Preference Robust Generalized Shortfall Risk Measure Based on the Cumulative Prospect Theory When the Value Function and Weighting Functions Are Ambiguous *

Sainan Zhang and Huifu Xu †

December 21, 2021

Abstract

The utility-based shortfall risk (SR) measure introduced by Föllmer and Schied [15] has been recently extended by Mao and Cai [29] to cumulative prospect theory (CPT) based SR in order to better capture a decision maker’s utility/risk preference. In this paper, we consider a situation where information on the value function and/or the weighting functions in the CPT based SR is incomplete. Instead of using partially available information to construct an approximate value function and weighting functions, we propose a robust approach to define a generalized shortfall risk which is based on a tuple of the worst case value/weighting functions from ambiguity sets of plausible value/weighting functions identified via available information. The ambiguity set may be constructed with elicited preference information (e.g. pairwise comparison questionnaires) and subjective judgement, and the ambiguity reduces as more and more preference information is obtained. Under some moderate conditions, we demonstrate how the subproblem of the robust shortfall risk measure can be calculated by solving a linear programming problem. To examine the performance of the proposed robust model and computational scheme, we carry out some numerical experiments and report preliminary test results. This work may be viewed as an extension of recent research of preference robust optimization models to the CPT based SR.

Key words. Preference robust GSR, ambiguity set, pairwise comparison, tractable formulation

1 Introduction

Let $\xi : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ be a random variable representing a financial loss and $l : \mathbb{R} \rightarrow \mathbb{R}$ be a convex, non-decreasing and non-constant loss/disutility function. The utility-based shortfall risk measure (SR) of $\xi$ is defined as

$$\text{SR}^{\mathbb{P}}_{l,\lambda}(\xi) := \inf \{ x \in \mathbb{R} : \mathbb{E}_{\mathbb{P}}[l(\xi - x)] \leq \lambda \}, \quad \text{(1.1)}$$

*This project is supported by Hong Kong RGC grant 14500620.
†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Email: snzhang@se.cuhk.edu.hk, hfxu@se.cuhk.edu.hk.
where the expectation is taken w.r.t. the probability distribution of $\xi$, i.e., $P = \mathbb{P} \circ \xi^{-1}$ and $\lambda$ is the maximal tolerable expected utility loss. The shortfall risk, denoted by $\text{SR}_P^\lambda(\xi)$, is the minimum cash injection to the financial position ($-\xi$) such that the expected loss of the new position falls below the specified level $\lambda$. The definition is introduced by Föllmer and Schied [15] and has received wide attentions in risk management and finance not only because it covers some well-known risk measures such as value at risk (VaR), conditional value at risk (CVaR) and entropic risk measure, but also better captures tail losses (by a proper choice of loss function) [16]. Moreover, it satisfies invariance under randomization and can be used more appropriately for dynamic measurement of risks over time, see [17, 42].

The SR model is essentially based on Von Neumann-Morgenstein’s (VNM) expected utility theory [31] if we regard the loss function as a disutility function ($l(y) := -u(-y)$). The expected utility theory assumes that the decision maker’s preference satisfies four axiomatic properties (completeness, transitivity, continuity, and independence). In the literature of behavioural economics, it is well known that the theory contradicts some experimental studies where most decision makers tend to underweight medium and high probabilities and overweight low probabilities of extreme outcomes (Elsberg [13] and Allais paradoxes [2]) and this prompts various remedies/modifications of the VNM’s expected utility model. Rank dependent expected utility (RDEU) theory ([35] and [36, Chaper 5]) and cumulative prospect theory (Tversky and Kahneman [38], [41, page 158] and [28, Definition 1]) are subsequently proposed to address the drawbacks of the VNM’s classical model. The definition of SR in (1.1) inherits the drawbacks of the VNM’s expected utility model and this prompts introduction of generalized shortfall risk measure (GSR) based on CPT:

$$
\text{(GSR-CPT) } \text{SR}_{w^-,w^+}^{v,\lambda}(\xi) := \inf_{x \in \mathbb{R}} \quad \text{s.t. } \mathbb{E}_{w^-}^{w^+}[v(\xi - x)] \leq \lambda,
$$

