set of optimal solutions of BSD-P.

THEOREM 2. Suppose $u_{ref} \in \mathfrak{U}^m(\epsilon)$. For any $\delta \in (0,1)$, choose the degree n of BSD-P such that

$$\delta \geq \Lambda(n) := \frac{4M}{1-\epsilon} u_{ref}\left(\frac{1}{\sqrt{n}} + \underline{\theta}\right)\left(1 + \frac{\overline{\theta} - \underline{\theta}}{n}\right). \tag{2.16}$$

Then

$$\Psi_n^m(\delta) \subseteq \Psi^m(0) \subseteq \Psi_n^m(0) \subseteq \Psi^m(-\delta). \tag{2.17}$$

REMARK 1. Theorem 2 gives a required n to reach the desired accuracy of the feasible region $\Psi_n^m(0)$ of BSD-P. That is, $\Psi^m(0) \subseteq \Psi_n^m(0) \subseteq \Psi^m(-\Lambda(n))$. By the Taylor-series expansion of u_{ref} at $\underline{\theta}$, we have

$$\Lambda(n) = \frac{4M u_{ref}(\underline{\theta})}{1-\epsilon} + \frac{4M u'_{ref}(\underline{\theta})}{1-\epsilon} n^{-1/2} + o\left(n^{-1/2}\right) \tag{2.18}$$

$$= \frac{4M u'_{ref}(\underline{\theta})}{1-\epsilon} n^{-1/2} + o\left(n^{-1/2}\right). \tag{2.19}$$

This shows that the convergence rate of $\Psi_n^m(0)$ is $n^{-1/2}$. □

We now discuss the asymptotic convergence of the optimal value and the set of optimal solutions of the approximation BSD-P.

THEOREM 3. Let ξ^m and Ξ^m be the optimal value and the set of optimal solutions of RSD-P, ξ_n^m and Ξ_n^m be the optimal value and the set of optimal solutions of BSD-P with degree n. Suppose that (i) $u_{ref} \in \mathfrak{U}^m(\epsilon)$, (ii) the set $\mathfrak{Z}$ is convex and compact, (iii) the random function $X(z)$ is concave in $\mathfrak{Z}$, (iv) there exists an interior point $\tilde{z} \in \mathfrak{Z}$ such that $\pi(\tilde{z}) \geq 0$, and (v) the objective function f is continuous in $\Psi^m(0)$, then $\xi_n^m \to \xi^m$ and $\mathbb{D}(\Xi_n^m, \Xi^m) := \max_{x \in \Xi_n^m} \min_{y \in \Xi^m} \|x - y\|_\infty \to 0$ as $n \to \infty$.

2.4 Cut Generation Algorithm for Model BSD-P

We now develop a cut-generation algorithm to solve BSD-P. This algorithm uses a sequence of coefficient cuts $c^i \in \mathfrak{C}_n^m(\epsilon)$ for $i = 1, \ldots, k$. We solve a sequence of the problems

$$\begin{aligned} &\max_z f(z) \\ &\text{s.t. } \Pi_n(z, c^i) \geq 0, \quad i = 1, \ldots, k, \\ &\quad z \in \mathfrak{Z}, \end{aligned} \tag{2.20}$$

which are the relaxations of BSD-P over the subset of $\Psi_n^m(0)$ consisting of the generated cuts. Note that, at the initial iteration $k = 0$, there are no cut constraints applied in problem (2.20). If $\mathfrak{Z}$ is convex and both $f(z)$ and $X(z)$ are concave in $\mathfrak{Z}$, problem (2.20) is a convex stochastic program, which can be solved via well-studied Monte-Carlo sampling-based methods such as sample average approximation (SAA), stochastic approximation (SA), etc. (see (Birge & Louveaux, 2011; de Mello & Bayraksan, 2014; Kim et al., 2014; Kushner & Clark, 1978; Kushner & Yin, 2003; Shapiro et al., 2014) and reference therein). Under mild assumptions, the SAA method has been shown to have fast convergence rates (Shapiro & de Mello, 2000; Shapiro et al., 2002). Also, solution quality is assessed by bounding the optimality gap of a candidate solution in the multiple replication procedure (Mak et al., 1999). Stochastic quasi-gradient (SQG) algorithms are the typical SA methods to solve convex stochastic programming problems (Ermoliev, 1983). In particular, the Mirror-Descent SQG methods introduced in (Lan, 2011; Nemirovski et al., 2009) exhibit excellent performance on convergence rate and computational time.

At the optimal solution z_k^* of problem (2.20), we calculate $\pi_n^m(z_k^*)$ by solving problem (2.14). For the sake of readability, we rewrite problem (2.14) in the explicit quadratically constrained linear programming (QCP) form as follows:

$$\pi_n^m(z) = \min_c c^T \mathbb{E}[\phi(X(z)) - \phi(Y)] \tag{2.21a}$$

$$\text{s.t. } (-1)^{i-1} \sum_{k=0}^{i} (-1)^k \begin{pmatrix} k \\ i \end{pmatrix} c_{j+i-k} \geq 0, \tag{2.21b}$$

$$j = 0, \ldots, n-i, \; i = 1, \ldots, m.$$

$$c_0 = 0, \; c_n = 1, \tag{2.21c}$$

$$c^T A c + g^T c + r \leq \epsilon^2 \tag{2.21d}$$

$$c_j \leq \min\left\{1,\ \frac{M}{1-\epsilon}u_{ref}(\underline{\theta} + j\frac{\overline{\theta}-\underline{\theta}}{n})\right\}, \quad j = 0, \ldots, n, \tag{2.21e}$$

where ϕ in the objective function (2.21a) is the vector-based function of Bernstein polynomial bases given in (2.11), and constraints (2.21b)-(2.21e) repeat conditions (B1) - (B4) that describe the feasible region $\mathfrak{C}_n^m(\epsilon)$ of problem (2.14). See (Boyd & Vandenberghe, 2004; Martein & Schaible, 1987; van de Panne, 1966) and reference therein for detail discussions on properties and solution methods of QCP problems.

Denote by c_k^* the optimal solution of problem (2.14). If $\pi_n^m(z_k^*) \geq 0$, z_k^* is the optimal solution of BSD-P. Otherwise, the constraint $\Pi_n(z, c_k^*) \geq 0$ is added to the master problem (2.20) as a valid cut. Algorithm 1 formally describes this cut-generation algorithm.

Algorithm 1 Cut-Generation Algorithm for Model BSD-P

`Step 1`	Choose $\gamma > 0$ and $k = 0$.
`Step 2`	Find the optimal solution z_k^* of problem (2.20).
`Step 3`	Calculate $\pi_n^m(z_k^*)$ by solving problem (2.14). Let c_k^* be the optimal solution.
`Step 4`	If $\pi_n(z_k^*) \geq -\gamma$, exit. Otherwise, let $c^{k+1} = c_k^*$ and $k = k+1$. Then go to step 2.

The following theorem 4 shows that Algorithm 1 terminates in finitely many iterations. Let

$$\xi_n^m(-\gamma) := \max_{z \in \Psi_n^m(-\gamma)} f(z), \tag{2.22}$$

which is a relaxation of BSD-P for $\gamma > 0$. $\xi_n^m(0)$ is the optimal value of BSD-P.

THEOREM 4. Algorithm 1 terminates in finitely many iterations. Let $\widetilde{\xi}_n^m$ be the optimal value of problem (2.20) at the last iteration where Algorithm 1 terminates. Then $\xi_n^m(0) \leq \widetilde{\xi}_n^m \leq \xi_n^m(-\gamma)$.

2.5 Optimization Problem for the Most Robust Preference

The concept of RSD can also be used to seek a valid solution, at which the random outcome is the most robustly preferred to the random benchmark. This optimization problem is specified as

$$\max_{z \in \mathfrak{Z}} \ \mathcal{E}^{(m)}(X(z), Y; u_{ref}). \tag{MDL-P}$$

Model MDL-P maximizes the maximum dominance level of the random outcome $X(z)$ at $z \in \mathfrak{Z}$ over the benchmark Y. By Definition 2, MDL-P is reformulated as

$$\begin{aligned} \max_{z,\epsilon} \ & \epsilon \\ \text{s.t. } & X(z) \succeq^{\epsilon}_{(m)} Y \text{ w.r.t. } u_{ref}, \\ & z \in \mathfrak{Z}, \ \epsilon \in [0,1]. \end{aligned} \tag{2.23}$$

For discussing the effect of the dominance level ϵ in MDL-P conveniently, we re-denote $\pi^m(z)$ in (2.9) and $\pi^m_n(z)$ in (2.14) as follows:

$$\pi^m(z;\epsilon) := \min_{u \in \mathfrak{U}^m(\epsilon)} \Pi(z,u), \text{ and} \tag{2.24}$$

$$\pi^m_n(z;\epsilon) := \min_{c \in \mathfrak{C}^m_n(\epsilon)} \Pi_n(z,c). \tag{2.25}$$

Substituting $\pi^m(z;\epsilon) \geq 0$ for the RSD constraint in problem (2.23), we rewrite MDL-P as

$$\begin{aligned} \max_{z,\epsilon} \ & \epsilon \\ \text{s.t. } & \pi^m(z;\epsilon) \geq 0, \\ & z \in \mathfrak{Z}, \ \epsilon \in [0,1], \end{aligned} \tag{2.26}$$

of which the approximation using Bernstein polynomials is correspondingly given as

$$\begin{aligned} \max_{z,\epsilon} \ & \epsilon \\ \text{s.t. } & \pi^m_n(z;\epsilon) \geq 0, \\ & z \in \mathfrak{Z}, \ \epsilon \in [0,1]. \end{aligned} \tag{2.27}$$

We now adapt Algorithm 1 to solve problem (2.27). The bisection method is used to seek the largest ϵ in problem (2.27). For a given ϵ, if there exists $z \in \mathfrak{Z}$ such that

$\pi_n^m(z;\epsilon) \geq 0$, we increase ϵ; otherwise, we decrease ϵ. To verify the existence of such z satisfying $\pi_n^m(z;\epsilon) \geq 0$ for a given ϵ, we develop a cut-generation approach analogous to Algorithm 1. In this approach, we solve a sequence of the problems

$$\begin{aligned} &\max_{z,t}\, t \\ &\text{s.t. } \Pi_n(z, c^i) \geq t, \quad i = 1, \ldots, k, \\ &\qquad z \in \mathfrak{Z}, \end{aligned} \tag{2.28}$$

where $c^i \in \mathfrak{C}_n^m(\epsilon)$ are generated coefficient cuts. Denote by (z_k^*, t_k^*), the optimal solution of problem (2.28). Note that t_k^* is also the optimal value of problem (2.28). If $t_k^* < 0$, it is verified that none of $z \in \mathfrak{Z}$ can satisfy $\pi_n^m(z;\epsilon) \geq 0$. Otherwise, we calculate $\pi_n^m(z_k^*;\epsilon)$ by solving problem (2.25), and obtain an optimal solution c_k^*. If $\pi_n^m(z_k^*;\epsilon) \geq 0$, z_k^* is the sought solution satisfying the condition; otherwise, the constraint $\Pi_n(z_k^*, c_k^*) \geq t$ is added into the master problem (2.28) as a valid cut. Algorithm 2 formally describes this bisection and cut-generation approach.

Algorithm 2 Bisection and Cut-Generation Algorithm for Model (2.27)

`Step 1` Choose $\gamma > 0$, $\underline{\epsilon} = 0$, and $\overline{\epsilon} = 1$.
`Step 2` If $\overline{\epsilon} - \underline{\epsilon} \leq \gamma$, then exit. Otherwise, let $\epsilon = (\overline{\epsilon} - \underline{\epsilon})/2$, and $k = 0$.
`Step 3` Find the optimal solution (z_k^*, t_k^*) of problem (2.28). If $t_k^* < -\gamma$, then let $\overline{\epsilon} = \epsilon$, and goto step 2.
`Step 4` Calculate $\pi_n^m(z_k^*;\epsilon)$ by solving problem (2.25). Let c_k^* be the optimal solution. If $\pi_n^m(z_k^*;\epsilon) \geq -\gamma$, then let $\underline{\epsilon} = \epsilon$, and goto step 2.
`Step 5` Let $c^{k+1} = c_k^*$ and $k = k + 1$. Go to step 3.

2.6 Case Study: Portfolio Investment

In this section we apply RSD-P to the portfolio optimization problem given in (Dentcheva & Ruszczynski, 2003) involving $N(= 8)$ assets: (S1) U.S. three-month treasury bills, (S2) U.S. long-term government bonds, (S3) S&P 500, (S4) Willshire 5000, (S5) NASDAQ, (S6) Lehmann Brothers corporate bond index, (S7) EAFE foreign stock index, and (S8) gold. This problem uses $M(= 22)$ yearly returns r_{ij} $(i = 1, \ldots, M,\ j = 1, \ldots, N)$ of these assets as equally probable realizations (See Table 2.2).

Using RSD-P to model this problem, we have

$$\max_{z} \left(1 + \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{N} z_j r_{ij}\right) \tag{2.29a}$$

$$\text{s.t.}\ \frac{1}{M}\sum_{i=1}^{M} u\left(1+\sum_{j=1}^{N} z_j r_{ij}\right) \geq \frac{1}{M}\sum_{i=1}^{M} u\left(1+\sum_{j=1}^{N} z_j^Y r_{ij}\right), \quad \forall u \in \mathfrak{U}^m(\epsilon), \tag{2.29b}$$

$$\sum_{j=1}^{N} z_j = 1, \tag{2.29c}$$

$$z_j \geq 0, \quad j = 1, \ldots, N. \tag{2.29d}$$

In the above problem, the objective (2.29a) seeks the best asset allocation to maximize the expected total wealth, while the RSD constraint (2.29b) requires that this allocation should be sufficiently preferred to the benchmark z^Y.

We discuss model (2.29) in four cases. In case (i), we consider the 2nd order RSD constraint by setting $m = 2$ in (2.29b). The option of investing all money in S1 is used as the benchmark, i.e., $z^{Y_1} := \{1, 0, \ldots, 0\}$. S1 is a risk-free asset, using which the RSD constraint (2.29b) guarantees the investment on risky assets to reach a given level of safety. We let the support $\Theta = [0, 2]$ and choose the CRRA utility function, $u^1_{ref}(x) := \sqrt{\frac{x}{2}}$, as the reference, which is consistent with Examples 1 and 2 of Hannah's purchasing lottery tickets. Case (i) is default in this study, and we adapt it to the other three cases: case (ii) substitutes the CARA reference utility function $u^2_{ref}(x) := \frac{e^{-x}-1}{e^{-2}-1}$; case (iii) uses an alternative benchmark, $z^{Y_2} := \left\{\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}\right\}$, which equally invests on every asset; and case (iv) discusses the 3rd order RSD constraint by letting $m = 3$. Table 2.3 summarizes the different configurations. In this study, we let the measure μ in condition (A3) be the uniform distribution on Θ.

Model (2.29) is approximated by the corresponding BSD-P. This study first tests the computational complexity of Algorithm 1 with $\gamma = 10^{-10}$, and next analyzes the model performance by adjusting dominance level ϵ. Finally, we discuss the application of model (2.29) to Hannah's investment.

Table 2.2: Asset Returns (in %) in (Dentcheva & Ruszczynski, 2003)

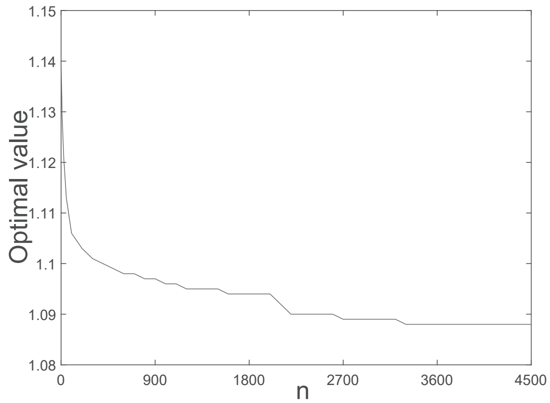
Year	S1	S2	S3	S4	S5	S6	S7	S8
1	7.5	-5.8	-14.8	-18.5	-30.2	2.3	-14.9	67.7
2	8.4	2.0	-26.5	-28.4	-33.8	0.2	-23.2	72.2
3	6.1	5.6	37.1	38.5	31.8	12.3	35.4	-24.0
4	5.2	17.5	23.6	26.6	28.0	15.6	2.5	-4.0
5	5.5	0.2	-7.4	-2.6	9.3	3.0	18.1	20.0
6	7.7	-1.8	6.4	9.3	14.6	1.2	32.6	29.5
7	10.9	-2.2	18.4	25.6	30.7	2.3	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

Table 2.3: Configurations in the Four Studied Cases

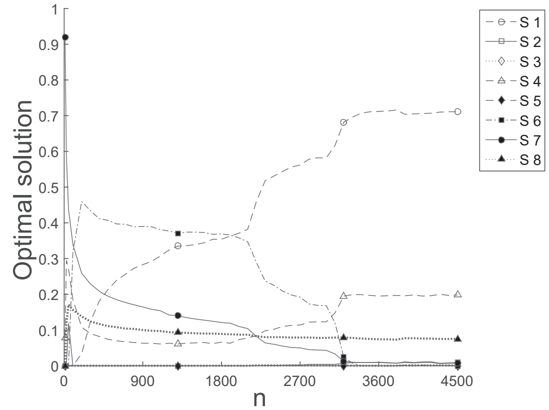
	Reference utility	Benchmark	RSD order
Case (i)	u^1_{ref}	z^{Y_1}	2
Case (ii)	u^2_{ref}	z^{Y_1}	2
Case (iii)	u^1_{ref}	z^{Y_2}	2
Case (iv)	u^1_{ref}	z^{Y_1}	3

2.6.1 Computational Analysis

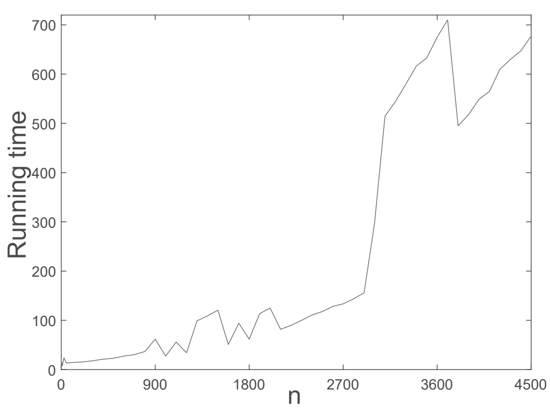
All experiments are conducted on a laptop with Intel Core i7 processor with 4 physical cores and hyper-threading on each core. The maximum frequency is 2.4 GHz with the boost at specific core up to 3.2 GHz. The maximum amount of RAM allowed for computation is 2 GB for each core. Step 2 in Algorithm 1 is coded using the optimization toolbox of Matlab R2015a, and step 3 is solved via the QCP solver of CPLEX 12.6. Both of the solvers work



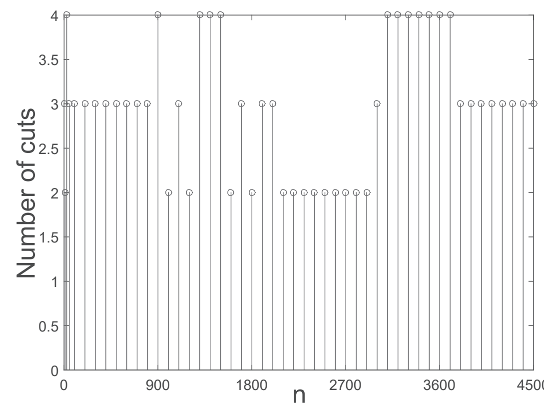
(a) Optimal Value



(b) Optimal Solution



(c) Runnning Time (in sec).



(d) Number of Generated Cuts.

Figure 2.1: Computational Results of Algorithm 1 for the Degree of BSD-P with u^1_{ref} and z^{Y_1}.

in parallel modes, which create 4 clusters for Matlab and 4 threads for CPLEX. Algorithm 1 is tested for cases (i)-(iv) with $\epsilon = 0.1$. We first discuss the performance of Algorithm 1 for case (i), and then compare the other cases.

Computational Results of Case (i). Figure A.1 reports the optimal values, optimal solutions, running times, and numbers of generated cuts when the degree n of BSD-P increases from 10 to 4500. Recall that the degree of BSD-P is the degree of Bernstein polynomials approximating the set $\mathfrak{U}^m(\epsilon)$. Shown in Figures 2.1a and A.1b, the optimal value and solution fluctuate at the low degrees, but become stable for the degrees larger than 3500. Theorem 2 implies that the approximate RSD constraint, $\pi^m_n(z) \geq 0$, in BSD-P can be regarded as a relaxation of the true counterpart, $\pi^m(z) \geq 0$, in RSD-P. As n increases, the effect of $\pi^m_n(z) \geq 0$ is enhanced, and the approximate solution is stabilized for the

Table 2.4: Computational Results of Cases (i)-(iv)

	Case (i)				Case (ii)			
n	1500	2500	3500	4500	1500	2500	3500	4500
Opt Sol	0.339	0.544	0.711	0.711	0.506	0.696	0.720	0.722
	0	0.001	0.009	0.009	0	0.008	0.001	0.005
	0	0.001	0	0	0.002	0.003	0.001	0
	0.064	0.110	0.197	0.198	0.117	0.192	0.200	0.195
	0	0	0	0	0	0	0	0
	0.374	0.209	0	0	0.220	0.011	0	0
	0.131	0.055	0.009	0.007	0.070	0.013	0.005	0.006
	0.091	0.080	0.075	0.075	0.084	0.076	0.072	0.070
Opt val	1.095	1.090	1.088	1.088	1.091	1.089	1.088	1.088
Running time	108.97	110.72	616.42	647.01	113.82	190.26	512.21	759.33
Num of cuts	4	2	4	3	4	3	3	3
	Case (iii)				Case (iv)			
n	1500	2500	3500	4500	1500	2500	3500	4500
Opt Sol	0	0	0	0	0.304	0.344	0.353	0.354
	0	0	0	0	0	0	0	0.001
	0	0	0	0	0	0	0	0
	0.225	0.240	0.251	0.279	0.061	0.067	0.074	0.078
	0	0	0	0	0	0	0	0
	0.349	0.355	0.373	0.362	0.388	0.383	0.389	0.397
	0.290	0.271	0.245	0.221	0.151	0.117	0.098	0.085
	0.136	0.133	0.130	0.138	0.096	0.089	0.086	0.084
Opt Val	1.112	1.112	1.111	1.110	1.096	1.094	1.094	1.093
Running time	121.84	367.44	569.98	616.48	86.47	293.96	584.74	687.84
Num of cuts	5	5	5	4	3	3	3	3

convergence to the true solution. This is observed in Figure A.1b that the investment in the optimal solution is highly diversified for a small n, while the risk-free asset S1, invested solely in the benchmark z^{Y_1}, gets an overwhelming proportion of investment when n is above 3500.

The running time of Algorithm 1 is related to not only the degree but also the number of generated cuts. Note that the degree decides the number of decision variables in BSD-P, and the number of generated cuts is the total iterations run by Algorithm 1. Figure 2.1c reflects the tendency that a longer running time is needed as the degree increases. However, the number of generated cuts is independent of the degree shown in Figure 2.1d. Particularly, Algorithm 1 generates 4 cuts when the degree varies in [3200, 3800], but there are only 3 cuts for the degree in [3900, 4500]. As the results, the running time reaches the peak, which is 710.11 seconds, for the degree is 3800, while it falls down to 494.94 seconds for the degree is 3900. Then the running time grows again to 647.01 seconds as the degree increases to 4500.

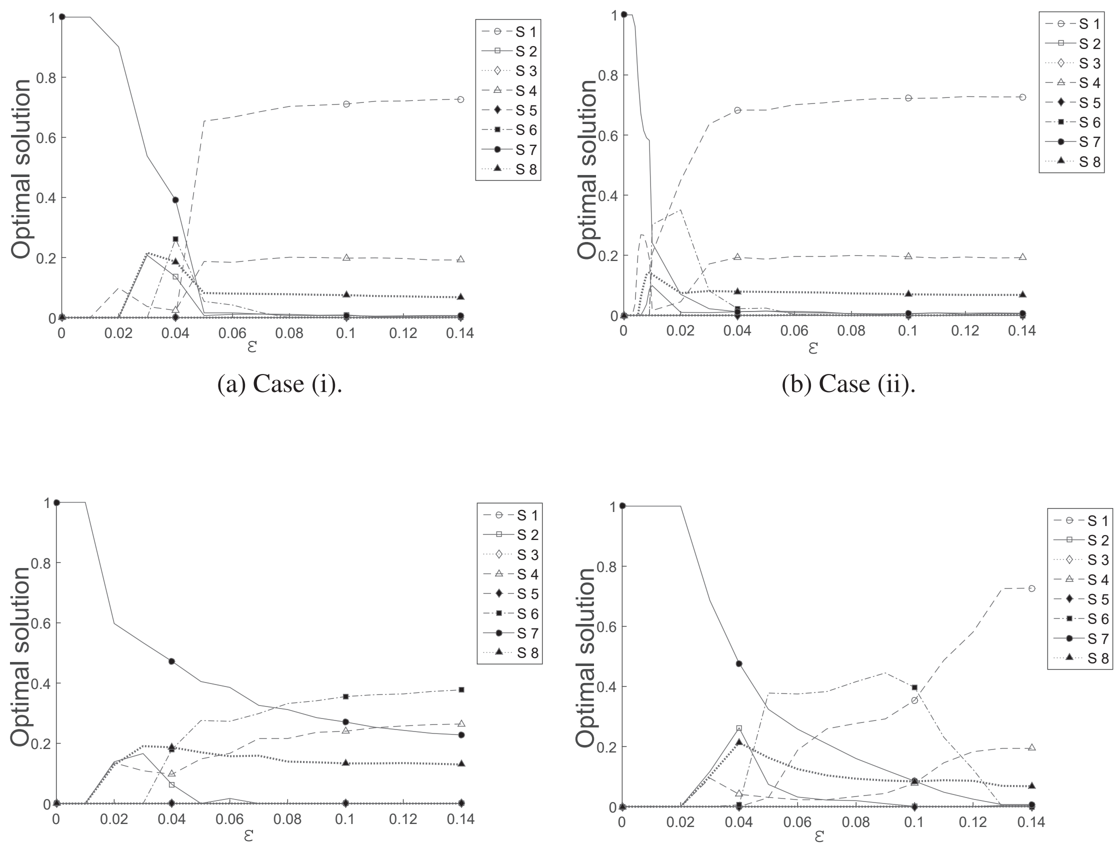
Comparison between Computational Results of cases (i)-(iv). Table 2.4 lists the optimal values, optimal solutions, running times, and numbers of generated cuts for cases (i)-(iv) when the degree n increases from 1500 to 4500. The four cases have similar running times and numbers of generated cuts. In average, Algorithm 2 runs 107.77, 240.73, 570.84, and 677.67 seconds, and generates 4, 3.25, 3.75, and 3.25 cuts, when n equals 1500, 2500, 3500, and 4500, respectively.

Cases (i) and (ii) require n to be at least 3500 in order to get the stable optimal values and optimal solutions, while cases (iii) and (iv) need n to be 1500. Cases (i)-(ii) use the RSD constraints with z^{Y_1}. These constraints tell that the decision maker strongly desires a low risk portfolio, in which S1 should be given an overwhelming proportion of investment. In comparison, z^{Y_2} uniformly invests money to all the assets, and this diversification makes the RSD constraint in case (iii) weaker than in cases (i)-(ii). Similarly, by Proposition 3, the 3rd order RSD constraint in case (iv) is also a relaxation of the 2nd order one in case (i). The test results in Table 2.4 show that it needs a large n to approximate the RSD constraints with z^{Y_1}. Indeed, for a small n, the approximate RSD constraint produces a highly diversified portfolio, which is unable to research the strong requirement on low risk. With the weaker RSD constraints, cases (iii) and (iv) do not need n to be very large.

2.6.2 Effect Analysis of the RSD Constraint

We now analyze cases (i)-(iv) to test the effect of the RSD constraint (2.29b). The results are given in Figures 3.3 and 2.3. In this test, the degree of approximation is 4500, and the dominance level ϵ is adjusted in [0, 0.14]. In each case we divide this interval into three sub-intervals — weak region, mild region, and strong region — due to the strength of the RSD constraint (2.29b). In general, ϵ in the weak region is very small such that the optimal value and solutions of model (2.29) are identical to ones given by only using the reference utility function (i.e., $\epsilon = 0$). Indeed, the RSD constraint (2.29b) has a limited impact on the performance of model (2.29) in the weak region. The strong region is opposite, for ϵ is rather large. The corresponding optimal value and solutions are very stable, and are indifferent to the reference utility function. In contrast, model (2.29) is sensitive to ϵ in the mild region. A small change on ϵ may incur completely different asset allocations in the

I



(a) Case (i).

(b) Case (ii).

(c) Case (iii).

(d) Case (iv).

Figure 2.2: Impact of RSD on Optimal Asset Allocations for Cases (i)-(iv).

Case (i). In this case, the RSD constraint (2.29b) compares risky asset allocations with the benchmark z^{Y_1}, which only invests the risk-free asset S1. Shown in Figures 2.2a and 2.3, the weak region is [0, 0.01], the mild region is (0.01, 0.08), and the strong region is [0.08, 0.14]. Model (2.29) with ϵ in the weak region suggests that S7 obtain 100% of the total investment and yields the highest expected total wealth 1.141. As we increase ϵ to the mild region, the investment is diversified. For example, at $\epsilon = 0.04$, the percentage of S7 in the portfolio dramatically decreases from 100% to 39.1%, while the percentages of S2, S4, S6, and S8 rise to 13.6%, 2.4%, 26.2%, and 18.6%, respectively. At $\epsilon = 0.06$, S1 becomes crucial in the portfolio, owning 70.3% of the total investment and overwhelming S7 of which the percentage reduces to 1.2%. These results reflect the fact that, for satisfying the sufficient preference over the benchmark, the RSD constraint (2.29b) requires a large

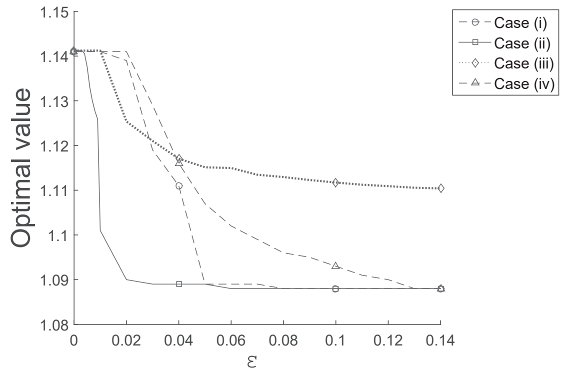


Figure 2.3: Expected Total Wealths of Cases (i)-(iv).

percentage of the total investment on the risk-free asset S1 to reduce the investment risk. As the effect of the RSD constraint (2.29b) is enhanced by increasing ϵ, S1 gets more percentage until the strong region is reached. In the strong region, the investment is stable at (72.7%, 0.5%, 0.1%, 19.2%, 0%, 0%, 0.7%, 6.8%), which is very close to the solution (72.7%, 0.4%, 0%, 19.3%, 0%, 0%, 0.7%, 6.8%) suggested by the classical second order stochastic dominance ($\epsilon = 1$). In addition, the decrease in the investment risk greatly reduces the expected total wealth, which rapidly decreases from 1.141 to 1.089 as ϵ changes from 0 to 0.05, and then slowly changes to 1.088. The over-conservativeness of stochastic dominance results in a very low yield, in comparison with the risk-free investment on S1 yielding the expected total wealth 1.078.

Case (ii). This case is designed to test the effect of the reference utility function in the RSD constraint (2.29b). We substitute the CARA utility function u^2_{ref} for the CRRA utility function u^1_{ref}. In contrast, for the total wealth more than 1, u^2_{ref} has a higher Arrow and Pratt's measure of risk-aversion than u^1_{ref}, i.e.,

$$-\frac{(u^2_{ref})''(x)}{(u^2_{ref})'(x)} = 1 > -\frac{(u^1_{ref})''(x)}{(u^1_{ref})'(x)} = \frac{1}{x}, \quad \text{for } x > 1. \tag{2.30}$$

Hence, in this study, u^2_{ref} characterizes stronger preference for low-risk investment than u^1_{ref}. It can be seen in Figures 2.2b and 2.3 that this substitution shrinks the weak region to [0, 0.003]. A subtle change on $\epsilon = 0$ has a big impact on the investment proportion and total wealth. Analogous to case (i), the investment is diversified to hedge the risk in the mild region (0.003, 0.08). However, case (ii) has a much faster diversification rate. The asset allocation at $\epsilon = 0.04$ is (68.3%, 1.2%, 0%, 19.3%, 0%, 2.2%, 1.2%, 7.9%), in which S1 has become a major invested asset, compared to 0% of the total investment on S1 in case (i). Also, this allocation is close to the stable solution, (72.7%, 0.5%, 0.1%, 19.2%, 0%, 0%, 0.7%, 6.8%), obtained in the strong region [0.08, 0.14]. Cases (i) and (ii) have the same asset allocation and total wealth in the strong region. This observation verifies that, for a sufficiently large ϵ, the RSD constraint (2.29b) is indifferent to the reference utility function, and approaches to the classical stochastic dominance.

Case (iii). In this case the equal allocation benchmark z^{Y_2} is substituted for the risk-free investment z^{Y_1}. Model (2.29) suggests a completely different investment policy without over-emphasizing the safety of investment. Shown in Figure 2.2c, S1 is not invested on, but S7 always obtains more than 22.8%. Figure 2.3 also indicates that the risky investment greatly raises the expected total wealth. In this case, the weak region is [0, 0.01] and the mild region is (0.01, 0.14]. Since the RSD constraint (2.29b) is ineffective in the weak region, the best policy is still 100% of the total investment on S7. As ϵ increases, the investment on S7 monotonically decreases, while S2, S4, S6, and S8 obtain more. The asset allocation is (0%, 6.3%, 0%, 9.7%, 0%, 18%, 47.2%, 18.7%) at $\epsilon = 0.04$, and changes to (0%, 0%, 0%, 26.4%, 0%, 37.7%, 22.8%, 13%) at $\epsilon = 0.14$. Different from case (i), there is not an overwhelming asset in case (iii), restricted by the benchmark z^{Y_2} where every asset is equally treated.

Case (iv). This case tests the 3rd order RSD. As indicated by Proposition 1, case (iv) with the 3rd order RSD constraint (2.29b) is the relaxation of case (i). As the result, shown in Figures 2.2d and 2.3, the weak region is enlarged to [0, 0.02], the strong region shrinks to [0.13, 0.14], and the diversification rate is much slower in the widely spanned mild region. Moreover, the curve of the total wealth in case (iv) is always above the curve in case (i). At $\epsilon = 0.04$, the asset allocation is (0%, 26.2%, 0%, 4.1%, 0%, 0.6%, 47.7%, 21.3%), in which case (iv) suggest 8.6% (= 47.7% - 39.1%) of the total investment on S7 more than case (i). At $\epsilon = 0.06$, the allocation is (18.8%, 3.2%, 0%, 2.3%, 0%, 37.5%, 25.8%, 12.5%), in which S7 still gets more investment than S1 and S6 has the largest percentage.

2.6.3 Hannah's Decision

We now discuss an investor's desired dominance level in this portfolio problem. Recall that Example 2 estimates Hannah's desired dominance level. To simply the discussion, we assume that Hannah suggests the same probability region for $\vartheta_1 = \$0.5$, $\vartheta_2 = \$1$, and $\vartheta_3 = \$1.5$. The example discusses the comparison between the lottery tickets $\bar{X}_i$, $i = 1, \ldots, 9$, and non-purchase option ϑ_2 in Table 2.1. We ask Hannah to select the lottery tickets which she is not reluctant to purchase or must not purchase, and then evaluate her desired dominance level for risky investment. In Example 2, Hannah selects $\bar{X}_1$, $\bar{X}_2$, $\bar{X}_3$, $\bar{X}_8$, and $\bar{X}_9$, and her desired dominance level is estimated to be

$$\left[\mathcal{E}^2(\bar{X}_3, \vartheta_2; \bar{u}_{ref}) + \mathcal{E}^2(\vartheta_2, \bar{X}_8; \bar{u}_{ref})\right]/2 = 0.040. \tag{2.31}$$

Since $\bar{u}_{ref}$ in Example 2 is u^1_{ref} used in this portfolio problem, we are able to use Hannah's desired dominance level to decide her best investment policy in case (i). Figure 2.2a shows that 0.040 is in the mild region of case (i). Her preferred asset allocation is (0%, 13.6%, 0%, 2.4%, 0%. 26.2%, 39.1%, 18.6%) and the expected total wealth is 1.111.

Suppose that Hannah is a veteran investor, and is able to accurately characterize her own risk attitude. Then under $\bar{u}_{ref}$, she will select all $\bar{X}_i$ ($i = 1, \ldots, 9$) in Table 2.1. Her desired dominance level is

$$\left[\mathcal{E}^2(\bar{X}_5, \vartheta_2; \bar{u}_{ref}) + \mathcal{E}^2(\vartheta_2, \bar{X}_6; \bar{u}_{ref})\right]/2 = 0.004, \tag{2.32}$$

which is in the weak region. It means that Hannah puts all investment on S7, and the expected total wealth rises to 1.141. Oppositely, Hannah has hesitating attitude. There may be an investment expert to design a utility assessment questionnaire to help characterize her risk preference. Notwithstanding Hannah's reference utility function is estimated to be $\bar{u}_{ref}$, she is not sure on her answers to the assessment questions. Actually, she shows interest in only purchasing $\bar{X}_1$ and $\bar{X}_2$, and insists in not purchasing $\bar{X}_9$. Such hesitating attitude could incur high ambiguity in the assessment procedure. Hence, her desire dominant level is suggested to be

$$\left[\mathcal{E}^2(\bar{X}_2, \vartheta_2; \bar{u}_{ref}) + \mathcal{E}^2(\vartheta_2, \bar{X}_9; \bar{u}_{ref})\right]/2 = 0.063. \tag{2.33}$$

This level is in the strong region. Her preferred asset allocation is (66.7%, 1.1%, 0%, 18.4%, 0%, 4.2%, 1.6%, 8%) and the expected total wealth decreases to 1.089. Recall that the expected total wealth is 1.078 for the risk-free investment.

2.7 Conclusions

This study has introduced a novel concept of reference-based almost stochastic dominance (RSD) and its application in risk averse optimization problems. In the $\mathcal{L}_2$-normed space, we have specified a subset of the general class of risk averse utility functions. This subset consists of nonparametric shape-preserving perturbations around a given reference utility function. The RSD represents a preference relation that a preferred uncertain prospect should have the larger expected utility over the perturbation subset. We have also defined the maximum dominance level, which quantifies the decision maker's preference between alternative choices in the context of robustness.

We have proposed the RSD constrained stochastic optimization model and studied its solution method. An approximation approach based on Bernstein polynomials has been developed. This approach resorts to a cut-generation algorithm. We have discussed the convergence rate of the approximate feasible region, shown the asymptotic convergence of the optimal value and the set of optimal solutions obtained in this approach, and proved that the algorithm has finitely many iterations. In addition, we have adapted the cut-generation algorithm to seek a valid option most robustly preferred to a random benchmark.

The portfolio optimization problem given in (Dentcheva & Ruszczynski, 2003) has been used to analyze the computational complexity of the approximation approach and to illustrate the effect of the RSD constraint. We have compared four cases with different benchmarks, reference utility functions, and dominance orders. In addition, we have discussed the impact of an investor's desired dominance level on asset allocations.

Chapter 3

Balance of Preference Maximization and Preference Ambiguity Sensitivity Analysis in Risk-Averse Decision Making

3.1 Introduction

Risk-averse stochastic optimization is a technique commonly used by the decision makers in risk management problems with uncertain outcomes. “Risk” is the possibility of change in the value of decision due to changes in the value of the underlying components on which that decision was made (McNeil et al., 2015). Obviously, risk is associated with uncertainties associated with outcomes. However, uncertain outcomes are not the only sources of the risk. The value of the decision can change also from the point of view of decision maker himself, due to preference ambiguity. To manage this risk sources, in this chapter we require the optimal decision to be robust with regards to both sources of risk- uncertainties in outcomes and decision maker’s preference ambiguity. We achieve robustness of the solutions by incorporating sensitivity analysis with regards to decision maker’s preference characteristics, in a classical expected utility maximization model.

Expected utility maximization is one of well implemented frameworks of risk-averse stochastic optimization. Applications of such utility maximization models can be found in almost any area of industry, including finance (Benidis et al., 2018; Bruni et al., 2014; Cornuejols & Tutuncu, 2006; Hu et al., 2011; Roman et al., 2013), transportation (Li et al., 2014; Nie et al., 2011), response systems (Noyan, 2010), energy management (Gollmer et al., 2008), etc... In this framework a rational decision maker’s risk preference is represented by a unique utility function (see Friedman & Savage, 1948; von Neumann & Morgenstern, 2007, and others in references). Thus, the risks associated to decision maker’s preference characteristics are entangled with his utility function. Classically, before solving the expected utility maximization model, the decision maker is tested to satisfy 3:A-3:C fundamental axioms of utility theory (see von Neumann & Morgenstern, 2007). If these axioms are satisfied, then the decision maker's preference can be mapped to a unique utility function. Then that utility function is assessed using some kind of parametric or non-parametric approach.

However, in expected utility maximization, due to preference ambiguity, may produce a solution which may produce a solution which may not be preferred by decision maker. We refer to such solutions as "sensitive with regards to preference ambiguity" (correspondingly, we will refer to solutions insensitive to preference ambiguity as "stable with regards to preference ambiguity"). Besides the aforementioned studies aimed at detecting the solutions sensitivity to preference of the decision maker, we present in section 3.1.3 of literature review several real world studies hindered by the problems arising from the (Chl 2) and (Chl 1) sources of risks.

The solution's sensitivity to preference is not a new problem in decision analysis (see Sage, 1981; Samson, 1988, comprehensive overview). But the sensitivity analysis is usually performed as a post-optimality step. In this chapter we go further, and suggest to incorporate the sensitivity analysis with respect to the decision maker's preference in the optimization framework, instead of cyclic application of sensitivity analysis of the solution in post-optimality stage. Suggested ideology is similar to the incorporation of the risk measure in the optimization framework suggested by Markowitz (1952). In our model we couple risk-aversion and the sensitivity analysis in bi-objective optimization model for a utility maximizer decision maker.

3.1.1 Balance of Preference Maximization and Sensitivity-Averseness

It is important to note, that despite the inconsistencies in preferences, the majority of decision maker's have "core" behavioral traits which they strive to follow, such as preference maximization and/or risk-averseness. In other words, a preference maximizing (and/or risk-averse) decision maker for any given choice will try to pick an option which has highest utility (and/or lowest outcome risk) from their point of view (Agranov & Ortoleva, 2017; Echenique et al., 2011; Loomes, 2005; Tversky & Kahneman, 1981). The emphasis being on “their point of view”, which might not be consistent.

Thus, in general, preferring options which maximize the decision maker’s preference is still logically correct and shall not be abandoned. The assessed utility function might not fully represent the decision maker, but it may capture the core properties of his behavior. To emphasize this, we refer to the assessed utility function as a decision maker’s “reference utility function”, which we denote as $u^{\#}$.

We accomplish the modeling of a preference maximizing decision maker using a classical method of having a preference relation as an objective function in stochastic optimization models. For a given decision maker the model tries to find an option X which will be maximally preferred by him to some bottom line option Y called “benchmark”. Examples of benchmarks are non-purchase (or certain outcome) options in optimal lottery purchase problems in decision analysis (Hu & Stepanyan, 2017; Leshno & Levy, 2002), given market index in enhanced indexation problems in finance (Benidis et al., 2018; Bruni et al., 2014; Cornuejols & Tutuncu, 2006), 100% investment in non-risky asset or equal investment in all assets in portfolio optimization problems (Armbruster & Delage, 2015; Dentcheva & Ruszczynski, 2003; Hu & Stepanyan, 2017), or the outcomes of placebo drug in health care problems (Giancristofaro & Bonnini, 2008).