(1.2)

where $v$ is a value function corresponding to disutility function and $w^-, w^+$ are probability weighting functions and

$$
\mathbb{E}_{w^-}^{w^+}[v(\xi)] = \int_0^{\infty} w^+ (\mathbb{P}(v(\xi) \geq t))dt - \int_{-\infty}^{0} w^- (\mathbb{P}(v(\xi) \leq t))dt.
$$

(1.3)

Note that when $w^-(t) = 1 - w^+(1 - t), \forall t \in [0, 1]$ and $v$ is a utility function, the cumulative prospect theory reduces to rank dependent expected utility theory (RDEU) [37, page 19]. Consequently, we call the resulting shortfall risk calculated by (1.2) the GSR-RDEU. The GSR-CPT model is first introduced by Mao and Cai [29] although their actual definition is slightly more complex. Since the weighting functions distort/modify the cumulative distribution function of $v(\xi)$, $\mathbb{E}_{w^-}^{w^+}[v(\xi)]$ is also called the CPT-based distortion risk measure and consequently the GSR-CPT is also known as a generalized shortfall risk measure induced by the CPT based distortion risk measure, see [29] for details. Mao and Cai [29] motivate their CPT based GSR by interpreting it as the regulator’s requirement of the minimal capital to cover a risk faced by a financial institution or an insurance company. In that case, the regulator may have different attitudes to the shortfall risk and the over-required capital risk and subsequently adopt different value functions and distortion functions to evaluate their losses and related probabilities.

The formulation of SR in (1.1) is closely related to insurance premium calculation model. To see this, consider the case that $\xi : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}_+$ where $\xi$ represents an insurer’s liability and $u : \mathbb{R} \rightarrow \mathbb{R}$
is a non-decreasing utility function. The insurance premium problem is to find \( x =: \Pi(\xi) \) such that \( x \) solves the indifference equation

\[
\mathbb{E}[u(c_0 + x - \xi)] = u(c_0), \tag{1.4}
\]

where \( c_0 \) is the initial wealth of the insurer and the indifference equation means that the insurer’s expected utility of the overall wealth is unchanged after committing to the insurance. Under some moderate conditions of \( u \), the indifference equation is equivalent to

\[
\Pi(\xi) = \inf \{ x \in \mathbb{R} : \mathbb{E}_p[-u(c_0 - \xi + x)] \leq -u(c_0) \}, \tag{1.5}
\]

where the minimum is taken in the sense that the insurer is keen to minimize the premium in order to attract more customers. Obviously (1.5) corresponds to the formulation (1.1) with \( l(y) = -u(-y) \), \( \lambda = -u(c_0) \) and \( c_0 = 0 \), see details in our earlier paper [43]. Like the definition of SR, the insurance premium model (1.4) is also based on VNM’s expected utility theory and hence has drawbacks as the SR model.

Heilpern [21] seems to be the first to realize the issue and propose a new insurance premium calculation model based on the RDEU. Specifically he defines a RDEU generalization of zero utility principle as the solution of

\[
\mathbb{E}_w[u(c_0 + x - \xi)] = u(c_0),
\]

where \( \mathbb{E}_w[u(\xi)] := \int \mathbb{R} u(x)dw(F_\xi(x)) \) based on a single weighting function \( w \), and derives the properties of the defined insurance premium, such as translation invariance, positive homogeneity, independent additiveness and sub-additiveness. Kaluszka and Krzyszowiec [25] extend the formulation to CPT based premium principal by lifting the constraints on the relationship between the weighting functions, and demonstrate that their premium principle enjoys the important properties such as non-excessive loading, no unjustified risk loading and translation invariance and positive homogeneity. Nardon and Pianca [30] introduce a premium principal based on the continuous cumulative prospect theory, show that the principal retains translation invariance and positive homogeneity and discuss several applications to the pricing of insurance contracts. Since the new definitions can be put under the general mathematical framework of (GSR-CPT), we regard all of them as generalized SR although their focuses might be different.