According to Von Neumann–Morgenstern expected utility theory the preference of X to Y, denoted as $X \succeq_{u^{\#}} Y$, is defined as $\mathbb{E}[u^{\#}(X)] \geq \mathbb{E}[u^{\#}(Y)]$. We also denote the strict preference of these two options, $\mathbb{E}[u^{\#}(X)] > \mathbb{E}[u^{\#}(Y)]$, as $X \succ_{u^{\#}} Y$ and the equivalence, $\mathbb{E}[u^{\#}(X)] = \mathbb{E}[u^{\#}(Y)]$, as $X \simeq_{u^{\#}} Y$. With this notation, the objective function of an preference maximizer decision maker will have following form

$$\max_{X} X \succeq_{u^{\#}} Y. \tag{3.1}$$

At the same time, we argue that it is also logical for a decision maker to avoid solutions which are sensitive with regards to ambiguity his preference. We call this kind of behavior a “sensitive-averse”. To model a sensitive-averse decision maker we use a sensitivity measure of preference relation regarding decision maker’s reference utility function (referred as SMU and denoted by $\mathbb{M}$). The measure, that we use was first introduced in (Insua et al., 1997). It measures the maximal perturbation on the decision maker's reference utility function while keeping the preference of X to Y intact. We restrict it to preserve "core" behavioral patterns in perturbations and maximize its value in stochastic optimization model as a second objective function.

$$\max_{X} \mathbb{M}[X \succeq_{u^\#} Y]. \tag{3.2}$$

A decision maker's personal preferences may typically have lower maximal sensitivity levels and these two objective functions are in conflict with each other. Thus, instead of a single optimal value, we often get a whole set of Pareto-optimal frontier. The decision maker's final goal is to pick a solution which will balance his preference of X to Y and its sensitivity.

3.1.2 The Sensitivity Measure

In the utilitarian theory there are two major ways of modeling decision maker's choice inconsistencies: random utility theory (RUT) and random preference theory (RPT). RUT assumes that the decision maker makes his decisions according to his "true" utility function. But, the utility of any preference fluctuates by adding some random value. The RUT allows to model decision maker's core behavioral traits in the utility function. Additionally by controlling the variance of the random error term results in high level of accuracy (see Hess et al., 2018; Hey & Orme, 1994; McFadden, 1974). On the other hand, RPT assumes that when making decision decision maker can be described by some function from a functional set. Compared to RUT, it has an advantage of not imposing any probability distribution on error term, but it has a disadvantage of requiring computationally harder algorithms to solve (see Loomes, 2005, for a comprehensive comparison).

In our optimization models decision maker's choice inconsistencies are measured with a special sensitivity measure (SMU). This measure was first introduced by Insua et al. (1997) in the Bayesian statistical decision theory under the name of "$\epsilon_{u^\#}$-robust sensitivity measure". It models inconsistent decision maker similar to RUT and RPT. It borrows from RUT the ideas of having a core behavioral traits encoded in reference utility function and having a fluctuation on it. But unlike RUT, the reference utility function fluctuations are constructed as a convex sum of it and a function from functional set $\mathfrak{V}$, similar to RPT. The functional set $\mathfrak{V}$ encodes core behavioral traits as decision maker's reference utility function. But since those behavioral traits are very basic fundamentals (such as risk attitude or preference maximization), it contains all the functions satisfying them. To illustrate the idea of the construction of functional set $\mathfrak{V}$, we present bellow simple example similar to example 1 from Insua et al. (1997).

EXAMPLE 3. *A simple functional set $\mathfrak{V}$ based on assessed finite utility points.*

Suppose that decision maker's reference utility function is assessed using some classical non-parametric standard utility assessment questionnaire, which produced following initial utility assessments:

$$\begin{aligned} &u^{\#}(0) = 0, u^{\#}(0.38) = 0.5, u^{\#}(0.58) = 0.7, \\ &u^{\#}(0.7) = 0.8, u^{\#}(0.84) = 0.9, u^{\#}(1) = 1. \end{aligned} \tag{3.3}$$

From these values we can conclude, that the decision maker is a preference maximizer ($u^{\#}$ is increasing) with a risk-averse attitude ($u^{\#}$ is concave). Thus the set $\mathfrak{V}$ can be constructed as follows

$$\mathfrak{V} = \{v : v \text{ is increasing concave}, v(0) = 0, v(1) = 1\} \tag{3.4}$$

□

The SMU uses functions from $\mathfrak{V}$ set to perturb decision maker's reference utility function $u^{\#}$. And similar to classical sensitivity measures, the SMU measures the maximal amount of that perturbation on the reference utility function the given preference relationship of option X to Y can withstand.

3.1.3 Literature Review

Suggested in this chapter bi-objective model incorporating risk-aversion and sensitivity-aversion in one optimization framework to our knowledge is a novel approach of solving problems (Chl 1) and (Chl 2). However, the problems (Chl 1) and (Chl 2) themselves are not new in stochastic optimization.

Suggested by Dentcheva & Ruszczynski (2003) optimization models with stochastic dominance constraints can be viewed as a simplistic way of solving these problems. In these models the optimal outcome is required to be preferred to benchmark for all possible decision makers with specific core behavioral trait. For the first order stochastic dominance the core trait is the preference maximization, for the second it is risk-averseness towards the outcomes. It is obvious that optimal solutions of such models are extremely stable with regards to preference ambiguity, as the problems (Chl 1) and (Chl 2) are eliminated by considering all possible utility functions describing corresponding core behavioral trait.

The shortcoming of this approach is its over conservativeness and lack of individualism. By requiring for the solution to be preferable for all possible preference maximizer (for second order also risk-averse) decision makers, these models completely lose any individualism encoded in decision maker's reference utility function (see Armbruster & Delage, 2015; Hu & Mehrotra, 2014; Hu & Stepanyan, 2017). Additionally, this models are well known of considering utility functions describing rather inhumane behaviors (see Hu et al., 2014; Leshno & Levy, 2002; Lizyayev & Ruszczyński, 2012). The aforementioned and other studies from the stochastic optimization with almost stochastic dominance field try to resolve over conservativeness and/or lack of decision maker's individualism, by limiting considered utility functions in the stochastic dominance constraints. However, it also means that problems (Chl 1) and (Chl 2) can arise again and make the solutions sensitive to decision maker's preference ambiguity.

Another approach of solving problems (Chl 1) and (Chl 2) in stochastic optimization is the sensitivity analysis of optimal solution with regards to decision maker's reference utility function. The sensitivity analysis is usually performed in post-optimality stage, where the optimal solution is tested to be stable with regards to preference ambiguity by perturbing different parameters of the stochastic optimization models and solving them again in a cyclic iterative decision process (French & Insua, 1989; Insua & French, 1991; Sage, 1981; Samson, 1988, ,etc...).

In the majority of the literature, for multiattribute decision models sensitivity analysis with regards to the decision maker's utility function is performed on aggregation weights only. This is due to computational efficiency of the solution methods, compared to sensitivity analysis performed with regards to general function forms. Alexander (1989) presents five measures of sensitivity designed for solution ranking of complex decision making models. The study tries to quantify the sensitivity of the solutions of the multi-criteria, multiattribute decision models with respect to the weights in additive objective function. Four of suggested measures are designed to use the ranking numbers, while one- the "Alexander's S", uses the utility of random outcome in the ranking. Similar to Samson (1988) authors suggest to use iterative process for complex decision models, with several runs of the model and sensitivity analysis of the solution. Barron & Schmidt (1988) provide two models for the computation of weights, which change the current solution to another while satisfying specific distance based criteria. First method is based on maximum entropy principle and produces results which tend to be equally distributed, while the second is based on least squares distance, which produces results as close as possible to the original weights. The study provides closed form formulas for first and in special cases for the second model. Triantaphyllou & Sánchez (1997) uses sensitivity analysis to find the most important attributes and criteria in deterministic multi-criteria decision models. The study defines the sensitivity of the criteria as the reciprocal of the smallest percent by which its weight must change to induce change in the existing ranking. Similarly the sensitivity of the alternative in terms of a criterion in a given ranking is defined the reciprocal of the smallest percent by which its value in that criterion must change to induce change in its rank in the ranking.

Obviously, the shortcoming of such approaches are the lack of sensitivity analysis on the functional forms of final utility function.

Sensitivity analysis done with regards to general form of utility functions is much more scarce in the literature. Even though the necessity of consideration of which is mentioned in almost all guideline studies (Insua & French, 1991; Sage, 1981; Samson, 1988, , etc...).

In (French & Insua, 1989; Insua & French, 1991) suggested analyzing the sensitivity of the ranking with respect to some parameter in objective function in generalized framework of sensitivity analysis in discrete multi-objective decision making. Similar to Samson (1988) in these studies the decision analysis is a repetitive cycle of stages of judgmental elicitation, computation and sensitivity analysis. In the sensitivity analysis phase, it is assumed that the decision maker cannot provide further information in the elicitation phase, without further insight and reflection. Sensitivity analyses is viewed as a means of stimulating him to think more closely about the problem. The analysis tries to find other possibly optimal alternatives by varying a parameter in objective function. Unlike the other studies this parameter is not bound to be an aggregation weight. The study tries to quantify the sensitivity of the chosen alternative on the basis of this variation. If the small variation in a parameter induces change then the solution is considered sensitive. The minimal amount of variation then is taken as absolute measure of sensitivity of the alternative with respect to that parameter. The absolute sensitivity value of the chosen alternative can be normalized by dividing it by the sensitivity of the most stable alternative.

Based on this foundation Insua et al. (1997) presented the generalized sensitivity measure for a solution in decision analysis with respect to preference of decision maker using a utility set. They measure maximum perturbation of the given preference relation in some functional set. The parameter controls the linear contamination of decision maker's assessed utility function. This method is well known as ϵ-contamination in Bayesian sensitivity analysis (see Berger et al., 2000). The idea is to construct a perturbation neighborhood around that parameter value and solve the problem based on common ground. If the conclusions agree then the parameter is robust (see Berger et al., 1994). However, in Bayesian sensitivity and decision analysis the most work has concentrated on investigations about sensitivity with respect to the prior. And even though, the necessity of sensitivity analysis with respect to both the prior and the utility function is notes in several studies such as (Insua et al., 1997; Smith, 1994), to our knowledge Insua et al. (1997) is the only paper which addresses this problems. This is probably, due to the disconnect between the Bayesian Statistics and statistical decision theory mentioned in Liese & Miescke (2007), or due to inherit computational difficulty of the algorithms of non-parametric stochastic optimization with functional sets.

In their study, Insua et al. (1997) note that computations of robust neighborhoods are difficult in the general case of imprecision in both the utility and the prior. They present only upper and lower bounds for the computation of sensitivity measure. In this study we incorporate it in two bi-objective optimization problems, then by using some linearization methods we were able to reformulate them as a mixed-integer and linear programming models.

Examples of the preference inconsistencies in the field.

The problems (Chl 1) and (Chl 2) have been constantly registered throughout the years in various real studies in decision making. Below we present some of our finding.

In 2016 Dean & Martin (2016) analyzed the data of grocery purchases recorded by Aguiar & Hurst (2007) for the 1993-1995 years in the Denver metropolitan area retail outlets. It was found, 71% of 977 representative households have inconsistent results in the their decisions of purchases. Similarly, in 2011 Echenique et al. (2011) analyzed widely used Stanford Basket Dataset containing grocery purchases from four stores in an urban area of a large U.S. mid-western city for the 1991-1993 years. It was found that 80% of unique 494 households purchase decisions cannot be described with classical single utility functions because of inconsistencies.

In the book of Decision Making in Health Care Chapman & Elstein (2013) presented an overview of several examples of preference inconsistencies. Particularly in 1997 Stalmeier et al. (1997) conducted three experiments to test the preferences of in total of 176 students of the University of Nijmegen in Netherlands. In the first experiment 33 university students and 17 high school students were selected by the condition that they preferred living 25 years to living 50 years with metastasized breast cancer. After being selected, women were asked how many years in good health they consider equivalent to the outcomes. In contrary the classical expected utility theory, 23 of the 33 university students and 14 of the 17 high school students assigned higher value to 50 years. Since the students were informed about the potency of the metastasized breast cancer, it was assumed that living 25 or 50 years with metastasis may not be realistic. Which could have impacted students decisions. Thus in the second experiment the breast cancer was replaced with much realistic health state of having chronic migraines on average of 4.5 days per week. 16 students were selected by the condition that they have preferred living 10 years to living 20 years with such condition. And again students were asked how many years in good health they consider equivalent to the outcomes. Out of 16 students 15 assigned a larger values to the option of 20 years, therefore again presenting inconsistent preferences. To test the nature of inconsistent preferences in the third experiment 27 students were confronted with the preference reversals and asked whether they wanted to change their response patterns. All 27 students indicated that they understand that their answers were inconsistent and only 4 of 27 participants undid their original preference reversals. Finally, to test the possibility of inconsistent preferences arising because of inexperience of the students with health states in the questionnaires in 2001 Stalmeier et al. (2001) published results of interviews from June 1996 until May 1997 conducted with the patients from the Neurology Department outpatient clinic at the University of Illinois Hospital in Chicago. All patients had a history of migraines and were thus highly familiar with the health state of living with migraines. 10 out of 14 patients were selected by the condition that they have preferred living 10 years to living 20 years with chronic migraines for 5.9 days per week. Of the 10 patients with such preferences, seven patients exhibited a preference reversal and after confrontation about it only 2 changed their answers. As a result (Stalmeier et al., 2001, 1997) conclude that preference reversals within a single dimension suggests irrationalities at a more fundamental level.

3.1.4 Organization

This chapter is structured the following way. Section 3.2 pretenses the sensitivity measure, it's properties in economic theory, the Frechet derivative and the idea of sensitivity aversion. Section 3.3 presents the multi-objective multi-attribute optimization model and it's properties for additively independent attributes. Section 3.4 presents the tractable reformulations of the general model for increasing and increasing, concave utility functions. Section 3.5 presents the application of constructed model for Homeland Security budget allocation study.

3.2 Sensitivity Measure of Utility (SMU)

Consider two N-dimensional random vectors $\bar{X}, \bar{Y} \in (\Omega, \mathcal{F}, P; \Theta)$ with the support $\Theta \subseteq \mathbb{R}^N$. We write $\bar{X} = (X_1, \ldots, X_N)$ and $\bar{Y} = (Y_1, \ldots, Y_N)$ where X_i and Y_i are the ith components of $\bar{X}$ and $\bar{Y}$. As presented in section 3.1.2, in this study we use the sensitivity measure described in (Insua et al., 1997). They propose the ϵ_u-robust measure allowing perturbations in both the utility function and the prior distribution. The perturbations are linearly added on the current assessment within classes modeling imperfect judgment. In this paper we restrict our discussion to such utility function, while assuming that the probability distribution of $\bar{X}$ and $\bar{Y}$ is known. In the following definition of the ϵ_u-robust measure the reference utility function $u^{\#}$ is linearly perturbed within the set $\mathfrak{V}$ representing ambiguity in utility assessment.

DEFINITION 3. The *sensitivity measure* of the preference relation $\bar{X} \succeq_{u^{\#}} \bar{Y}$ regarding the reference utility function $u^{\#}$ is

$$\mathbb{M}[\bar{X} \succeq_{u^{\#}} \bar{Y}] := \max\{\epsilon \mid \bar{X} \succeq_u \bar{Y},\ u \in \mathfrak{U}(u^{\#}, \epsilon),\ \epsilon \in [0,1]\}, \tag{SMU}$$

where

$$\mathfrak{U}(u^{\#}, \epsilon) := \{u \mid u = (1-\epsilon)u^{\#} + \epsilon v,\ v \in \mathfrak{V}\}. \tag{3.5}$$

The sensitivity measure quantifies the maximum amount of perturbation of $u^{\#}$ that the preference relation can withstand.

REMARK 2. If $\bar{Y}$ is preferred to $\bar{X}$ by the decision maker's reference utility function $u^{\#}$, the set $\{\epsilon \mid \bar{X} \succeq_u \bar{Y},\ u \in \mathfrak{U}(u^{\#}, \epsilon),\ \epsilon \in [0,1]\}$ is empty. By Definition 3, $\mathbb{M}[\bar{X} \succeq_{u^{\#}} \bar{Y}]$ is the objective value of the infeasible maximization problem equal to $-\infty$.

Remark 2 completes the definition of SMU mathematically; however, it is meaningless for the decision maker to be concerned about the level of sensitivity of an unfavorable relationship. In the late statement we limit our discussion to the case of $\bar{X} \succeq_{u^{\#}} \bar{Y}$ while addressing the properties of SMU.

3.2.1 Properties of SMU

We now investigate the properties of SMU. It straightforwardly follows by definition 3 that an equivalent representation of SMU is

$$
\begin{aligned}
\mathbb{M}[\bar{X} \succeq_{u^\#} \bar{Y}] = \max \quad & \epsilon \\
\text{subject to} \quad & (1-\epsilon)\mathbb{E}[u^\#(\bar{X}) - u^\#(\bar{Y})] + \epsilon \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})] \geq 0, \\
& \epsilon \in [0,1].
\end{aligned} \tag{3.6}
$$

Insua et al. (1997) give an alternative formulation of (3.6) in the following proposition.

PROPOSITION 4 (Insua et al. (1997)).

$$
\mathbb{M}[\bar{X} \succeq_{u^\#} \bar{Y}] = \min \left\{ 1,\ \frac{\mathbb{E}[u^\#(\bar{X}) - u^\#(\bar{Y})]}{\mathbb{E}[u^\#(\bar{X}) - u^\#(\bar{Y})] - \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})]} \right\}. \tag{3.7}
$$

We observe that, if $\bar{X} \succeq_{u^\#} \bar{Y}$ and there exists $v \in \mathfrak{V}$ such that $\bar{Y} \succ_v \bar{X}$,

$$
\begin{aligned}
\mathbb{M}[\bar{X} \succeq_{u^\#} \bar{Y}] &= \frac{\mathbb{E}[u^\#(\bar{X}) - u^\#(\bar{Y})]}{\mathbb{E}[u^\#(\bar{X}) - u^\#(\bar{Y})] - \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})]} \\
&= \frac{1}{1 + \frac{\max_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{Y})] - \mathbb{E}[v(\bar{X})]}{\mathbb{E}[u^\#(\bar{X})] - \mathbb{E}[u^\#(\bar{Y})]}}.
\end{aligned} \tag{3.8}
$$

By the second equality in (3.8), SMU can be explained as the normalized ratio of the maximal level of perturbation given by $\max_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{Y})] - \mathbb{E}[v(\bar{X})]$ and the level of the preference of $\bar{X}$ to $\bar{Y}$ with respect to $u^\#$. We also derive the following conditions for SMU to catch two boundary values.

(i) $\mathbb{M}[\bar{X} \succeq_{u^\#} \bar{Y}] = 1$ if and only if $\bar{X} \succeq_v \bar{Y}$ for all $v \in \mathfrak{V}$.

(ii) $\mathbb{M}[\bar{X} \succeq_{u^\#} \bar{Y}] = 0$ if and only if $\bar{X} \simeq_{u^\#} \bar{Y}$ and there exists $v \in \mathfrak{V}$ such that $\bar{Y} \succ_v \bar{X}$.

In case (i), $\bar{X}$ is preferred $\bar{Y}$ over the entire set $\mathfrak{V}$, which is the necessary weak preference relation called by Greco et al. (2007). Moreover, this is the first (second) order stochastic dominance if $\mathfrak{V}$ consists of all increasing (concave) utility functions (Bawa, 1975; Hadar & Russell, 1969; Hanoch & Levy, 1975; Müller & Stoyan, 2002; Shaked & Shanthikumar, 1994). Oppositely, case (ii) tells that the ranking of $\bar{X}$ and $\bar{Y}$ is very sensitive with regards to preference ambiguity, if they are indifferent regarding $u^\#$ but not over the entire set $\mathfrak{V}$.

3.2.2 Effect of the Reference Utility Function on SMU

The inconsistency in utility assessment and the decision maker's mutable attitude bring on the ambiguity in characterizing his risk preference. In this section we quantifies the impact of the reference utility function $u^{\#}$ on SMU. To write more concisely, we define two functionals on $\mathcal{L}_2$ for given random vectors $\bar{X}$ and $\bar{Y}$ as

$$\langle \mathcal{F}, u \rangle := \mathbb{E}[u(\bar{X}) - u(\bar{Y})], \qquad \langle \mathcal{M}, u \rangle := \mathbb{M}[\bar{X} \succeq_u \bar{Y}], \qquad \text{for } u \in \mathcal{L}_2, \tag{3.9}$$

and a function as

$$N(x) := \frac{x}{x - \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})]}. \tag{3.10}$$

The theorem below presents the Frechet derivative of SMU with respect to $u^{\#}$. This derivative gouges the rate of the change on SMU resulting from a linear perturbation on $u^{\#}$.