In this paper, we take on this stream of research but with a different focus. We consider a situation where the value function and/or the weighting functions in (GSR-CPT) are ambiguous. As noted by Mao and Cai in the motivation of their GSR model, there could be a case that the regulator of a market might have a number of value/weighting functions in mind but is short of complete information as to which one may capture its true risk preference. If the number of plausible value/weighting functions is finite, then one may use model (1.2) to compute the generalized shortfall risk measure one by one and then pick up the one which suits the DM’s preference best. However, this approach may not work when the number is infinite. Likewise, the primary goal of incorporating CPT in the insurance premium calculation model is to better reflect the insurer’s risk preference. However, when it comes to practical applications of the model, it is often difficult to determine which particular value function and/or pair of weighting functions will capture the insurer’s true risk preference. In other words, when the (GSR-CPT) model is applied to practical decision making problems, there could be an ambiguity of which particular tuple of value functions and weighting functions at hand to be used to cater for the decision maker’s risk preference.
There are potentially two ways to do this. One is to use partially available information such as pairwise comparisons to identify the values of the value functions and/or the weighting functions at a discrete set of points and then construct an approximate value function and a pair of approximate weighting functions by interpolation. This kind of approach is widely used in economics for identifying an approximate utility function, see for instance Clemen and Reilly [8]. The other is to use partially available information to construct a set of plausible value functions/weighting functions [23, 40] and base the shortfall risk on the worst ones. In this paper, we proceed with the latter which is known as robust approach. We call it preference robust model to distinguish it from distributionally robust models in the literature of shortfall risk measures and distortion measures (see e.g. Guo and Xu [18], Cai et al. [7], and Pesenti et al. [33]). Delage et al. [11] seem to be the first to consider a preference robust version of SR (1.1). By assuming that the decision maker’s risk preference can be characterized by SR or more generally convex/coherent risk measures (when \( l \) is convex, SR is a convex risk measure), they use DM’s risk preference information elicited via certainty equivalents and the DM’s risk preference on large tail losses to construct an ambiguity set of loss functions and then use the latter to define a preference robust SR. This paper takes a step further to consider the (GSR-CPT) model.

To this end, we assume the decision maker’s risk preference can be represented by the GSR based on CPT which is not convex in general. We then use the DM’s preferences characterized by the GSR to identify a class of value/weighting functions and subsequently use the latter to define/calculate the preference robust GSR (PRGSR). There two ways to define PRGSR. One is to use each pair of value/weighting functions from the ambiguity to calculate a GSR-CPT via (1.2) and then consider the worst one (with largest value), see (2.3). The other is to replace \( E_{w-w^+}[v(\xi - x)] \) in (1.2) by the worst distorted expected value function in (1.2) and define the corresponding shortfall risk value (the optimal value) as the PRGSR, see (2.1). We adopt the latter and demonstrate that the two definitions are equivalent (Proposition 2.1).

A key ingredient of the research is construction of the ambiguity set of value/weighting functions. Differing from Delage et al. [11], here a DM’s preference is described by both the value function and the weighting functions. As such, we must infer not only the value function (corresponding to the loss function in the SR model) but also the weighting functions based on the decision maker’s preferences. Moreover, the CPT based SR is not a convex risk measure in general when the value functions are strictly S-shaped. All these pose new challenges to the construction of the ambiguity set and development of computational schemes for calculating the preference robust shortfall risk measure, this paper aims to address them. We propose to use certainty equivalent and pairwise comparison approaches to infer the value function for a fixed pair of weighting functions and then for a fixed value function, we construct a ball of weighting functions centered at a nominal weighting function under some pseudo-metric. The latter is based on the fact that in the literature of behavioural economics, weighting functions of decision makers are mostly similar [38].