THEOREM 5. Denote by $\left\langle D^{\mathcal{M}}_{u^{\#}}, h \right\rangle$ and $\left\langle D^{\mathcal{F}}_{u^{\#}}, h \right\rangle$ the Frechet derivatives of functionals $\mathcal{M}$ and $\mathcal{F}$ with respect to $u^{\#}$ acting on the direction $h \in \mathcal{L}_2$, respectively. If (i) $\bar{X} \succeq_{u^{\#}} \bar{Y}$ and (ii) there exists $v \in \mathfrak{V}$ such that $\bar{Y} \succ_v \bar{X}$, then

$$\left\langle D^{\mathcal{M}}_{u^{\#}}, h \right\rangle = N' \left(\langle \mathcal{F}, u^{\#} \rangle \right) \left\langle D^{\mathcal{F}}_{u^{\#}}, h \right\rangle, \tag{3.11}$$

where

$$N' \left(\langle \mathcal{F}, u^{\#} \rangle \right) = - \frac{\min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})]}{\left(\langle \mathcal{F}, u^{\#} \rangle - \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}) - v(\bar{Y})] \right)^2}, \tag{3.12}$$

and

$$\left\langle D^{\mathcal{F}}_{u^{\#}}, h \right\rangle = \mathbb{E}[h(\bar{X}) - h(\bar{Y})]. \tag{3.13}$$

Proof: Equation (3.8) indicates that

$$\left\langle \mathcal{M}, u^{\#} \right\rangle = N \left(\langle \mathcal{F}, u^{\#} \rangle \right). \tag{3.14}$$

By the definition of the Frechet differentiability (Cheney (2001); Clarke (2013)), the functional $\mathcal{F}$ is Frechet differentiable on the direction h if there exists a bounded linear operator

$D^{\mathcal{F}}_{u^\#}$ such that

$$\lim_{\|h\|_{\mathcal{L}_2}\to 0} \frac{\left|\left\langle \mathcal{F}, u^\# + h\right\rangle - \left\langle \mathcal{F}, u^\#\right\rangle - \left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle\right|}{\|h\|_{\mathcal{L}_2}} = 0. \tag{3.15}$$

In that case, $D^{\mathcal{F}}_{u^\#}$ operator is unique (Theorem 1 in Cheney, 2001, pg. 116) and $\left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle$ is the Frechet derivative of $\mathcal{F}$ with respect to $u^\#$ acting on the direction h. It is easy to verify that the limit in the equation above is 0 when substituting $\mathbb{E}[h(\bar{X}) - h(\bar{Y})]$ for $\left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle$. Conditions (i) and (ii) guarantee the existence of the derivative $N'\left(\left\langle \mathcal{F}, u^\#\right\rangle\right)$ given by (3.12). Since the derivatives in equations (3.12) and (3.13) exist, equation (3.11) follows from the chain rule theorem of Frechet derivatives (See Theorem 1 in Cheney, 2001, pg. 121). □

The value $\left\langle \mathcal{F}, u^\#\right\rangle$ quantifies the level of the decision maker's individual preference of $\bar{X}$ to $\bar{Y}$ according to his taste $u^\#$, while $\left\langle \mathcal{M}, u^\#\right\rangle$ measures the sensitivity of this preference relationship. Their Frechet derivatives, $\left\langle D^{\mathcal{M}}_{u^\#}, h\right\rangle$ and $\left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle$, are the rates of the changes corresponding to the variation in his taste. The derivative $N'\left(\left\langle \mathcal{F}, u^\#\right\rangle\right)$ in the chain rule (3.11) links the two changes.

REMARK 3. Condition (ii) in Theorem 5 guarantee that $N'\left(\left\langle \mathcal{F}, u^\#\right\rangle\right) > 0$.

By Remark 3, $\left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle$ and $\left\langle D^{\mathcal{M}}_{u^\#}, h\right\rangle$ have the same sign. A change on the individual preference with $\left\langle D^{\mathcal{F}}_{u^\#}, h\right\rangle \geq 0$ implies that the decision maker becomes more confident with his choice. The higher confidence increases the sensitivity measure against perturbation, i.e., $\left\langle D^{\mathcal{M}}_{u^\#}, h\right\rangle \geq 0$.

3.3 Multi-Objective Optimization Model Using SMU

Expected utility maximization is a broadly-implemented risk-averse stochastic programming technique (Fishburn, 1979; Vickson & Ziemba, 2014; von Neumann & Morgenstern, 2007; Yan, 2018), which suggests an optimal solution due to the decision maker's risk preference. We propose a bi-objective optimization model to maximize not only the expected utility but also the sensitivity measure of the utility (ESM) as

$$\begin{aligned} \max \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})] \\ \max \quad & \mathbb{M}[\bar{X}(z) \succeq_{u^{\#}} \bar{Y}] \\ \text{subject to} \quad & z \in \mathfrak{Z}, \end{aligned} \tag{ESM}$$

where $\mathfrak{Z} \subseteq \mathbb{R}^G$ is a decision region, $\bar{X} : \mathfrak{Z} \mapsto (\Omega, \mathcal{F}, P; \Theta)$ represents a random outcome due to the decision, and $\bar{Y} \in (\Omega, \mathcal{F}, P; \Theta)$ is a given benchmark. In model ESM the first objective is to maximize expected utility approaching the risk arising in systemic stochasticity, while the second is to maximize the SMU for reducing the impact of the ambiguity of the reference utility.

When observed with equation (3.6), it can be seen that the SMU in the second objective function maximally perturbs the excess expected utility value from the first objective function in ESM model. The perturbation is achieved with a utility function from the set $\mathfrak{V}$ producing the worst preference of $\bar{X}(z)$ to $\bar{Y}$. This observation allows us to draw various use cases of model ESM, based on the construction of perturbation set $\mathfrak{V}$.

In this work we focus on cases, where the perturbation set $\mathfrak{V}$ is constructed to contain all the utility functions satisfying the core risk-attitude characteristics of a decision maker. Then the model ESM will produce Pareto-optimal solutions each of which results to a random performance maximally preferred to $\bar{Y}$ under the reference utility function $u^{\#}$, while having the best stability with regards to preference ambiguity. An example of such perturbation set for a risk-averse decision maker can be the following

$$\mathfrak{V} := \{v : v \text{ is increasing concave}\}. \tag{3.16}$$

In that case ESM model can also be viewed as a balancing framework between risk-aversion and sensitivity-aversion. We discuss this property of ESM model in detail in the section 3.3.1.

Other than the problems with preference ambiguity, ESM model can be used in social decision making problems. As an example the perturbation set $\mathfrak{V}$ can be constructed to contain all the utility functions of the public sector of interest. In that case the SMU value of second objective function can be used as a measure of the public's agreement relative to the decision maker's opinions.

Finally, ESM model can be used in group decision making problems, where the decision maker has multiple advisors. In those cases the perturbation set $\mathfrak{V}$ can be constructed to contain the assessed utility functions of the advisors. Similar to social decision making problems, in this scenarios the optimal solutions of ESM model will be maximally preferred by the decision maker, corresponding to different relational values of agreement between the advisors.

The derivative forms from section 3.2.2, allow to employ ESM model for hybrid problems concerning both preference ambiguity and group decision making. To illustrate this feature, in the section 3.5.3 of the case study we solve a group decision making for the decision maker who is being advised by two advisors. Unlike the approach suggested above, the ESM model is still used to handle the preference ambiguity. However, the decision maker is required to not only present a solution acceptable by advisors, but also to ensure its approximate stability with regards to their preference ambiguity. In this case the problem was solved by linearly approximated SMU values of the advisors, using the Frechet derivatives from section 3.2.2.

3.3.1 Risk-Aversion vs. Sensitivity-Aversion

The Von Neumann-Morgensetern theorem represents the decision maker's risk preference as a utility function. On this basis, Expected utility maximization hedges the risk resulting from systemic randomness (Fishburn, 1979; Vickson & Ziemba, 2014). It is traditionally assumed that utility functions are deterministic and choice are consistent (von Neumann & Morgenstern, 2007). However, these assumptions are persistently violated in practice (see Agranov & Ortoleva, 2017; H. Simon, 1955; Tversky, 1969, and others in the references). Indeed, it is usually observed that the decision maker exhibits mutable and ambivalent preference, particularly without sufficient knowledge of complex problems and uncertain environment in the real world (Agranov & Ortoleva, 2017; Dean & Martin, 2016; Echenique et al., 2011). The ambiguity in utility representation is also concomitant with risk. The maximization of SMU, the second objective in ESM, provide a sensitivity-averse approach where the risk arising in the ambiguity is measured by SMU and controlled by its maximization.

We now give a lottery selection example to illustrate difference between risk-aversion and sensitivity-aversion. Let lottery X value \$2 with probability of 0.8 and \$0 with probability of 0.2 and lottery Y be certain \$1. Let a hesitant decision maker's reference utility function be

$$u^{\#}(\theta) = \frac{1-e^{-\theta}}{1-e^{-2}}, \qquad \theta \in [0,2], \tag{3.17}$$

and also give his alternative preference as

$$\hat{u}(\theta) = \frac{1-e^{-2\theta}}{1-e^{-4}}, \qquad \theta \in [0,2]. \tag{3.18}$$

Move $u^{\#}$ towards $\hat{u}$ linearly and observe the changes on the level of risk-aversion and sensitivity-aversion. Denote a directional function as

$$d(\theta) := \hat{u}(\theta) - u^{\#}(\theta), \tag{3.19}$$

by which the linear movement can be expressed as

$$u(\theta;\delta) = u^{\#}(\theta) + \delta d(\theta), \tag{3.20}$$

for some $\delta \in [0,1]$. The level of risk-aversion of $u(\theta;\delta)$ is determined by the Arrow-Pratt absolute risk-aversion coefficient is

$$r^d_{u^{\#}}(\theta,\delta) := -\frac{u''(\theta;\delta)}{u'(\theta;\delta)}. \tag{3.21}$$

The utility function $\hat{u} = u(.;1)$ is more risk averse than $u^{\#} = u(.;0)$ since $r^d_{u^{\#}}(\theta,1) = 3 > r^d_{u^{\#}}(\theta,0) = 1.5$ for all $\theta \in [0,2]$. The movement incurs increase in the level of risk-aversion of $u(\theta;\delta)$ since

$$\frac{\partial}{\partial\delta} r^d_{u^{\#}}(\theta,\delta) = \frac{\hat{u}'(\theta)u^{\#\prime\prime}(\theta) - \hat{u}''(\theta)u^{\#\prime}(\theta)}{(u^{\#}(\theta)+\delta d(\theta))^2} \geq 0, \qquad \text{for all } \theta \in [0,2]. \tag{3.22}$$

The change on the decision maker's risk preference varies the selection of the lotteries X and Y. X is preferred regarding $u^{\#}$ since

$$\mathbb{E}[u^{\#}(X)] - \mathbb{E}[u^{\#}(Y)] = 0.8 - 0.7311 = 0.0689, \tag{3.23}$$

while Y is better regarding $\hat{u}$, which is more risk averse, for

$$\mathbb{E}[\hat{u}(X)] - \mathbb{E}[\hat{u}(Y)] = 0.8 - 0.8176 = -0.0176. \tag{3.24}$$

We next observe the change of SMU in the movement. In this example, let the perturbation set $\mathfrak{V}$ consist of all risk-averse (increasing concave) functions. Letting the variation $h = \delta d$, we have

$$\begin{aligned}\mathbb{E}[h(X) - h(Y)] &= \delta(\mathbb{E}[\hat{u}(X) - u^{\#}(X)] - \mathbb{E}[\hat{u}(Y) - u^{\#}(Y)]) \\ &= -0.09\delta < 0, \qquad \text{for all } \delta \in (0, 1]. \end{aligned} \tag{3.25}$$

It follows by Theorem 5 and Remark 3 that $\left\langle D^{\mathbb{M}}_{u^{\#}}, h \right\rangle < 0$. Hence, the level of sensitivity-aversion decreases monotonically, which is opposite to the increase in risk-aversion. This observation motives us to propose the bi-objective optimization model ESM for handling the risks arsing in not only systemic randomness but also ambiguity of utility representation.

3.3.2 ϵ-Solution Method

In general Model ESM is non-convex and we use the ϵ-method to depict the Pareto optimality as

$$\begin{aligned} \max_{z} \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})] \\ \text{subject to} \quad & \mathbb{M}[\bar{X}(z) \succeq_{u^{\#}} \bar{Y}] \geq \epsilon, \\ & z \in \mathfrak{Z}. \end{aligned} \tag{ϵ-ESM}$$

Using equation (3.6) of the SMU, we get another equivalent form of model (ϵ-ESM),

$$\begin{aligned}
\max_{z} \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})] \\
\text{subject to} \quad & \epsilon \min_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}(z)) - v(\bar{Y})] + (1-\epsilon)\mathbb{E}[u^{\#}(\bar{X}(z)) - u^{\#}(\bar{Y})] \geq 0, \\
& z \in \mathfrak{Z}.
\end{aligned} \tag{3.26}$$

Assume that model (ϵ-ESM) is feasible for all $\epsilon \in [0,1]$. Let $z^*(\epsilon)$ is an optimal solution of model (ϵ-ESM). Then the Pareto frontier of model ESM can be depicted by $(\mathbb{E}[u^{\#}(\bar{X}(z^*(\epsilon)))], \epsilon)$ varying ϵ from 0 to 1. Model (ϵ-ESM) with $\epsilon = 0$ seeks the maximal preference $\bar{X}(z)$ over Y with respect to $u^{\#}$, i.e., $\mathbb{E}[u^{\#}(\bar{X}(z^*(0))) - u^{\#}(\bar{Y})]$. For the case that $0 < \epsilon < 1$, we have

$$\mathbb{M}[\bar{X}(z^*(\epsilon)) \succeq_{u^{\#}} \bar{Y}] = \epsilon \tag{3.27}$$

It follows by Proposition 4 that

$$(1-\epsilon)\mathbb{E}[u^{\#}(\bar{X}(z^*(\epsilon))) - u^{\#}(\bar{Y})] = \epsilon\mathbb{E}[v_\epsilon(\bar{Y}) - v_\epsilon(\bar{X}(z^*(\epsilon)))], \tag{3.28}$$

where v_ϵ is the worst-case utility function over $\mathfrak{V}$ as

$$v_\epsilon := \operatorname{argmin}_{v \in \mathfrak{V}} \mathbb{E}[v(\bar{X}(z^*(\epsilon))) - v(\bar{Y})]. \tag{3.29}$$

The quantity $\mathbb{E}[u^{\#}(\bar{X}(z^*(\epsilon))) - u^{\#}(\bar{Y})]$ indicates the level of the preference of $\bar{X}(z^*(\epsilon))$ over $\bar{Y}$ with respect to $u^{\#}$, while $\mathbb{E}[v_\epsilon(\bar{Y}) - v_\epsilon(\bar{X}(z^*(\epsilon)))]$ shows how $\bar{Y}$ is preferred to $\bar{X}(z^*(\epsilon))$ in the worst case. In equation (3.28) the Pareto-efficient point requires to balance these two preference relations using the weight ϵ. When $\epsilon = 1$, Model (ϵ-ESM) finds the maximum of the expected reference utility $\mathbb{E}[u^{\#}(\bar{X}(z^*(1)))]$ under the condition that $\bar{X}(z^*(1))$ dominates $\bar{Y}$ over the set $\mathfrak{V}$.

We next discuss model ESM under the assumption of the additive independence of the attributes. Accordingly, all the utility functions in $\mathfrak{V}$, including the reference utility function, should have additive forms (Abbas, 2018; Dyer, 2005). The reference utility function is formulated as

$$u^{\#}(\theta) = \sum_{n=1}^{N} \omega_n u_n^{\#}(\theta_n), \tag{3.30}$$

where ω_n's are nonnegative trade-off weights satisfying $\sum_{n=1}^{N} \omega_n = 1$ and $u_n^{\#}$'s are marginal reference utility functions. Also denote by $\mathfrak{V}_n$ the set of the marginal utility functions regarding the nth attribute and by

$$\Lambda := \left\{ (\lambda_1, \ldots, \lambda_N) \;\middle|\; \sum_{n=1}^{N} \lambda_n \bar{a}_n \leq \bar{b},\ (\lambda_1, \ldots, \lambda_N) \geq 0 \right\} \tag{3.31}$$

a polyhedral set of trade-off weights among the attributes, where $\bar{b}$ and $\bar{a}_n$'s are given coefficient vectors. We now specify the set of utility functions as

$$\mathfrak{V} := \left\{ v \;\middle|\; v(\theta) = \sum_{n=1}^{N} \lambda_n v_n(\theta_n),\ (\lambda_1, \ldots, \lambda_n) \in \Lambda, v_n \in \mathfrak{V}_n \right\}. \tag{3.32}$$

PROPOSITION 5. Model (3.26) with $u^{\#}$ and $\mathfrak{V}$ given in (3.30) and (3.32), respectively, is equivalent to

$$\begin{aligned} \max_{z,\bar{\pi}} \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})] && (\Sigma\text{-ESM}) \\ \text{subject to} \quad & \epsilon \bar{b} \bar{\pi} + (1-\epsilon)\mathbb{E}[u^{\#}(\bar{X}(z)) - u^{\#}(\bar{Y})] \geq 0, && \\ & \bar{a}_n \bar{\pi} \leq \mathbb{E}[v_n(X_n(z)) - v_n(Y_n)], && v_n \in \mathfrak{V}_n,\ n = 1, \ldots, N, \\ & \bar{\pi} \leq 0, && \\ & z \in \mathfrak{Z}. && \end{aligned}$$

Proof: By the definitions given in (3.30) and (3.32), we can rewrite the inner optimization problem in model (3.26) as

$$\begin{aligned}\min_{v\in\mathfrak{V}}\mathbb{E}[v(\bar{X}(z))-v(\bar{Y})] = \min_{(\lambda_1,\ldots,\lambda_N)\in\Lambda} \quad & \sum_{n=1}^{N}\lambda_n\left\{\min_{v_n\in\mathfrak{V}_n}\mathbb{E}[v_n(X_n(z))-v_n(Y_n)]\right\}\\ \text{subject to} \quad & \sum_{n=1}^{N}\lambda_n\bar{a}_n \le \bar{b} \\ & (\lambda_1,\ldots,\lambda_N)\ge 0.\end{aligned} \tag{3.33}$$

The model above is linear for $(\lambda_1,\ldots,\lambda_N)$. Let $\bar{\pi}$ be the vector of its dual variables. The dual model is

$$\begin{aligned}\max_{\bar{\pi}} \quad & \bar{b}\bar{\pi} \\ \text{subject to} \quad & \bar{a}_n\bar{\pi} \le \mathbb{E}[v_n(X_n(z))-v_n(Y_n)], \qquad v_n\in\mathfrak{V}_n,\ n=1,\ldots,N \\ & \bar{\pi}\le 0.\end{aligned} \tag{3.34}$$

The proof is then complete by substituting the dual model for the inner optimization problem in model (3.26).

3.4 Reformulations of Σ-ESM

With additively independent attributes, Proposition 5 represents model (Σ-ESM) as an equivalent reformulation of (ϵ-ESM). Model (Σ-ESM) is a semi-infinite program specified on the sets $\mathfrak{V}_n$ of marginal utility functions. In this section we further develop solvable reformulations of model (Σ-ESM).

3.4.1 Assumptions

We first describe the assumptions needed in this section. Let $\Theta_n := [\theta_{n,0},\theta_{n,L_n}]$ for $n=1,\ldots,N$ and $\Theta=\Theta_1\otimes\cdots\otimes\Theta_n$ (the symbol $\otimes$ represents the Cartesian product) be an bounded subset of $\mathbb{R}^N$. Since the normative decision maker prefers more to less, we let the set $\mathfrak{V}_n$ consist of all increasing utility functions normalized on Θ_n, i.e., $v_n(\theta_{n,0})=0$ and $v_n(\theta_{n,L_n})=1$ for all $v_n\in\mathfrak{V}_n$. If the decision maker is risk-averse, we need to further restrict $\mathfrak{V}_n$ to the set of all increasing, concave, and continuous utility functions on Θ_n. Assume that each marginal utility function $u_n^{\#}$ is a piece-wise linear function with a finite number of break points at $\theta_{n,0} = \eta_{n,0}, \eta_{n,1} \ldots, \eta_{n,L_n-1}, \eta_{n,L_n} = \theta_{n,L_n}$, as follows:

$$u_n^{\#}(\theta_n) := 1 - \sum_{l=1}^{L_n} \alpha_{n,l}(\eta_{n,l} - \theta_n)_+, \tag{3.35}$$

where

$$\alpha_{n,l} := \begin{cases} \frac{u_n^{\#}(\eta_{n,l}) - u_n^{\#}(\eta_{n,l-1})}{\eta_{n,l} - \eta_{n,l-1}}, & \text{if } l = L_n, \\ \frac{u_n^{\#}(\eta_{n,l}) - u_n^{\#}(\eta_{n,l-1})}{\eta_{n,l} - \eta_{n,l-1}} - \frac{u_n^{\#}(\eta_{n,l+1}) - u_n^{\#}(\eta_{n,l})}{\eta_{n,l+1} - \eta_{n,l}}, & \text{if } l = 1, \ldots, L_n - 1. \end{cases} \tag{3.36}$$

This assumption does not lose generality, since the most used standard-gamble and paired-gamble utility assessment methods, such as preference comparison, probability equivalence, value equivalence, certainty equivalence, etc., yield piece-wise linear utility functions (Farquhar, 1984).

We also assume that $\bar{Y}$ and $\bar{X}(z)$ for any given z are discrete random vectors. Let

$$p^k := \mathbb{P}[\bar{X}(z) = \bar{x}^k(z)], \qquad k = 1, \ldots K, \tag{3.37}$$

$$q^t := \mathbb{P}[\bar{Y} = \bar{y}^t], \qquad t = 1, \ldots T, \tag{3.38}$$

where $\bar{x}^k(z) = (x_1^k(z), \ldots, x_N^k(z))$ and $\bar{y}^t = (y_1^t, \ldots, y_N^t)$ are the realizations of $\bar{X}(z)$ and $\bar{Y}$. We separately sort the realizations of the nth component of $\bar{Y}$ and the boundary points of Θ_n, i.e., $y_n^1, \ldots, y_n^{\mathcal{T}}, \theta_{n,0}$, and θ_{n,L_n}, in strictly ascending order as

$$\theta_{n,0} = y_n^{(0)} < \cdots < y_n^{(T_n+1)} = \theta_{n,L_n}, \qquad n = 1, \ldots, N. \tag{3.39}$$

Notice that T_n may be less than $\mathcal{T}$ since we combine equal numbers in the sorting. Accordingly, the marginal probabilities of the sorted sequence are

$$q_n^{(t)} := \mathbb{P}[Y_n = y_n^{(t)}], \qquad t = 0, \ldots, T_n + 1. \tag{3.40}$$

3.4.2 MILP Reformulation for Preference Maximizing Decision Maker

We are now dedicate to the case with each set $\mathfrak{V}_n$ consisting of all increasing utility functions. This following theorem reformulates model (Σ-ESM) to be a mixed-integer linear program.