Another important component of this work is to develop computational methods which can be effectively used to calculate the proposed preference robust shortfall risk measure. To this end, we derive reformulations of the robust shortfall risk measure when the ambiguity set of value functions and the ambiguity set of weighting functions take specific forms. Since the choice of worst case value function and worst case weighting functions inevitably results in an optimization problem with bilinear objective function, we reformulate the latter as a linear program by replacing the bilinear terms with new variables. To examine the effectiveness of the ways proposed for constructing the ambiguity
sets and the subsequent computational schemes, we carry out some numerical tests on an artificially constructed academic example. The preliminary results show that convergence of the worst case value functions and weighting functions to their true counterparts as the ambiguity reduces (with more preference information being elicited or the radius of the ball of weighting functions driven to zero).

The rest of the paper is organized as follows. Section 2 introduces the preference robust GSR model and discusses the equivalent definition and the main properties of the robust GSR. Section 3 discusses construction of the ambiguity sets of both the value functions and the weighting functions. Section 4 presents details of computational schemes for calculating the preference robust shortfall risk measure via reformulations. Finally, Section 5 reports numerical test results.

Throughout the paper, we use $V$ to denote the set of all value functions $v : \mathbb{R} \to \mathbb{R}$, which are strictly increasing with $v(0) = 0$. To facilitate properties of the value function over the negative half line and positive half, we write $v^-(t)$ for $v(t)$ when $t < 0$, and $v^+(t)$ for the value function when $t > 0$. Furthermore, we use $W$ to denote the set of all weighting/distortion functions $w : [0, 1] \to [0, 1]$, which are strictly increasing with $w(0) = 0, w(1) = 1$, $W_{inS}$ the subset of weighting functions which are inverse S-Shaped, that is, concave in $(0, p^*)$ and convex in $(p^*, 1)$ with some $p^* \in (0, 1)$. In some cases, we need to discuss the derivatives of the weighting functions and consequently use $\psi^-(t)$ and $\psi^+(t)$ to denote the derivatives of $w^-$ and $w^+$ respectively. Finally, by convention, we use $L^0(\mathbb{R})$ to represent the space of measurable random variables over $\mathbb{R}$.

In this paper, we use a number of acronyms for both new notions and mathematical models to facilitate exposition/reading. Below is a list.

- SR – shortfall risk measure
- CPT – cumulative prospect theory
- GSR – generalized shortfall risk measure
- GSR-CPT – generalized shortfall risk measure based on cumulative prospect theory
- GSR-RDEU – generalized shortfall risk measure based on rank dependent expected utility theory
- PRGSR – preference robust generalized shortfall risk measure
- PRGSR-CPT – preference robust generalized shortfall risk measure based on cumulative prospect theory
- Worst-GSR-CPT – worst generalized shortfall risk measure based on cumulative prospect theory
- PRGSR-CPT-S – specific preference robust generalized shortfall risk measure based on cumulative prospect theory
- PRGSR-CPT-S-Dis – specific preference robust generalized shortfall risk measure based on cumulative prospect theory for discrete random variable

When we refer to the mathematical formulation of a particular concept, we use a round parenthesis, e.g., (SR) refers to the shortfall risk model defined in (1.1).
2 Preference robust GSR-CPT

We begin by a formal definition of the preference robust generalized shortfall risk measure based on CPT (PRGSR-CPT).