THEOREM 6. Let $\mathfrak{V}_n$ consist of all increasing utility functions normalized on Θ_n for $n = 1, \ldots, N$. Model (Σ-ESM) is equivalent to

$$\max_{z,\bar{\pi},\psi,s} \quad \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \left(\mathbb{E}[(\eta_{n,l} - Y_n)_+] - \sum_{k=1}^{K} p^k s_{n,k,l} \right) \tag{3.41a}$$

$$\text{subject to} \quad -\epsilon \bar{b} \bar{\pi} + (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \sum_{k=1}^{K} p^k s_{n,k,l}$$

$$\leq (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \mathbb{E}[(\eta_{n,l} - Y_n)_+], \tag{3.41b}$$

$$x_n^k(z) + s_{n,k,l} \geq \eta_{n,l}, \quad n = 1, \ldots, N, \; l = 1, \ldots, L_n, \; k = 1, \ldots, K, \tag{3.41c}$$

$$\sum_{k=1}^{K} p^k \sum_{i=0}^{t-1} \psi_{n,k,i} + \bar{a}_n \bar{\pi} \leq \sum_{i=0}^{t-1} q_n^{(i)}, \quad n = 1, \ldots, N, \; t = 1, \ldots, T_n + 1, \tag{3.41d}$$

$$x_n^k(z) - \sum_{t=0}^{T_n+1} y_n^{(t)} \psi_{n,k,t} \geq 0, \quad n = 1, \ldots, N, \; k = 1, \ldots, K, \tag{3.41e}$$

$$\sum_{t=0}^{T_n+1} \psi_{n,k,t} = 1, \quad n = 1, \ldots, N, \; k = 1, \ldots, K \tag{3.41f}$$

$$\psi_{n,k,t} \in \{0, 1\}, \quad n = 1, \ldots, N, \; k = 1, \ldots, K$$

$$t = 0, \dots, T_n + 1 \tag{3.41g}$$

$$s_{n,k,l} \geq 0, \qquad n = 1, \dots, N, \quad l = 1, \dots, L_n, \quad k = 1, \dots, K, \tag{3.41h}$$

$$\bar{\pi} \leq 0, \tag{3.41i}$$

$$z \in \mathfrak{Z}. \tag{3.41j}$$

To prove Theorem 6, we need to state the technical lemma which represents the equivalent formulations of the problem minimizing the function $\Delta(X_n, Y_n; v_n)$ over the set of all increasing utility functions.

LEMMA 1. For $n = 1, \dots, N$, let $\mathbb{F}_{X_n}$ and $\mathbb{F}_{Y_n}$ be the CDFs of X_n and Y_n. Then

$$\min_{v_n \in \mathfrak{V}_n} \mathbb{E}[v_n(X_n) - v_n(Y_n)] = \min_{t=1,\dots,T_n+1} \mathbb{F}_{Y_n}(y_n^{(t-1)}) - \mathbb{F}_{X_n}(y_n^{(t)}-), \tag{3.42}$$

where $\mathbb{F}_{X_n}(y-) = \mathbb{P}[X_n < y]$.

Proof: For the sake of conciseness, we omit the subscript n in the proof. We first prove the first equality in the theorem

$$\min_{v \in \mathfrak{V}} \mathbb{E}[v(X) - v(Y)] = \min_{\theta \in [\theta_0, \theta_L]} \mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta). \tag{3.43}$$

Recall that $\mathfrak{V}$ consists of all increasing utility functions normalized on Θ_n. For any $v \in \mathfrak{V}$, we have

$$\begin{aligned} \mathbb{E}[v(X) - v(Y)] &= \int_{\theta_0}^{\theta_L} v(\theta) d\mathbb{F}_X(\theta) - \int_{\theta_0}^{\theta_L} v(\theta) d\mathbb{F}_Y(\theta) \\ &= \int_{\theta_0}^{\theta_L} \mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta) dv(\theta) \\ &\geq \min_{\theta \in [\theta_0, \theta_L]} \{\mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta)\} \int_{\theta_0}^{\theta_L} dv(\theta) \\ &= \min_{\theta \in [\theta_0, \theta_L]} \mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta). \end{aligned} \tag{3.44}$$

Observe, that for any $\theta \in [\theta_0, \theta_L]$ exists $t \in \{1, \ldots, T+1\}$ such that $\theta \in [y^{(t-1)}, y^{(t)})$. Since $\mathbb{F}_Y$ is right continuous,

$$\mathbb{F}_Y(\theta) = \mathbb{F}_Y(y^{(t-1)}). \tag{3.45}$$

In addition, since $\mathbb{F}_X$ is nondecreasing,

$$\mathbb{F}_X(\theta) \leq \mathbb{F}_X(y^{(t)}-). \tag{3.46}$$

We have

$$\min_{\theta \in [\theta_0, \theta_L]} \mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta) \geq \min_{t=1,\ldots,T+1} \mathbb{F}_Y(y^{(t-1)}) - \mathbb{F}_X(y^{(t)}-) \tag{3.47}$$

On the other hand, recall that $y^{(0)} = \theta_0$ and $y^{(T+1)} = \theta_L$ are defined in (3.39). We have $\mathbb{F}_X(y^{(0)}) = \mathbb{F}_Y(y^{(0)}) = 0$ and $\mathbb{F}_X(y^{(T+1)}) = \mathbb{F}_Y(y^{(T+1)}) = 1$. Thus $\mathbb{F}_Y(\theta) - \mathbb{F}_X(\theta) = 0$ when $\theta = y^{(0)}$ or $y^{(T+1)}$. It follows then

$$\begin{aligned}
\min_{t=1,\ldots,T+1} \mathbb{F}_Y(y^{(t-1)}) - \mathbb{F}_X(y^{(t)}-) &= \min_{t=1,\ldots,T+1} \mathbb{F}_Y(y^{(t)}-) - \mathbb{F}_X(y^{(t)}-) \\
&= \min_{t=0,\ldots,T} \mathbb{F}_Y(y^{(t)}) - \mathbb{F}_X(y^{(t)}) \\
&= \min_{t=0,\ldots,T+1} \mathbb{F}_Y(y^{(t)}) - \mathbb{F}_X(y^{(t)}).
\end{aligned} \tag{3.48}$$

For any $t \in \{0, \ldots, T+1\}$

$$\begin{aligned}
\mathbb{F}_Y(y^{(t)}) - \mathbb{F}_X(y^{(t)}) &= \mathbb{P}\left[X > y^{(t)}\right] - \mathbb{P}\left[Y > y^{(t)}\right] \\
&= \mathbb{E}\left[Id_{(y^{(t)}, \theta_L]}(X) - Id_{(y^{(t)}, \theta_L]}(Y)\right],
\end{aligned} \tag{3.49}$$

where the indicator function $Id_{(y^{(t)}, \theta_L]}(x)$ equals to 1 if $x \in (y^{(t)}, \theta_L]$ and 0 otherwise. Since $Id_{(y^{(t)}, \theta_L]} \in \mathfrak{V}$, we have

$$\mathbb{E}\left[Id_{(y^{(t)}, \theta_L]}(X) - Id_{(y^{(t)}, \theta_L]}(Y)\right] \geq \min_{v \in \mathfrak{V}} \mathbb{E}[v(X) - v(Y)]. \tag{3.50}$$

□

The proof of Theorem 6.

We first give the mixed-integer linear formulation of the constraint $\bar{a}_n\bar{\pi} \leq \mathbb{E}[v_n(X_n(z)) - v_n(Y_n)]$ for all $v_n \in \mathfrak{V}_n$ and $n = 1, \ldots, N$, in model Σ-ESM, combining the results in Lemma 1 with the formulation of the first order stochastic dominance constraint given in Luedtke (2008). It follows by Lemma 1 that

$$\bar{a}_n\bar{\pi} \leq \min_{v_n \in \mathcal{V}_n} \mathbb{E}[v_n(X_n(z)) - v_n(Y_n)] = \min_{t=1,\ldots,T_n+1} \mathbb{F}_n(y_n^{(t-1)}) - \mathbb{F}_{X_n(z)}(y_n^{(t)}-). \tag{3.51}$$

Using the PMF of Y_n given in (3.40), we have $\mathbb{F}_n(y_n^{(t-1)}) = \sum_{i=0}^{t-1} q_n^{(i)}$. Then the inequality above is equivalent to

$$\mathbb{F}_{X_n(z)}(y_n^{(t)}-) + \bar{a}_n\bar{\pi} \leq \sum_{i=0}^{t-1} q_n^{(i)}, \qquad t = 1, \ldots, T_n + 1. \tag{3.52}$$

To calculate $\mathbb{F}_{X_n(z)}(y_n^{(t)}-)$, we define integer variables as

$$\psi_{n,k,t} := \begin{cases} 1, & \text{if } y_n^{(t)} \leq x_n^k(z) < y_n^{(t+1)} \text{ for } t = 0, \ldots T_n, \\ & \quad \text{or if } x_n^k(z) = y_n^{(T_n+1)} \text{ for } t = T_n + 1 \\ 0, & \text{otherwise,} \end{cases} \tag{3.53}$$

for $k = 1, \ldots, K$ and $t = 0, \ldots, T_n + 1$. By this definition, $x_n^k(z) < y_n^{(t)}$ if and only if $\sum_{i=0}^{t-1} \psi_{n,k,i} = 1$. From here

$$\mathbb{F}_{X_n(z)}(y_n^{(t)}-) = \sum_{k \in \{j \in \{1,\ldots,K\} \mid x_n^j(z) < y_n^{(t)}\}} p^k = \sum_{k=1}^{K} p^k \sum_{i=0}^{t-1} \psi_{n,k,i}. \tag{3.54}$$

Replacing $F_{X(z)}(y^{(t)}-)$ in inequality (3.52) we get (3.41d). The constraints (3.41e)-(3.41g) are up to the definition of $\psi_{n,k,t}$.

We next substitute the piece-wise linear formulation (3.35) of $u^{\#}$ into the objective function $\mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})]$ and the constraint $\epsilon\bar{b}\bar{\pi} + (1-\epsilon)\mathbb{E}[u^{\#}(\bar{X}(z)) - u^{\#}(\bar{Y})] \geq$

0 in model Σ-ESM as

$$\mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})]$$
$$= \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \left(\mathbb{E}[(\eta_{n,l} - Y_n)_+] - \sum_{k=1}^{K} p^k (\eta_{n,l} - x_n^k(z))_+ \right), \quad (3.55)$$

and

$$-\epsilon \bar{b} \bar{\pi} + (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \sum_{k=1}^{K} p^k (\eta_{n,l} - x_n^k(z))_+$$
$$\leq (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \mathbb{E}[(\eta_{n,l} - Y_n)_+]. \quad (3.56)$$

Furthermore, introducing the intermediate variables $s^I_{n,k,l}$ to linearize $(\eta_{n,l} - x_n^k(z))_+$ in the two equations above, we have the objective (3.41a) and the constraints (3.41b), (3.41c), and (3.41h). □

3.4.3 LP Reformulation for Risk-Averse Decision Maker

We now address model (Σ-ESM) for the decision maker who has risk averse attitude towards random outcomes. In this section, we assume each set $\mathfrak{V}_n$ consists of all increasing, concave and continues marginal utility functions on Θ_n. The following theorem provides a linear programming reformulation of model (Σ-ESM).

THEOREM 7. Let $\mathfrak{V}_n$ consist of all increasing, concave and continues utility functions normalized on Θ_n for $n = 1, \ldots, N$. Model (Σ-ESM) is equivalent to

$$\max_{z,\bar{\pi},s^I,s^{II}} \quad \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \left(\mathbb{E}[(\eta_{n,l} - Y_n)_+] - \sum_{k=1}^{K} p^k s^I_{n,k,l} \right) \quad (3.57a)$$

$$\text{subject to} \quad -\epsilon \bar{b} \bar{\pi} + (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \sum_{k=1}^{K} p^k s^I_{n,k,l}$$
$$\leq (1-\epsilon) \sum_{n=1}^{N} \omega_n \sum_{l=1}^{L_n} \alpha_{n,l} \mathbb{E}[(\eta_{n,l} - Y_n)_+], \quad (3.57b)$$

$$\frac{1}{y_n^{(t)} - \theta_{n,0}} \sum_{k=1}^{K} p^k s^{II}_{n,k,t} + \bar{a}_n \bar{\pi} \leq \frac{1}{y_n^{(t)} - \theta_{n,0}} \mathbb{E}[(y_n^{(t)} - Y_n)_+],$$

$$n = 1, \ldots, N, \ t = 1, \ldots, T_n + 1 \quad (3.57c)$$

$$x_n^k(z) + s_{n,k,l}^I \geq \eta_{n,l},$$

$$n = 1, \ldots, N, \ l = 1, \ldots, L_n,$$

$$k = 1, \ldots, K, \quad (3.57d)$$

$$x_n^k(z) + s_{n,k,t}^{II} \geq y_n^{(t)},$$

$$n = 1, \ldots, N, \ k = 1, \ldots, K,$$

$$t = 1, \ldots, T_n + 1, \quad (3.57e)$$

$$s_{n,k,t}^{II} \geq 0,$$

$$n = 1, \ldots, N, \ k = 1, \ldots, K,$$

$$t = 1, \ldots, T_n + 1, \quad (3.57f)$$

$$s_{n,k,l}^I \geq 0,$$

$$n = 1, \ldots, N, \ l = 1, \ldots, L_n,$$

$$k = 1, \ldots, K, \quad (3.57g)$$

$$\bar{\pi} \leq 0, \quad (3.57h)$$

$$z \in \mathfrak{Z}. \quad (3.57i)$$

To prove Theorem 7, we need to state the following technical lemma which represents the equivalent formulations of the problem minimizing the function $\Delta(X_n, Y_n; v_n)$ over the set of all increasing, concave, and continuous utility functions $\mathfrak{V}_n$.

LEMMA 2. For $n = 1, \ldots, N$,

$$\begin{aligned} &\min_{v_n \in \mathfrak{V}_n} \mathbb{E}[v_n(X_n) - v_n(Y_n)] \\ &\qquad = \min_{t=1,\ldots,T_n+1} \frac{1}{y_n^{(t)} - \theta_{n,0}} \left(\mathbb{E}[(y_n^{(t)} - Y_n)_+] - \mathbb{E}[(y_n^{(t)} - X_n)_+] \right). \end{aligned} \quad (3.58)$$

Proof: For the sake of conciseness, we omit the subscript n in the proof without confusion.

We define the following functions

$$\begin{aligned}\phi\left(x;\frac{1}{\theta-\theta_0}\right) &:= 1-\left(1-(x-\theta_0)\frac{1}{\theta-\theta_0}\right)_+ = 1-\frac{1}{\theta-\theta_0}(\theta-x)_+ \quad \text{and} \\ \Delta(\bar{X},\bar{Y};v) &:= \mathbb{E}[v(\bar{X})-v(\bar{Y})].\end{aligned} \tag{3.59}$$

Observe that, for any given parameter of $\theta \in (\theta_0, \theta_L]$, $\phi\left(.;\frac{1}{\theta-\theta_0}\right)$ is an increasing, concave and continues function on $[\theta_0, \theta_L]$ and $\phi\left(\theta_0;\frac{1}{\theta-\theta_0}\right) = 0$, $\phi\left(\theta_L;\frac{1}{\theta-\theta_0}\right) = 1$. Thus, $\phi\left(.;\frac{1}{\theta-\theta_0}\right) \in \mathfrak{V}$ and

$$\begin{aligned}&\min_{t=1,\dots,T+1}\frac{1}{y^{(t)}-\theta_0}\left(\mathbb{E}[(y^{(t)}-Y)_+]-\mathbb{E}[(y^{(t)}-X)_+]\right) \\ &= \min_{t=1,\dots,T+1}\mathbb{E}\left[1-\frac{1}{y^{(t)}-\theta_0}(y^{(t)}-X)_+\right]-\mathbb{E}\left[1-\frac{1}{y^{(t)}-\theta_0}(y^{(t)}-Y)_+\right] \\ &= \min_{t=1,\dots,T+1}\Delta\left(X,Y;\phi\left(.;\frac{1}{y^{(t)}-\theta_0}\right)\right) \\ &\geq \min_{v\in\mathfrak{V}}\Delta(X,Y;v).\end{aligned} \tag{3.60}$$

On the other hand, for any $v \in \mathfrak{V}$, let

$$\beta_t := \begin{cases} \frac{v(y^{(T+1)})-v(y^{(T)})}{y^{(T+1)}-y^{(T)}}(y^{(T+1)}-\theta_0), & \text{if } t = T+1 \\ \frac{v(y^{(t)})-v(y^{(t-1)})}{y^{(t)}-y^{(t-1)}}(y^{(t)}-\theta_0)-\frac{v(y^{(t+1)})-v(y^{(t)})}{y^{(t+1)}-y^{(t)}}(y^{(t)}-\theta_0) & \text{if } t = 1,\dots,T, \end{cases} \tag{3.61}$$

and denote

$$v_T(x) := \sum_{t=1}^{T+1}\beta_t\phi\left(x;\frac{1}{y^{(t)}-\theta_0}\right). \tag{3.62}$$

Let us compare v and v_T. Since v is an increasing concave function, $v(x) \geq v_T(x)$ for all $x \in \Theta$ and $v(y^{(t)}) = v_T(y^{(t)})$ for $t = 0,\dots,T+1$. Therefore, $\mathbb{E}[v(X)] \geq \mathbb{E}[v_T(X)]$ and

$\mathbb{E}[v(Y)] \geq \mathbb{E}[v_T(Y)]$. Using these results we have

$$\begin{aligned}\Delta(X,Y;v) \geq\ & \Delta(X,Y;v_T) \\ =\ & \sum_{t=1}^{T} \beta_t \Delta\left(X,Y;\phi\left(.;\frac{1}{y^{(t)}-\theta_0}\right)\right) \\ \geq\ & \min_{t=1,\ldots,T+1} \Delta\left(X,Y;\phi\left(.;\frac{1}{y^{(t)}-\theta_0}\right)\right) \sum_{t=1}^{T+1} \beta_t \\ =\ & \min_{t=1,\ldots,T+1} \Delta\left(X,Y;\phi\left(.;\frac{1}{y^{(t)}-\theta_0}\right)\right).\end{aligned} \tag{3.63}$$

□

The proof of Theorem 7.

We first linearize the constraints $\bar{a}_n\bar{\pi} \leq \mathbb{E}[v(X_n(z)) - v_n(Y_n)]$ for $v_n \in \mathfrak{V}_n$ and $n = 1,\ldots,N$, in model Σ-ESM. By Lemma 2 we have

$$\begin{aligned}\bar{a}_n\bar{\pi} &\leq \min_{v_n\in\mathfrak{V}_n} \mathbb{E}[v(X_n(z)) - v_n(Y_n)] \\ &= \min_{t=1,\ldots,T_n+1} \frac{1}{y_n^{(t)}-\theta_{n,0}} \left(\mathbb{E}[(y_n^{(t)}-Y_n)_+] - \mathbb{E}[(y_n^{(t)}-X_n(z))_+]\right).\end{aligned} \tag{3.64}$$

Introducing the intermediate variables $s^{II}_{n,k,t} \geq 0$, we linearize $(y_n^{(t)} - X_n(z))_+$ to obtain (3.57c), (3.57e), and (3.57f).

We next substitute the piece-wise linear formulation (3.35) of $u^{\#}$ into the objective function $\mathbb{E}[u^{\#}(\bar{X}(z))]$ and the constraint $\epsilon\bar{b}\bar{\pi} + (1-\epsilon)\Delta(\bar{X}(z),\bar{Y};u^{\#}) \geq 0$ in model Σ-ESM as

$$\begin{aligned}&\mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y})] \\ &\qquad = \sum_{n=1}^{N}\omega_n \sum_{l=1}^{L_n}\alpha_{n,l}\left(\mathbb{E}[(\eta_{n,l}-Y_n)_+] - \sum_{k=1}^{K} p^k(\eta_{n,l}-x_n^k(z))_+\right),\end{aligned} \tag{3.65}$$

and

$$-\epsilon\bar{b}\bar{\pi} + (1-\epsilon)\sum_{n=1}^{N}\omega_n\sum_{l=1}^{L_n}\alpha_{n,l}\sum_{k=1}^{K}p^k(\eta_{n,l} - x_n^k(z))_+ \\ \leq (1-\epsilon)\sum_{n=1}^{N}\omega_n\sum_{l=1}^{L_n}\alpha_{n,l}\mathbb{E}[(\eta_{n,l} - Y_n)_+]. \qquad (3.66)$$

Furthermore, introducing the intermediate variables $s^I_{n,k,l}$ to linearize $(\eta_{n,l} - x_n^k(z))_+$ in the two equations above, we have the objective (3.57a) and the constraints (3.57b), (3.57d), and (3.57g). □

3.5 Case Study: Homeland Security Budget Allocation

In this section we apply ESM model to the budget allocation optimization problem described in Hu et al. (2011). The study is about the budget allocation to urban areas in the United States under the Urban Areas Security Initiative (UASI) for prevention, response, and recovery from national catastrophes and terrorist attacks. The total budget under this initiative for 2009 was $799,000,000. This budget was allocated to 62 urban areas, with 60% of the money allocated to the 10 highest risk urban areas (see Table 3.1). The goal is to find such an allocation vector z, which is at least as preferred by the decision maker as the average government allocation for 2005-2009 years, presented as a benchmark z^{Gov} in Table 3.1.

Table 3.1: Budget Allocation Benchmark

Urban Area Id (denoted with g)	Urban Area	Avg. of government allocations (z^{Gov})
1	New York	31.93
2	Chicago	10.42
3	Bay Area	6.83
4	Washington, DC-MD-VA-WV	12.94
5	Los Angeles-Long Beach	15.24
6	Philadelphia, PA-NJ	4.19
7	Boston, MA-NH	3.74
8	Houston	5.88
9	Newark	6.60
10	Seattle-Bellevue-Everett	2.24

Logically, the site's budget share shall be calculated proportional to its vulnerability relative to other sites. Since the national catastrophes and terrorist attacks usually have unpredictable nature and multiple targets (i.e., the same site can have very protected infrastructure, but at the same time it can be vulnerable to attacks on pedestrians), the site's vulnerability usually is computed using multiple measures (indicators), which have random nature. Following Hu et al. (2011), the urban sites are enumerated from 1 to 10, we denote the site's id number with g (see Table 3.1). Each g-th site's share is computed proportional to four vulnerability indicators measuring shares of risks of having (i) property losses (C_1^g), (ii) fatalities (C_2^g), (iii) air traffic losses (C_3^g) and (iv) average daily bridge traffic losses (C_4^g). For example, $C_2^2 = 0.2$ will mean that in the total risk of having fatalities during a national catastrophic or terrorist attack Chicago's share is 20%. For the sake of conciseness we refer to these indicators as the g site's "vulnerability risk-share"-s. To incorporate the random nature of terrors attacks, vulnerability risk-shares are calculated using historical data and are treated as random variables with distributions specific to them.

Due to their nature, different sites are vulnerable to different aspects, which means that usually it is impossible to find an allocation with shares proportional to all of the vulnerability risk-shares for all sites. Thus, if Chicago's budget share z_2 is less than the value of vulnerability risk-share C_3^2 of having air traffic losses, then the $C_3^2 - z_2$ difference is the missing budget share necessary to recover from the attack on air traffic. This missing budget share is called the budget's "missallocation" for air traffic losses in Chicago -

$$X_3^2(z) := -(C_3^2 - z_2)_+. \tag{3.67}$$

Following (Hu et al., 2011; Willis et al., 2005), the total missallocation for any vulnerability aspect n is defined as the sum of missallocations of that aspect in all sites:

$$X_n(z) = \sum_{g=1}^{10} X_n^g(z). \tag{3.68}$$

This measure of missallocation is called negative semi-deviation and is widely used in budget management. Since for each site we have four vulnerability risk-shares, in this study we will have four missallocation attributes calculated according the formula (3.68) for each vulnerability. Combining them gives us the vector of the overall budget missallocation

$$\bar{X}(z) = (X_1(z), X_2(z), X_3(z), X_4(z)) \,. \tag{3.69}$$

We measure the preference of an budget allocation z for a decision maker with reference utility function $u^{\#}$, as the expected utility of this missallocation vector, denoted as $\mathbb{E}[u^{\#}(\bar{X}(z))]$. Since the goal is to find an allocation policy at least as preferred as the average of previous allocations, we get our first objective function in Model 1 by maximizing the budget *allocation's preference* with regards to z^{Gov} benchmark from the point of view of given decision maker. For the sake of coherence with the theory, we denote the missallocation vector produced by z^{Gov} as $\bar{Y}^{Gov}$ (i.e. $\bar{Y}^{Gov} := \bar{X}(z^{Gov})$). Thus, in this case study for the decision maker with the reference utility function $u^{\#}$ the first objective function has the following form

$$\max \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y}^{Gov})]. \tag{3.70}$$

On the other hand, the decision maker can find the optimal allocation of this initial problem not-preferable to z^{Gov} due to problems arising from risk sources (Chl 2) and (Chl 2). Then, in the best scenario he will need to reconstruct his utility function again and solve the optimization problem again. In the worst case, he may need to do reallocation of already implemented budget between the urban areas. To avoid possible reallocations of the budget, we pair decision maker's allocation preference maximization with the maximization of his preference stability, represented with the second objective function in Model 1.

$$\begin{aligned} \max \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y}^{Gov})] \\ \max \quad & \mathbb{M}[\bar{X}(z) \succeq_{u^{\#}} \bar{Y}^{Gov}], \\ \text{s.t.} \quad & \|z\|_1 = 1, \\ & z \geq 0, \end{aligned} \tag{Model 1}$$

where $\bar{Y}^{Gov} = \bar{X}(z^{Gov})$ represents the missallocation vector of the benchmark z^{Gov}. The Model 1 is an application of the theoretical ESM model presented in the section 3.3. In this study we will solve it using linear ϵ-reformulation model (3.57) described in section 3.4.3 for a risk averse decision makers. We will investigate Model 1's efficient frontier, corresponding Pareto-optimal solutions, and the effects of various parameters in it.

3.5.1 Problem Setup

The solution method of Model 1 for a risk-averse decision maker was presented in the model (3.57) in subsection 3.4.3. Following the necessary constructions described in the subsection 3.4.1, we construct $\bar{x}^k(z) = (x_1^k(z), x_2^k(z), x_3^k(z), x_4^k(z)),\ k = 1, \ldots, 400$ discrete realizations of the overall budget missallocation vector ($\bar{X}(z)$). In the Appendix A we show that the sample size of 400 is enough for the results of Model 1 to not be affected by any more samples. Since $\bar{X}(z)$'s each element is calculated from all urban areas' vulnerability indicators (see eq. (3.68)), we select $\{(c_1^{g,k}, c_2^{g,k}, c_3^{g,k}, c_4^{g,k}),\ g = 1, \ldots, 10\},\ k = 1, \ldots, 400$ i.i.d samples from $\{(C_1^g, C_2^g, C_3^g, C_4^g),\ g = 1, \ldots, 10\}$ group of all 10 urban areas vulnerability risk-share vectors.

The probability distributions of the risk-share vector elements are considered to be based on information about foreign terrorist organizations, domestic threat groups and different beliefs about the terrorist motivations and capabilities (see Hu et al., 2011, for in detail construction of the probability distributions).

From the selected samples then the elements of discrete realization $\bar{x}^k(z)$ are constructed from the equation (3.68)

$$x_n^k(z) = -\sum_{g=1}^{10}(c_n^{g,k} - z_n)_+,\ n = 1, \ldots, 4. \tag{3.71}$$

We treat this sample outcomes as equally likely, thus in our case the probability of those realization will be

$$p^k = \mathbb{P}[\bar{X}(z) = \bar{x}^k(z)] = \frac{1}{K} = \frac{1}{400}. \tag{3.72}$$

Since in our problem the Y^{Gov} benchmark is the budget missallocations from the average of previous allocation policies, it will also have $T = 400$ discrete realizations

$$q^t = \mathbb{P}[\bar{Y}^{Gov} = \bar{y}^t] = \mathbb{P}[\bar{X}(z^{Gov}) = \bar{x}^t(z^{Gov})] = \frac{1}{T} = \frac{1}{400}. \tag{3.73}$$

Aggregation Weights Set Generation.

In order to use model (3.57) we first must specify matrix A and vector $\bar{b}$ for the wight region Λ. Following the case study in Hu et al. (2011), we chose the set Λ to be the convex hull of 4 vertexes

$$\begin{aligned}
&(0.25 + \gamma^r, 0.25 - \gamma^r/3, 0.25 - \gamma^r/3, 0.25 - \gamma^r/3),\\
&(0.25 - \gamma^r/3, 0.25 + \gamma^r, 0.25 - \gamma^r/3, 0.25 - \gamma^r/3),\\
&(0.25 - \gamma^r/3, 0.25 - \gamma^r/3, 0.25 + \gamma^r, 0.25 - \gamma^r/3),\\
&(0.25 - \gamma^r/3, 0.25 - \gamma^r/3, 0.25 - \gamma^r/3, 0.25 + \gamma^r).
\end{aligned} \tag{3.74}$$

This convex hull is a polytope with circumradius of $2\gamma^r/\sqrt{3}$ and circumcenter of $(0.25, 0.25, 0.25, 0.25)$. In Hu et al. (2011) this set is called Equality-Centered as it does not produce bias towards any specific aspect's perturbation. By changing the parameter γ^r we can control the size of the set Λ, thus we will refer to it as a radius. In order to avoid negative weights γ^r is defined to be from $[0, 0.75]$. The parameters A and $\bar{b}$ are generated from here using facet enumeration algorithm for bounded convex polytopes. Following Hu et al. (2011), we select weight region radius of $\gamma^r = 0.15$ as default. But in section 3.5.2 we discuss its effect on the solutions of Model 1.

Parameter Sets.

To finalize the problem setup we present below the parameter cases which we study in this chapter (see Table 3.2).

Case (i) is considered the default one. In this configuration the benchmark is selected to be the missallocation Y^{Gov} produced by the average of government allocations for 2005-2009 years z^{Gov} (see Table 3.1). We pair this benchmark with an unbiased decision maker. For the sake of exposition, let the decision maker be named Eric, with a reference utility

Table 3.2: Configurations in Four Studied Cases

	Reference utility	Benchmark
Case (i)	$u_e^{\#}$	Y^{Gov}
Case (ii)	$u_i^{\#}$	Y^{Gov}
Case (iii)	$u_p^{\#}$	Y^{Gov}
Case (iv)	$u_e^{\#}$	Y^{Rand}

function

$$u_e^{\#}(\bar{X}(z)) = \sum_{n=1}^{4} 0.25 \frac{e^{-X_n(z)} - e}{1 - e}. \tag{3.75}$$

The utility function $u_e^{\#}$ is a weighted sum of the same exponential utility function, which describe a risk-averse attitudes towards the corresponding budget missallocation attributes. Those exponential functions are well known to have constant absolute risk aversion coefficient of 1 (CARA utility functions). Additionally, the aggregation weights in his utility function are all equal to 0.25, thus no bias is introduced towards any missallocation attribute. We use this configuration of Model 1 in section 3.5.2, where we discuss the main features of Pareto-optimal frontier and solutions. We also used this configuration in appendix A for the computational analysis of the Model 1 and the selection of the appropriate sample size of input data.

Case (ii) and (iii) are dedicated to investigation of results of Model 1 for decision makers with different preferences. In the section 3.5.3 the Model 1 is solved for two additional decision makers (hereby named as Irvin and Paula). Unlike Eric, Irvin is selected to value missallocations in infrastructure attributes ($X_3(z)$ and $X_4(z)$) more then in other areas. While Paula is selected to have opposite views, valuing more the private property ($X_1(z)$) and fatality ($X_2(z)$) attributes (See Table 3.3). Additionally, we discuss the situation of conflicting suggestions, where Irvin and Paula are acting as advisors of Eric. We investigate the effect of those suggestions on the results of Model 1 solved for Eric.

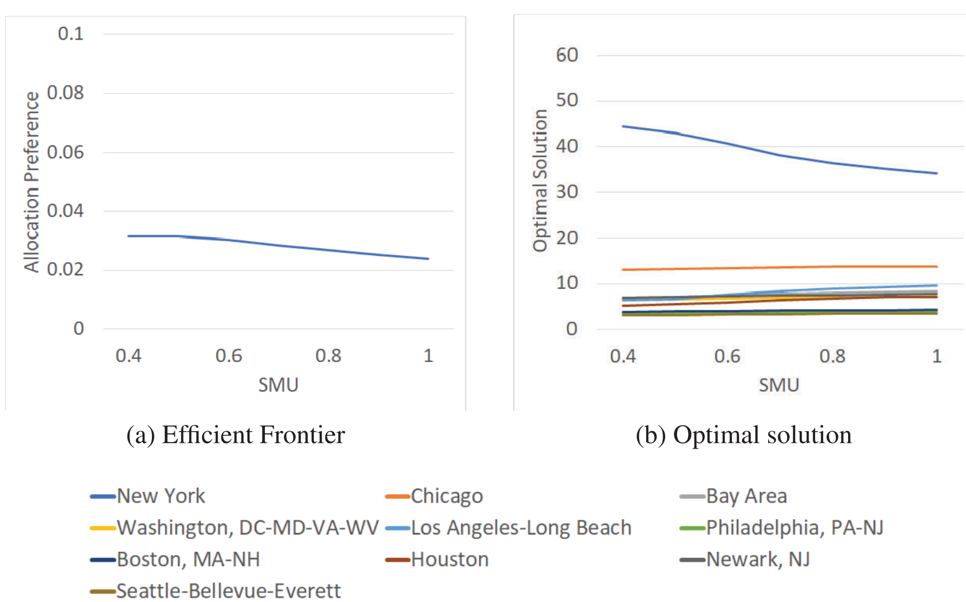
Case (iv) is dedicated to investigation of the effects of different benchmarks on Model 1. In their study Willis et al. (2005) suggest an allocation which is known to have a significant disagreement with z^{Gov}. In section 3.5.4 we use both of this allocations as benchmarks in Model 1 for an unbiased decision maker (Eric). We investigate the point of divergence, the most stable solutions with regards to preference ambiguity and other properties of the results.

Table 3.3: Eric's, Irvin's and Paula's Reference Utility Functions

		Eric $u_e^{\#}(\bar{x})$	Irvin $u_i^{\#}(\bar{x})$	Paula $u_p^{\#}(\bar{x})$
Property Losses	Weight (ω_1)	0.25	0.1	0.4
	Util.Fun. ($u_1^{\#}$)	$(e^{-x_1} - e)/(1 - e)$	$(e^{-x_1} - e)/(1 - e)$	$(e^{-4x_1} - e^4)/(1 - e^4)$
Fatalities	Weight (ω_2)	0.25	0.1	0.4
	Util.Fun. ($u_2^{\#}$)	$(e^{-x_2} - e)/(1 - e)$	$(e^{-x_2} - e)/(1 - e)$	$(e^{-4x_2} - e^4)/(1 - e^4)$
Air Departures	Weight (ω_3)	0.25	0.4	0.1
	Util.Fun. ($u_3^{\#}$)	$(e^{-x_3} - e)/(1 - e)$	$(e^{-4x_3} - e^4)/(1 - e^4)$	$(e^{-x_3} - e)/(1 - e)$
Avg. daily bridge traffic	Weight (ω_4)	0.25	0.4	0.1
	Util.Fun. ($u_4^{\#}$)	$(e^{-x_4} - e)/(1 - e)$	$(e^{-4x_4} - e^4)/(1 - e^4)$	$(e^{-x_4} - e)/(1 - e)$

3.5.2 Results of Base Case

In this subsection we investigate the results of Model 1 with default parameter configuration described in Table 3.2 as Case (i). Figure 3.1a present the efficient frontier of the Model 1 and figure 3.1b contains the corresponding Pareto-optimal solutions.



(a) Efficient Frontier (b) Optimal solution

Figure 3.1: Results of Model 1 With Default Parameter Set.

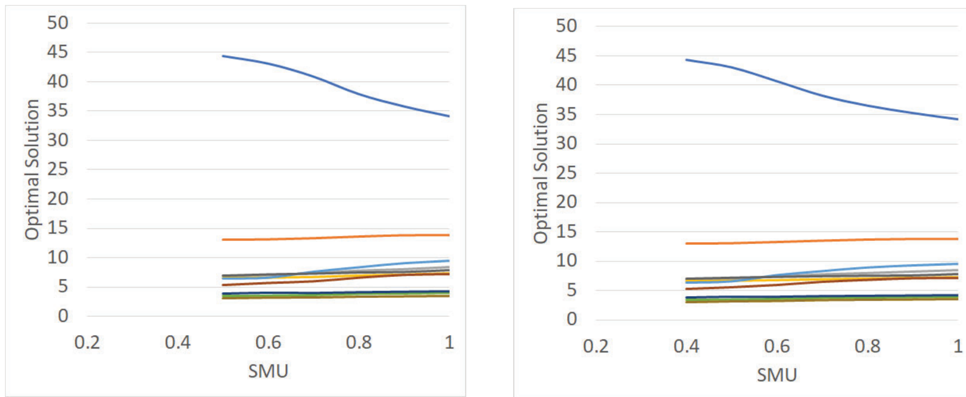
The efficient frontier has characteristic shape of a decreasing curve, due to the conflicting nature of objective function in Model 1. To reach increasingly stable solutions with regards to preference ambiguity, SMU measure requires Model 1 to produce solutions which are preferable by increasing amount of new utility functions. Thus, obviously, shrinking the set of feasible solutions. This shrunken set may not contain solutions which Eric prefers the most anymore, decreasing his preference of new optimal solution.

The efficient frontier starts at SMU value of 0.4, which corresponds to the sensitivity level of the initial solution picked with consideration of only Eric's reference utility function $u_e^{\#}$. The solutions for SMU values of 0 to 0.4 are equal to the initial optimal values (thus are not depicted in Pareto-optimal solution set in figure 3.1b). In all of them the biggest share is the New York urban area's with 44.35%, followed by Chicago with 13.08%. Eric's preference of this initial solution is 0.03.

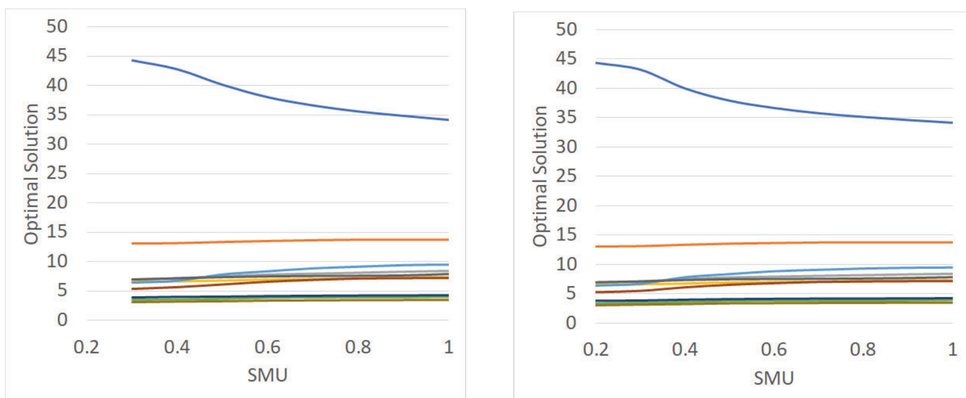
On figure 3.1a, as we move right on efficient frontier passed the $(0.4, 0.03)$ point, the corresponding solutions solutions start to change. Due to consideration of more utility functions in the perturbation set in the SMU the share of New York urban area is steadily decreasing from 44.35% to 34.17%. Which is reached at SMU value of 1. At this point the solution is the most stable with regards to preference ambiguity and will be preferred to benchmark Y^{Gov} by any risk-averse decision maker from the perturbation set. Unfortunately, it also means that in this solution the missallocation attributes are required to be preferable to the corresponding attributes in Y^{Gov} benchmark by risk-averse utility function describing hypothetical decision makers with rather inhumane risk attitudes. The consideration of such risk-attitudes was the main challenge which in chapter 2 we tried to solve.

Other than the reference utility function, benchmark and attribute-wise core traits encoded in the perturbation region, the only parameter affecting Model 1 results is the aggregation weight region. Since we do not want to introduce any bias between the attributes, the only parameter controlling aggregation weight region is its radius- γ^r. In the base case it is selected to have value of 0.15. To see how does the aggregation weight region size affect the optimal solutions we have applied Model 1 for Eric with different values of γ^r weight region radius. The figure 3.2 presents the Pareto-optimal solution sets for γ^r values of 0, 0.15, 0.4 and 0.75.

In all three figures the solutions start from Eric's most preferred allocation and reach



(a) Optimal Solutions for $\gamma^r = 0$. (b) Optimal Solutions for $\gamma^r = 0.15$.



(c) Optimal Solutions for $\gamma^r = 0.4$. (d) Optimal Solutions for $\gamma^r = 0.75$.

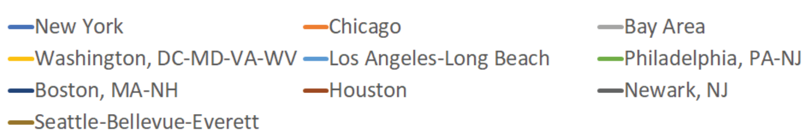


Figure 3.2: Impact of Perturbation Size on Optimal Asset Allocations.

the SMU value of 1, converging to the same values - Eric's most stable solution with regards to preference ambiguity (see figure 3.1b). In general case this effect may not appear as with the increase of weight region, we increase the considered utility functions in the perturbation set. Thus in general sense, the increase of weight region may cause the most stable solutions with regards to preference ambiguity to be different.

On the other hand, the comparison of figures 3.2a-3.2d reveals more general, negative relationship of the aggregation weight region radius γ^r and the SMU level of Eric's most preferred initial solution. The conflict in this relationship is explained with the nature of weight region. As the weight region increases in size (i.e. as γ^r is increased), the number of aggregated utility functions used in SMU measure increases as well. This results in introduction of additional constraints on a feasible solutions. Those constraints shrink the feasibility region of the solutions, which may not contain the initial optimal solution anymore. Thus resulting in lower SMU level for that solution in the efficient frontier of Model 1.

Returning to the observation of Eric's efficient frontier in figure 3.1a, one can observe that it is almost flat. The difference of preference values between the most and least stable allocations with regards to preference ambiguity is only 0.007. This means that the missallocations from average of government solutions z^{Gov} and the solution from Eric's reference utility function $u_e^{\#}$ are very close. Thus $u_e^{\#}$ unbiased reference utility function can be used to describe an government's average allocation preference. This observation was one of the reasons we picked $u_e^{\#}$ as a default reference utility function in this study. In the next subsection we compare it with the other functions described in Table 3.3.

But how to pick a solution? When given the efficient frontier and corresponding solution (see figure 3.1), it is natural to ask: Which solution Eric must pick?

It is obvious that Eric shall pick solutions which are not corresponding to endpoints on the efficient frontier. As the solution corresponding to the first point will usually be highly sensitive with regards to preference ambiguity, while the last point will usually be very conservative.

In budget allocation problems, such as one considered in this case study, the decision maker usually is advised to pick a solution within a given variation tolerance when compared to the most stable solution with regards to preference ambiguity. For example, suppose that Eric can have $\pm 5\%$ maximum variation between the corresponding shares of the same urban area in the most stable allocation with regards to preference ambiguity and the allocation he picked. Then, he shall start from the allocation corresponding to SMU value of 1 and move left on SMU level axis reducing its value. He shall stop when the maximum difference between the new allocation and the allocation corresponding to SMU value of 1 on some urban area becomes larger than 5%. With this technique, for this case, Eric must pick the allocation corresponding to SMU level of 0.7. At this point the New York's share is 38.22%, which is already 4.05% higher than 34.16% in the most stable allocation with regards to preference ambiguity. With $\pm 5\%$ restriction Eric cannot chose allocation at SMU level of 0.6, as at that point the difference for the New York's shares already increases to 6.5%.

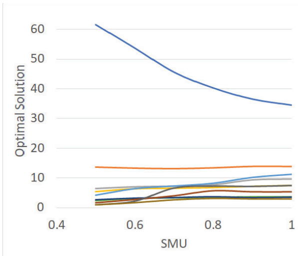
3.5.3 Impact of Reference Utility Function

We now investigate cases (ii) and (iii) to analyze the effect of decision maker's reference utility function on the results of Model 1. For this sake we select two additional decision makers- Irvin and Paula, and solve the Model 1 for them individually.

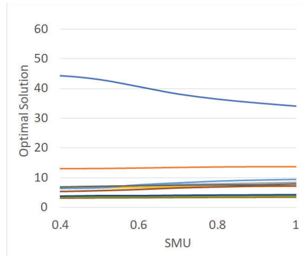
Irvin is selected to value missallocations in infrastructure attributes ($X_3(z)$ and $X_4(z)$) more then in other areas. Accordingly, in his reference utility function $u_i^{\#}$ the missallocation utilities corresponding to air departures and average daily bridge traffic attributes have higher weights of 0.4 and Arrow and Pratt's risk-aversion coefficients of 4. Additionally, the missallocation utilities corresponding to private property losses and fatalities have lower weights of 0.1 and risk-aversion coefficients of 1 (See Table 3.3). On the other hand, Paula is represented as a complete opposite to Irvin, valuing missallocations in private property ($X_1(z)$) and fatalities ($X_2(z)$) more then in infrastructure attributes. Correspondingly, her reference utility function ($u_p^{\#}$) in Table 3.3 has mirrored to Irvin weights and risk-aversion coefficients.

The figure 3.3 presents the Pareto-efficient frontiers and optimal solutions for Paula, Irvin and Eric generated by Model 1. The efficient frontiers in the figures 3.3d and 3.3f are characteristically decreasing as the personally preferable solutions of Irvin and Paula are being asked to become acceptable for wider set of utility functions.

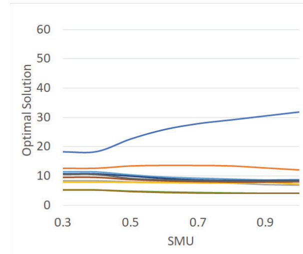
For Paula the efficient frontier starts at SMU value of 0.5. At this point the corresponding optimal solution is the most preferable by her with the allocation preference of 0.17 with regards to Y^{Gov} benchmark. And compared to Eric and Irvin, Paula's personal solution is more stable with regards to preference ambiguity as their most preferable solutions start at SMU value of 0.4 and 0.3 respectively. The solution from figure 3.3a corresponding to this point of $(0.5, 0.17)$ in the efficient frontier allocates 61.7% to New York urban area. This allocation is followed by Chicago comprising 13.7% of the budget. Thus, compared to Eric's initial solution, Paula's contains increased allocations only for New York



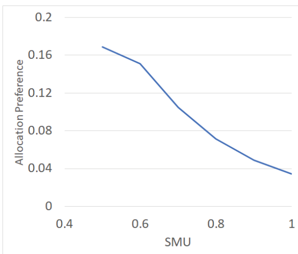
(a) Optimal Solutions for Paula.



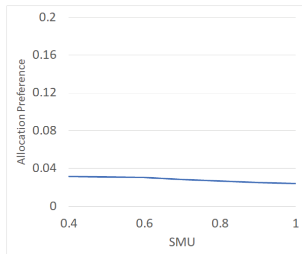
(b) Optimal Solutions for Eric.



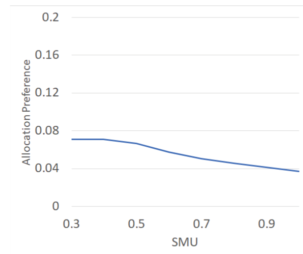
(c) Optimal Solutions for Irvin.



(d) Efficient Frontier for Paula.



(e) Efficient Frontier for Eric.



(f) Efficient Frontier for Irvin.

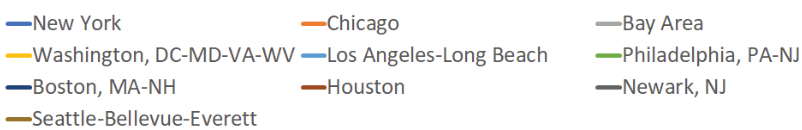


Figure 3.3: Impact of Inconsistency on Optimal Asset Allocations.

(39% increase) and Chicago (4.8% increase) urban areas, in the expense of others. Notably Newark, NJ with -86.6%, Seattle-Bellevue-Everett with -71.4% and Huston -68.3% urban areas share reductions. Those changes are completely expected based on Paula's values depicted in her reference utility function $u_p^\#$ in Table 3.3. As she values the missallocations on private property and population losses more then other attributes, it is only natural for her solution to allocate significant portions of the budget to densely populated urban areas with high population.

As the SMU value in the efficient frontier is increased beyond 0.5, Paula's preference of solution drops fast from its initial value of 0.17 to 0.03. The drop speed indicates that Paula has strong preference towards her initial solution, but it is in conflict with the new risk-averse utility functions in the perturbation set. The decrease in Paula's preference values and the changes in the Pareto-optimal solutions can be traced up to the SMU value of 1. At this point solution is the most stable with regards to preference ambiguity, preferred by all possible risk-averse decision maker's described by the utility functions in the perturbation set. The solution allocates 34.6% to New York and 13.9% to Chicago. It is interesting to observe that LA-Long Beach urban area's budget share is increased from 4.1% to 11.2%, which is the highest increase, followed by Newark, NJ rising from 0.9% to 7.5%. Compared to Eric's the most sable solution with regards to his preference ambiguity, Paula's most sable allocation with regards to her preference ambiguity differs in a very small amount. The Euclidean distance between solutions is only 3% and the maximal difference is 3.1% for Houston urban area.

Let us now observe the figures 3.3f and 3.3c depicting the efficient frontier and Pareto-optimal solutions for Irvin generated by Model 1. The efficient frontier starts at $(0.3, 0.07)$. At this point the solution is the most preferred one by Irvin, with allocation preference of 0.07. This solution allocates the budget in a relatively even manner. The highest allocation share is still New York urban area's (18.3%), followed by Chicago (12.6%) and LA-Long Beach (11.4%). Unlike Paula and Eric, Irvin's initial solution allocates 10.8% budget share to Boston (Paula does 2.7% and Eric 3.8%), MA-NH and 9.6% to Houston (Paula does 1.7% and Eric 5.3%) urban areas. This allocations are characteristic to decision maker like Irvin, and can be expected based on their reference utility function features. In the case of Irvin, in his reference utility function $u_i^{\#}$ in Table 3.3 indicates that he values the critical infrastructure of the urban area more then the other attributes.

As the value of SMU is increased in the efficient frontier Irvin's preference of the solutions drops from 0.07 to 0.03. This drop although is not as steep as in Paula's case, but still is noticeable compared to Eric. Unlike the Eric and Paula, in the corresponding solutions the New York's share is increasing (from 18.3% to 31.9%). Finally, at SMU value of 1, solution becomes the most stable with regards to preference ambiguity, preferred to Y^{Gov} benchmark by all risk-averse utility functions in the perturbation set. The Euclidean distance between Irvin's and Eric's most stable solutions with regards to preference ambiguity is 5.6% and the maximal difference is 18.8% for Boston, MA-NH urban area.

Targeted Sensitivity Analysis.

Let us assume that Irvin and Paula are selected to advise Eric in picking an appropriate solution. From Figure 3.3 we observe that Pareto-optimal solutions of Model 1 for Eric, Irvin and Paula are very different from each other. As Irvin and Paula are advisors of Eric, it is natural for him to investigate how the sensitivity of an allocation z^* he picked from his Pareto-optimal solutions will change for Irvin and Paula. That change can be described by Frechet derivative of the second objective value in the Model 1 with respect to Erick's reference utility function on the direction pointing towards Irvin's or Paula's reference utility functions. We have picked d_i and d_p unit directions pointing towards Irvin's ($u_i^\#$) and Paula's ($u_p^\#$) reference utility functions, correspondingly:

$$d_i(x) = \frac{u_i^\#(x) - u_e^\#(x)}{\|u_i^\# - u_e^\#\|_{\mathcal{L}_2}} \text{ and } d_p(x) = \frac{u_p^\#(x) - u_e^\#(x)}{\|u_p^\# - u_e^\#\|_{\mathcal{L}_2}}. \tag{3.76}$$

Using this directions in the equations (3.11), (3.12) and (3.13), we can use the corresponding Frechet derivatives to calculate linearly approximated values of SMU measure for Paula and Irvin. The table 3.4 presents those approximations and Frechet derivative values for four points, selected from Eric's efficient frontier (See Table 3.4).

In the table 3.4, solutions z_1^*-z_4^* are selected from Eric's Pareto-optimal solutions of Model 1. The z_1^* is the most sensitive solution with regards to preference ambiguity corresponding to point with SMU value of 0.4 on his efficient frontier (see figure 3.1). On the other hand, z_4^* is the most stable solution with regards to preference ambiguity corresponding to point with SMU value of 1. Without loss of generality, we have chosen z_2^* and z_3^* Pareto-optimal solutions to correspond to points on Eric's efficient frontier with equally spaced 0.6 and 0.8 values of SMU.

From table 3.4, we can observe that solution z_1^* has 0.4 SMU value for Eric and approximately 0.7 for Paula. However it is approximated to be very unstable with regards to preference ambiguity for Irvin with SMU value of 0. As we move from z_1^* to z_4^*, we can observe that the solutions are approximated to become more stable with regards to preference ambiguity for Paula and Irvin. This effect is expected as the more stable solutions regards to preference ambiguity are required to stay preferable for a utility function from

Table 3.4: Approximate SMU Levels and Frechet Derivatives on Directions Pointing to Irvin's and Paula's Reference Utility Functions

Urban Area	z_1^*	z_2^*	z_3^*	z_4^*
New York	44.35	40.7	36.52	34.17
Chicago	13.08	13.32	13.7	13.78
Bay Area	6.94	7.43	7.96	8.42
Washington, DC-MD-VA-WV	6.56	6.74	7.03	7.31
Los Angeles-Long Beach	6.42	7.66	8.92	9.5
Philadelphia, PA-NJ	3.43	3.59	3.86	3.97
Boston, MA-NH	3.85	3.98	4.17	4.28
Houston	5.30	5.98	6.85	7.18
Newark, NJ	7	7.34	7.56	7.85
Seattle-Bellevue-Everett	3.08	3.25	3.43	3.53
Erick's Allocation Preference	0.03	0.03	0.03	0.02
Erick's Allocation SMU Level	0.4	0.6	0.8	1
Irvin's Allocation Preference	0.03	0.03	0.02	0.02
Irvin's Allocation's Approx SMU Level	0	0.21	0.66	1
Paula's Allocation Preference	0.03	0.03	0.02	0.02
Paula's Allocation's Approx SMU Level	0.7	0.82	0.88	1
$\left\langle D^{\mathcal{F}}_{u_e^{\#}}, d_p \right\rangle$ (change Eric $\rightarrow$ Paula)	15.58	10.79	5.29	2.22
$\left\langle D^{\mathcal{M}}_{u_e^{\#}}, d_p \right\rangle$ (change Eric $\rightarrow$ Paula)	118.62	85.28	31.59	0
$\left\langle D^{\mathcal{F}}_{u_e^{\#}}, d_i \right\rangle$ (change Eric $\rightarrow$ Irvin)	-15.58	-10.79	-5.29	-2.22
$\left\langle D^{\mathcal{M}}_{u_e^{\#}}, d_i \right\rangle$ (change Eric $\rightarrow$ Irvin)	-118.62	-85.28	-31.59	0
Derivative Ratio ($N'\left(\langle \mathcal{F}, u^{\#} \rangle\right)$)	7.61	7.90	5.97	0

wider perturbation sets.

Interestingly, we can observe that the solution z_4^* is approximated to have SMU values of 1 for both Irvin and Paula. It must be noted, that this does not mean that it is the most optimal stable solution for both of them. However, since Paula and Irvin share the same perturbation area as Eric (i.e. they all are risk-averse), we can claim that z_4^* will have SMU value of 1 for all of them and is one of the candidates of the most optimal stable solution with regards to corresponding preference ambiguities.

On the other hand, we can observe that z_4^* is also the most disliked version by all three decision makers. We argue that for this kind of situations the solution z_3^* maybe allocation

Eric would like to choose, as it is approximated to be preferable by all three decision makers, while also having decent SMU values for all of them.

The table 3.4 also highlights the conflict between Irvin's and Paula's risk preferences. As we see for the same solutions their approximate sensitivity levels are always opposing to each other. For any solution Irvin's SMU levels are lower and Paula's are higher then Eric's. This is due their values depicted in their reference utility functions, as they are selected to be compete opposite to each other. Their conflict in values is probably best presented in comparison of the corresponding Frechet derivatives. As one can observe, all of the corresponding derivatives are opposite numbers. The derivative values towards Paula's reference utility function $\left(\left\langle D^{\mathcal{F}}_{u_e^{\#}}, d_p \right\rangle \text{ and } \left\langle D^{\mathcal{M}}_{u_e^{\#}}, d_p \right\rangle\right)$ are positives, while values towards Irvin's reference utility function $\left(\left\langle D^{\mathcal{F}}_{u_e^{\#}}, d_i \right\rangle \text{ and } \left\langle D^{\mathcal{M}}_{u_e^{\#}}, d_i \right\rangle\right)$ are negatives.

It is also interesting to observe that for any given allocation z^* the ratio between values of Frechet derivatives of allocation and SMU measure does not change for Irvin and Paula. That ratio is, in fact, the value of $N'\left(\left\langle \mathcal{F}, u^{\#} \right\rangle\right)$ derivative from the equation (3.12), indicating the ratio of change between the first and the second objective functions in the Model 1. The stability of this ratio is due to symmetry in Irvin's and Paula's risk-attitudes with regards to Eric.

3.5.4 Impact of the Benchmark

In the allocation problems, it is typical for the decision makers to have a bottom-line policy with which the optimal solutions are compared. In the theoretical ESM model their outcomes are refereed as benchmarks. By default, in this case study the ESM model was configured with

$$z^{Gov} = (31.93\%, 10.42\%, 6.83\%, 12.94\%, 15.24\%, 4.19\%, 3.74\%, 5.88\%, 6.6\%, 2.24\%)$$

allocation and corresponding to it $\bar{Y}^{Gov}$ missallocation benchmark (see Table 3.1 and Model 1). In this section we study the Case (iv) from configuration table 3.2, where we change the z^{Gov} allocation to

$$z^{Rand} := (58.61\%, 16.39\%, 7.91\%, 5.06\%, 4.95\%, 2.42\%, 2.11\%, 1.21\%, 0.66\%, 0.67\%).$$

This allocation was suggested by RAND Corporation in the (Willis et al., 2005) for the same time period. As one can observe, this benchmarks are vastly different. The biggest difference is in the share of New York urban area, which is 58.61% - 31.93% = 26.68% of the budget, amounting to $2.364 billion over 5 years. The average difference is 6.74% of the budget. Denote the outcome produced by z^{Rand} as $\bar{Y}^{Rand} = \bar{X}(z^{Rand})$. With this benchmark the ESM model will have following form

$$\begin{aligned} \max \quad & \mathbb{E}[u^{\#}(\bar{X}(z))] - \mathbb{E}[u^{\#}(\bar{Y}^{Rand})] \\ \max \quad & \mathbb{M}[\bar{X}(z) \succeq_{u^{\#}} \bar{Y}^{Rand}], \\ \text{s.t.} \quad & \|z\|_1 = 1, \\ & z \geq 0, \end{aligned} \tag{Model 2}$$

In the ESM model, the optimal solutions are required to be more or equally preferable to the benchmark. Therefore different benchmarks will generate different feasible solutions sets. Since

$$\begin{aligned} &\bar{X}(z^{Gov}) \succeq_{u^{\#}} \bar{X}(z^{Gov}) = \bar{Y}^{Gov} \text{ and} \\ &\bar{X}(z^{Rand}) \succeq_{u^{\#}} \bar{X}(z^{Rand}) = \bar{Y}^{Rand}, \end{aligned} \tag{3.77}$$

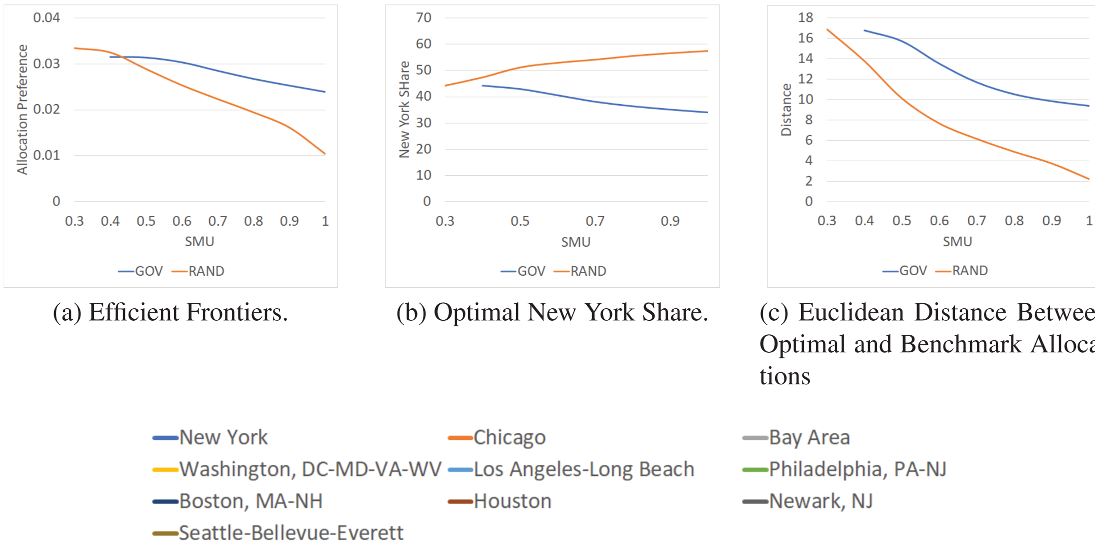
the benchmarks themselves are always members of feasibility sets of Model 1 and Model 2, respectively.

Interestingly, Eric's the most preferred allocation in Model 1 is also his the most preferred allocation in the Model 2. Since it is his initial most sensitive to preference ambiguity allocation we denote to it as

$$z^*_{sensitive} = (44.4\%, 13.08\%, 6.94\%, 6.56\%, 6.42\%, 3.43\%, 3.85\%, 5.3\%, 7.0\%, 3.08\%).$$

This allocation is almost equally distanced from z^{Gov} and z^{Rand} benchmark allocations, with 16.78% and 16.92% Euclidean distances, respectively. Major differences are in the share of New York urban area, with 12.41% increase compared to z^{Gov} and 14.27% decrease compared to z^{Rand}. The $z^*_{sensitive}$ allocation preference with regards to z^{Gov} is equal to 0.0315, while with regards to z^{Rand} is equal to 0.0334. Thus, we can conclude that

Eric prefers z^{Gov} to z^{Rand}. This preference can also be observed in starting solutions on corresponding efficient frontiers in figure 3.4a.



(a) Efficient Frontiers. (b) Optimal New York Share. (c) Euclidean Distance Between Optimal and Benchmark Allocations

Figure 3.4: Results of Model 1 and Model 2 for z^{Gov} and z^{Rand} Benchmark Allocations.

With the increase of SMU value the feasibility set of ESM models shrink. Since the benchmark allocations are in the those sets, this means that optimal solutions become more and more similar to them. This effect can be clearly observed in figure 3.4c, depicting the monotonically decreasing euclidean distances between the optimal solutions and corresponding benchmark allocations in Model 1 and Model 2. Observe, that the distance between optimal solutions and z^{Rand} allocation benchmarks in Model 2 decrease much faster, than the distance between optimal solutions and z^{Gov} in Model 1. This is combined effect of Eric's preference characteristics encoded in his reference utility function and the shrinkage of the feasibility set. Apparently, in the feasibility set of Model 1 preferred by Eric solutions are located relatively far from z^{Gov}, than ones in the feasibility set of Model 2 with regards to z^{Rand}.

It is interesting to observe completely opposite optimal policies in the solutions of Model 1 and Model 2 for the same decision maker. This effect is caused by the difference in feasibility sets of those models due to different benchmark policies. The most striking differences can be found in the share of New York urban area (see figure 3.4b). Even though, the initial solutions are the same, the increase of SMU value causes a monotonic decrease and a monotonic increase in the shares of New York urban area in the Model 1 and Model 2, respectively. At the solution corresponding to SMU value of 1, in the Model 1 the New York's share is decreased by 10.18%, while in the Model 2 it is increased by 13.17%. The overall euclidean distance of those solutions have increased from initial 0 to 25.63%.

3.6 Conclusion

In this chapter we have studied the sensitivity of stochastic optimization models with regards to the preference ambiguity in the decision maker's utility function. We have started with the study and analysis of decision maker's preference ambiguity. Isolating studies where the solutions of stochastic optimization models were registered to be sensitivity with regards to preference ambiguities.

To measure this sensitivity we have chosen a special sensitivity measure of a preference relation with regards to preference ambiguity — SMU. We have studied its boundary values and other properties from the utilitarian point of view. Classically the sensitivity analysis is considered in the post-optimality step. Which forces to re-solve the optimization models iteratively, until the satisfactory results are found. In this study, we have incorporated the SMU sensitivity measure in the optimization model itself. This step ensures, that the produced solutions are maximally stable with regards to decision maker's preference ambiguity.

To preserve the decision maker's individualism. We have incorporated expected utility maximization in the same optimization model. Allowing to ensure the optimality of the solutions from the point of view of classical risk-averse, preference maximization. We have named the constructed bi-objective optimization model incorporating preference maximization and sensitivity analysis on decision maker's preference ambiguity — ESM. And presented two solution methods for it, using linearization techniques. For the preference maximizer decision maker we have presented mixed integer programming model equivalent to ESM. And for a risk-averse decision makers we have presented linear programming model equivalent to ESM.

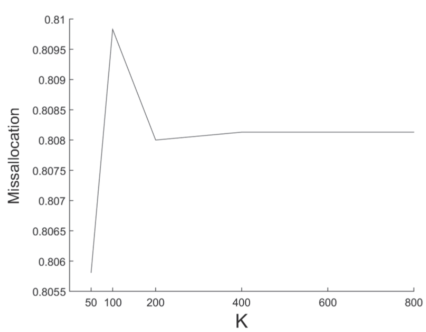
Finally, we have applied the ESM model for the budget allocation to ten major urban areas in the United States under the Urban Areas Security Initiative (UASI). We study the various solutions of the problem for different decision makers. Compare their allocation policies to the government's average allocation and to the allocation suggested by RAND corporation.

We have employed the closed form results for the Frechet derivative of the SMU sensitivity measure to conduct a group decision making analysis for a group of different decision makers. Employed targeted sensitivity analysis allowed to find a solution which was expected to suit all of the members of the group with an acceptable sensitivity values.

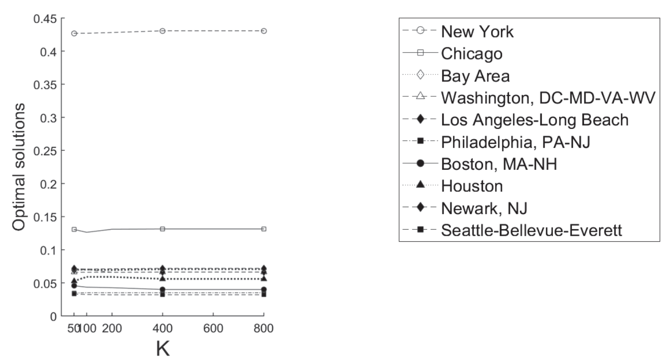
Appendix A

Computational Analysis of Model 1 from Chapter 3

All tests are conducted on a machine with Intel Core i7 3rd Gen processor with 4 physical cores and hyper-threading on each core. The maximum frequency is 2.4 GHz with the boost at specific core up to 3.2GHz. The maximum amount of RAM allowed for computations is 12 GB in total. The Model 1 model is solved in CPLEX Studio 12.6 in OPL language, with CPLEX barrier algorithm in deterministic parallel mode.



(a) Optimal value



(b) Optimal solution

Figure A.1: Computational Results of Model 1 for z^{Gov} Benchmark Allocation.

We solve the Model 1 model with the default parameter set (case (i)) from Table 3.2 while changing the sample size K from 50 to 800. The step size is selected to be 50 samples. The figure A.1 represents the results for the default parameter set for z^{Gov} government benchmark. At the sample sizes 50, 100 and 200 we can observe fluctuations in the optimal missallocation in figure A.1a and in the optimal solutions A.1b. This is clear indication that those sample sizes are not enough and cannot represent the whole population of possible values of $\{(C_1^g, C_2^g, C_3^g, C_4^g),\ g = 1, \ldots, 10\}$ group of all 10 urban areas valnurability risk-share vectors. However, starting from sample size 200 the fluctuations in the solutions decrease and after 400 the difference in the optimal solution are negligible. We observed similar behaviors with all the other parameter sets, including the case where we did consider Rand benchmark. This behavior indicates that the sample size of 400 is big enough the represent the whole population of possible values of $\{(C_1^g, C_2^g, C_3^g, C_4^g),\ g = 1, \ldots, 10\}$ group and new samples do not bring new effects on the results.

Bibliography

Aase, K. K. (2001). On the St. Petersburg paradox. *Scandinavian Actuarial Journal*, *2001*(1), 69–78. https://doi.org/10.1080/034612301750077356

Abbas, A. E. (2018). *Foundations of Multiattribute Utility*. New York, NY: Cambridge University Press.

Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, *125*(1), 40–68. https://doi.org/10.1086/689774

Aguiar, M., & Hurst, E. (2007). Life-cycle prices and production. *American Economic Review*, *97*(5), 1533–1559. https://doi.org/10.1257/aer.97.5.1533

Alexander, E. R. (1989). Sensitivity analysis in complex decision models. *Journal of the American Planning Association*, *55*(3), 323–333. https://doi.org/10.1080/01944368908975419

Armbruster, B., & Delage, E. (2015). Decision making under uncertainty when preference information is incomplete. *Management Science*, *61*(1), 111–128. https://doi.org/10.1287/mnsc.2014.2059

Barron, H., & Schmidt, C. P. (1988). Sensitivity analysis of additive multiattribute value models. *Operations Research*, *36*(1), 122–127. https://doi.org/10.1287/opre.36.1.122

Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, *2*(1), 95–121. https://doi.org/10.1016/0304-405x(75)90025-2

Benidis, K., Feng, Y., & Palomar, D. P. (2018). Optimization methods for financial index tracking: From theory to practice. *Foundations and Trends in Optimization*, *3*(3), 171-279. https://doi.org/10.1561/2400000021

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769–805. https://doi.org/10.1287/moor.23.4.769

Berger, J. O., Insua, D. R., & Ruggeri, F. (2000). Bayesian robustness. In D. R. Insua & F. Ruggeri (Eds.), *Robust bayesian analysis* (pp. 1–32). New York, NY: Springer Science+Business Media.

Berger, J. O., Moreno, E., Pericchi, L. R., Bayarri, M. J., Bernardo, J. M., Cano, J. A., ... Sivaganesan, S. (1994). An overview of robust bayesian analysis. *Test*, *3*(1), 5–124. https://doi.org/10.1007/BF02562676

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*(1), 23. https://doi.org/10.2307/1909829

Bertsimas, D., Doan, X. V., Natarajan, K., & Teo, C.-P. (2010). Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research*, *35*(3), 580–602. https://doi.org/10.1287/moor.1100.0445

Bertsimas, D., Pachamanova, D., & Sim, M. (2004). Robust linear optimization under general norms. *Operations Research Letters*, *32*(6), 510–516. https://doi.org/10.1016/j.orl.2003.12.007

Birge, J. R., & Louveaux, F. (2011). *Introduction to Stochastic Programming*. New York, NY: Springer Science+Business Media.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. New York, NY: Cambridge University Press.

Brockett, P. L., & Golden, L. L. (1987). A class of utility functions containing all the common utility functions. *Management Science*, *33*(8), 955–964. https://doi.org/10.1287/mnsc.33.8.955

Bruni, R., Cesarone, F., Scozzari, A., & Tardella, F. (2014). A linear risk-return model for enhanced indexation in portfolio optimization. *OR Spectrum*, *37*(3), 735–759. https://doi.org/10.1007/s00291-014-0383-6

Chajewska, U., & Koller, D. (2000). Utilities as random variables: Density estimation and structure discovery. In C. Boutilier & M. Goldszmidt (Eds.), *Proceedings of the sixteenth national conference on artificial intelligence* (pp. 63–71). Stanford, California: Morgan Kaufmann Publishers.

Chapman, G. B., & Elstein, A. S. (2013). Cognitive processes and biases in medical decision making. In G. B. Chapman & F. A. Sonnenberg (Eds.), *Decision Making in Health Care: Theory, Psychology, and Applications* (p. 183-210). New York, NY: Cambridge University Press.

Cheney, W. (2001). *Analysis for applied mathematics*. New York, NY: Springer.

Clarke, F. (2013). *Functional analysis, calculus of variations and optimal control*. London: Springer Science+Business Media.

Cornuejols, G., & Tutuncu, R. (2006). *Optimization methods in finance*. New York, NY: Cambridge University Press.

Dean, M., & Martin, D. (2016). Measuring rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics*, *98*(3), 524–534. https://doi.org/10.1162/REST_a_00542

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, *58*(3), 595–612. https://doi.org/10.1287/opre.1090.0741 de Mello, T. H., & Bayraksan, G. (2014). Monte Carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, *19*(1), 56–85. https://doi.org/10.1016/j.sorms.2014.05.001

Dentcheva, D., Henrion, R., & Ruszczyński, A. (2007). Stability and sensitivity of optimization problems with first order stochastic dominance constraints. *SIAM Journal on Optimization*, *18*(1), 322–337. https://doi.org/10.1137/060650118

Dentcheva, D., & Ruszczynski, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, *14*(2), 548–566. https://doi.org/10.1137/s1052623402420528

Dentcheva, D., & Ruszczynski, A. (2004). Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. *Mathematical Programming*, *99*(2), 329–350. https://doi.org/10.1007/s10107-003-0453-z

Dentcheva, D., & Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, *30*(2), 433–451. https://doi.org/10.1016/j.jbankfin.2005.04.024

Ditzian, Z. (1989). Best polynomial approximation and bernstein polynomial approximation on a simplex. *Indagationes Mathematicae (Proceedings)*, *92*(3), 243–256. https://doi.org/10.1016/s1385-7258(89)80002-2

Drapkin, D., & Schultz, R. (2010). An algorithm for stochastic programs with first-order dominance constraints induced by linear recourse. *Discrete Applied Mathematics*, *158*(4), 291–297. https://doi.org/10.1016/j.dam.2009.07.008

Dyer, J. S. (2005). Maut — multiattribute utility theory. In J. Figueira, S. Greco, & M. Ehrgott (Eds.), *International series in operations research & management science* (pp. 265–292). New York, NY: Springer.

Echenique, F., Lee, S., & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, *119*(6), 1201–1223. https://doi.org/10.1086/665011

el Ghaoui, L., & Lebret, H. (1997). Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, *18*(4), 1035–1064. https://doi.org/10.1137/s0895479896298130

Ermoliev, Y. (1983). Stochastic quasigradient methods and their application to system optimization†. *Stochastics*, *9*(1-2), 1–36. https://doi.org/10.1080/17442508308833246

Farquhar, P. H. (1984). State of the art—utility assessment methods. *Management Science*, *30*(11), 1283–1300. https://doi.org/10.1287/mnsc.30.11.1283

Fishburn, P. (1979). *Utility theory for decision making*. Huntington, NY: R.E. Krieger Pub. Co.

French, S., & Insua, D. R. (1989). Partial information and sensitivity analysis in multi-objective decision making. In I. G. Lockett A.G. (Ed.), *Lecture notes in economics and mathematical systems* (pp. 424–433). Berlin, Heidelberg: Springer.

Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, *56*(4), 279-304. https://doi.org/10.1086/256692

Giancristofaro, R. A., & Bonnini, S. (2008). Moment-based multivariate permutation tests for ordinal categorical data. *Journal of Nonparametric Statistics*, *20*(5), 383–393. https://doi.org/10.1080/10485250802195440

Gollmer, R., Neise, F., & Schultz, R. (2008). Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. *SIAM Journal on Optimization*, *19*(2), 552–571. https://doi.org/10.1137/060678051

Greco, S., Mousseau, V., & Slowinski, R. (2007). Robust multiple criteria ranking using a set of additive value functions. *Abbales Du Lamsade*(7). Retrieved from https://hal.archives-ouvertes.fr/hal-00180026 (Working paper)

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *American Economic Review*, *59*(1), 25-34. Retrieved from https://www.jstor.org/stable/1811090

Hanoch, G., & Levy, H. (1975). The efficiency analysis of choices involving risk. In *Stochastic optimization models in finance* (pp. 89–100). New York, NY: Elsevier.

Haskell, W. B., & Jain, R. (2015). A convex analytic approach to risk-aware Markov decision processes. *SIAM Journal on Control and Optimization*, *53*(3), 1569–1598. https://doi.org/10.1137/140969221

Hess, S., Daly, A., & Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, *84*(2), 181–204. https://doi.org/10.1007/s11238-017-9651-7

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, *62*(6), 1291. https://doi.org/10.2307/2951750

Hu, J., de Mello, T. H., & Mehrotra, S. (2010). Sample average approximation of stochastic dominance constrained programs. *Mathematical Programming*, *133*(1-2), 171–201. https://doi.org/10.1007/s10107-010-0428-9

Hu, J., de Mello, T. H., & Mehrotra, S. (2011). Risk-adjusted budget allocation models with application in homeland security. *IIE Transactions*, *43*(12), 819–839. https://doi.org/10.1080/0740817x.2011.578610

Hu, J., de Mello, T. H., & Mehrotra, S. (2014). Stochastically weighted stochastic dominance concepts with an application in capital budgeting. *European Journal of Operational Research*, *232*(3), 572–583. https://doi.org/10.1016/j.ejor.2013.08.007

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. https://doi.org/10.1287/opre.2019.1849

Hu, J., & Mehrotra, S. (2014). 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. https://doi.org/10.1080/0740817x.2014.919045

Hu, J., & Stepanyan, G. (2017). Optimization with reference-based robust preference constraints. *SIAM Journal on Optimization*, *27*(4), 2230–2257. https://doi.org/10.1137/16m1105050

Insua, D. R., & French, S. (1991). A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research*, *54*(2), 176–190. https://doi.org/10.1016/0377-2217(91)90296-8

Insua, D. R., Martin, J., Proll, L., French, S., & Salhi, A. (1997). Sensitivity analysis in statistical decision theory: A decision analytic view. *Journal of Statistical Computation and Simulation*, *57*(1-4), 197–218. https://doi.org/10.1080/00949659708811808

Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance*, *21*(1), 61–72. https://doi.org/10.1016/0030-5073(78)90039-9

Karoui, N. E., & Meziou, A. (2006). Constrained optimization with respect to stochastic dominance: Application to portfolio insurance. *Mathematical Finance*, *16*(1), 103–117. https://doi.org/10.1111/j.1467-9965.2006.00263.x

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives*. New York, NY: Cambridge University Press.

Kim, S., Pasupathy, R., & Henderson, S. G. (2014). A guide to sample average approximation. In M. C. Fu (Ed.), *Handbook of simulation optimization* (pp. 207–243). New York, NY: Springer.

Kushner, H. J., & Clark, D. S. (1978). *Stochastic approximation methods for constrained and unconstrained systems*. New York, NY: Springer.

Kushner, H. J., & Yin, G. G. (2003). *Stochastic approximation and recursive algorithms and applications* (2nd ed.). New York, NY: Springer.

Lan, G. (2011). An optimal method for stochastic composite optimization. *Mathematical Programming*, *133*(1-2), 365–397. https://doi.org/10.1007/s10107-010-0434-y

Lean, H. H., McAleer, M., & Wong, W.-K. (2010). Market efficiency of oil spot and futures: A mean-variance and stochastic dominance approach. *Energy Economics*, *32*(5), 979–986. https://doi.org/10.1016/j.eneco.2010.05.001

Leshno, M., & Levy, H. (2002). Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance. *Management Science*, *48*(8), 1074–1085. https://doi.org/10.1287/mnsc.48.8.1074.169

Levy, H. (2016). *Stochastic dominance*. Switzerland: Springer International Publishing.

Li, Q., Nie, Y., Vallamsundar, S., Lin, J., & de Mello, T. H. (2014). Finding efficient and environmentally friendly paths for risk-averse freight carriers. *Networks and Spatial Economics*, *16*(1), 255–275. https://doi.org/10.1007/s11067-013-9220-8

Liese, F., & Miescke, K. J. (2007). *Statistical decision theory*. New York, NY: Springer.

Lizyayev, A., & Ruszczyński, A. (2012). Tractable almost stochastic dominance. *European Journal of Operational Research*, *218*(2), 448–455. https://doi.org/10.1016/j.ejor.2011.11.019

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics*, *8*(4), 301–323. https://doi.org/10.1007/s10683-005-5372-9

Lorentz, G. (2012). *Bernstein polynomials*. New York, NY: Chelsea Pub Co.

Luedtke, J. (2008). New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, *19*(3), 1433–1450. https://doi.org/10.1137/070707956

Mak, W.-K., Morton, D. P., & Wood, R. (1999). Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, *24*(1-2), 47–56. https://doi.org/10.1016/s0167-6377(98)00054-6

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77. https://doi.org/10.2307/2975974

Martein, L., & Schaible, S. (1987). On solving a linear program with one quadratic constraint. *Rivista di Matematica per le Scienze Economiche e Sociali*, *10*(1-2), 75–90. https://doi.org/10.1007/bf02090479

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (p. 105-142). New York, NY: Academic Press.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management*. Princeton, NJ: Princeton University Press.

Müller, A., & Stoyan, D. (2002). *Comparison methods for stochastic models and risks*. Chichester, West Sussex, UK: John Wiley & Sons.

Nemirovski, A., Juditsky, A., Lan, G., & Shapiro, A. (2009). Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, *19*(4), 1574–1609. https://doi.org/10.1137/070704277

Nie, Y., Wu, X., & de Mello, T. H. (2011). Optimal path problems with second-order stochastic dominance constraints. *Networks and Spatial Economics*, *12*(4), 561–587. https://doi.org/10.1007/s11067-011-9167-6

Noyan, N. (2010). Alternate risk measures for emergency medical service system design. *Annals of Operations Research*, *181*(1), 559–589. https://doi.org/10.1007/s10479-010-0787-x

Noyan, N., & Rudolf, G. (2013). Optimization with multivariate conditional value-at-risk constraints. *Operations Research*, *61*(4), 990–1013. https://doi.org/10.1287/opre.2013.1186

Ray, S. S. (2018). *Numerical analysis with algorithms and programming*. Boca Raton, FL: Chapman and Hall/CRC.

Rivlin, T. J. (2010). *An introduction to the approximation of functions*. New York, NY: Dover Pubn inc.

Roman, D., Darby-Dowman, K., & Mitra, G. (2006). Portfolio construction based on stochastic dominance and target return distributions. *Mathematical Programming*, *108*(2-3), 541–569. https://doi.org/10.1007/s10107-006-0722-8

Roman, D., Mitra, G., & Zverovich, V. (2013). Enhanced indexation based on second-order stochastic dominance. *European Journal of Operational Research*, *228*(1), 273–281. https://doi.org/10.1016/j.ejor.2013.01.035

Sage, A. P. (1981). Sensitivity analysis in systems for planning and decision support. *Journal of the Franklin Institute*, *312*(3-4), 265–291. https://doi.org/10.1016/0016-0032(81)90064-8

Samson, D. (1988). *Managerial decision analysis*. Homewood, Ill: Irwin.

Savage, L. J. (2012). *The foundations of statistics*. Garden City, NY: Dover Publications.

Scarf, H. (1958). A min-max solution of an inventory problem. In *Studies in the mathematical theory of inventory and production* (pp. 201–209). Santa Monica, CA: Rand Corporation.

Shaked, M., & Shanthikumar, J. G. (1994). *Stochastic orders and their applications*. United Kingdom: Academic Press.

Shapiro, A., & Ahmed, S. (2004). On a class of minimax stochastic programs. *SIAM Journal on Optimization*, *14*(4), 1237–1249. https://doi.org/10.1137/s1052623403434012

Shapiro, A., & de Mello, T. H. (2000). On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization*, *11*(1), 70–86. https://doi.org/10.1137/s1052623498349541

Shapiro, A., de Mello, T. H., & Kim, J. (2002). Conditioning of convex piecewise linear stochastic programs. *Mathematical Programming*, *94*(1), 1–19. https://doi.org/10.1007/s10107-002-0313-2

Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2014). *Lectures on stochastic programming: Modeling and theory, second edition*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Simon, H. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10.2307/1884852

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Smith, J. M. B. A. M. (1994). *Bayesian theory* (J. M. Bernardo & A. F. M. Smith, Eds.). Chichestcr, West Sussex, UK: John Wiley & Sons, Inc.

Stalmeier, P. F. M., Chapman, G. B., de Boer, A. G. E. M., & van Lanschot, J. J. B. (2001). A fallacy of the multiplicative qaly model for low-quality weights in students and patients judging hypothetical health states. *International Journal of Technology Assessment in Health Care*, *17*(4), 488–496. https://doi.org/10.1017/s026646230110704x

Stalmeier, P. F. M., Wakker, P. P., & Bezembinder, T. G. G. (1997). Preference reversals: Violations of unidimensional procedure invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 1196–1205. https://doi.org/10.1037/0096-1523.23.4.1196

Sun, H., & Xu, H. (2013). Convergence analysis of stationary points in sample average approximation of stochastic programs with second order stochastic dominance constraints. *Mathematical Programming*, *143*(1-2), 31–59. https://doi.org/10.1007/s10107-013-0711-7

Thurstone, L. L. (1994). A law of comparative judgment. *Psychological Review*, *101*(2), 266–270. https://doi.org/10.1037/0033-295x.101.2.266

Triantaphyllou, E., & Sánchez, A. (1997). A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences*, *28*(1), 151–194. https://doi.org/10.1111/j.1540-5915.1997.tb01306.x

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, *76*(1), 31–48. https://doi.org/10.1037/h0026750

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

van de Panne, C. (1966). Programming with a quadratic constraint. *Management Science*, *12*(11), 798–815. https://doi.org/10.1287/mnsc.12.11.798

Vickson, R. G., & Ziemba, W. T. (2014). *Stochastic optimization models in finance*. Maryland Heights, MO: Elsevier Science & Techn.

von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior (60th anniversary commemorative edition)* (2nd ed.). Princeton, NJ: Princeton University Press.

Wakker, P., & Deneffe, D. (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, *42*(8), 1131–1150. https://doi.org/10.1287/mnsc.42.8.1131

Weber, M., & Borcherding, K. (1993). Behavioral influences on weight judgments in multiattribute decision making. *European Journal of Operational Research*, *67*(1), 1–12. https://doi.org/10.1016/0377-2217(93)90318-h

Willis, H. H., Morral, A. R., Kelly, T. K., & Medby, J. J. (2005). *Estimating terrorism risk*. Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/monographs/2005/RAND_MG388.pdf

Yan, J.-A. (2018). *Introduction to stochastic finance*. Singapore: Springer.