**Definition 2.1 (PRGSR-CPT)** Let $\mathcal{V} \times \mathcal{W} \times \mathcal{W} \subset \mathcal{V} \times \mathcal{W} \times \mathcal{W}$ be ambiguity sets of the value function and weighting functions. The preference robust generalized shortfall risk measure based on cumulative prospect theory of random variable $\xi \in \mathcal{L}^0(\mathbb{R})$ associated with $\mathcal{V} \times \mathcal{W} \times \mathcal{W}$ is defined as

\[
\rho_{\mathcal{V} \times \mathcal{W}}(\xi) = \inf_{x \in \mathbb{R}} \{ x : \sup_{(v, w^-, w^+) \in \mathcal{V} \times \mathcal{W} \times \mathcal{W}} E_{w^- - w^+} [v(\xi - x)] \leq 0 \}.
\] (2.1)

In this definition, the worst case distorted expected value of $v(\xi - x)$ is used in the constraint. Here we make a blanket assumption that the robust constraint function in (2.1) is well-defined. Note that in formulation (2.1), we focus on the case that $w^-, w^+$ belong to the same ambiguity set $\mathcal{W}$ to simplify the discussions. All of the results developed in this paper can be applied to ambiguity sets of weighting functions without this restriction. Note also that the ambiguity sets in this definition are given. We will come back shortly to discuss about how they may be constructed based on a decision maker’s preference information.

There is an alternative way to define the PRGSR. For each tuple $(v, w^-, w^+) \in \mathcal{V} \times \mathcal{W} \times \mathcal{W}$, we calculate the CPT based shortfall risk measure:

\[
\rho_{(v, w^-, w^+)}(\xi) := \inf_{x \in \mathbb{R}} \{ x : \sup_{(v, w^-, w^+) \in \mathcal{V} \times \mathcal{W} \times \mathcal{W}} E_{w^- - w^+} [v(\xi - x)] \leq 0 \}.
\] (2.2)

and then define the PRGSR as the worst (largest) GSR-CPT:

\[
\Pi_{\mathcal{V} \times \mathcal{W}}(\xi) := \sup_{(v, w^-, w^+) \in \mathcal{V} \times \mathcal{W} \times \mathcal{W}} \rho_{(v, w^-, w^+)}(\xi).
\] (2.3)

Note that the GSR-CPT model defined in (2.2) differs slightly from the one in (1.2) in that we set $\lambda = 0$ here. This is not only for the simplification in the forthcoming discussions but also aligning the formulation to the definition of GSR in Mao and Cai [29] (see Definition 2.1 there). In the rest of the paper, (GSR-CPT) always refers to problem (2.2). Note also that under some circumstances, i.e., $\xi$ is restricted to taking non-negative values, $w^-(p) = 1 - w^+(1 - p)$, $\forall p \in [0, 1]$ and $v(-t) = -v(t)$, $\forall t \in \mathbb{R}$, $E_{w^- - w^+} [v(\xi - x)] \leq 0$ is equivalent to $E_{w^- - w^+} [v(x - \xi)] \geq 0$. Consequently the (GSR-CPT) reduces to CPT based insurance premium calculation model [43].

The next proposition states that the two definitions (see (PRGSR-CPT) in (2.1) and (Worst-GSR-CPT) in (2.3)) are equivalent.

**Proposition 2.1** Let $\rho_{\mathcal{V} \times \mathcal{W}}(\xi)$ and $\Pi_{\mathcal{V} \times \mathcal{W}}(\xi)$ be defined as in (2.1) and (2.3) respectively. Then

\[
\rho_{\mathcal{V} \times \mathcal{W}}(\xi) = \Pi_{\mathcal{V} \times \mathcal{W}}(\xi).
\] (2.4)
\textbf{Proof.} We begin by showing that the equality (2.4) holds when one of the quantities is $+\infty$. First, consider the case that $\rho_{V \times W}(\xi) = +\infty$. Then the feasible set of problem (2.1) is empty. Assume for the sake of a contradiction that $\hat{x} := \Pi_{V \times W}(\xi) < +\infty$. Then $\rho_{(v, w^{-}, w^{+})}(\xi) \leq \hat{x} < +\infty$ for all $(v, w^{-}, w^{+}) \in V \times W \times W$. Thus for any positive number $\epsilon$, it follows by Lemma A.2 (i) that
\[ E_{w^{-}, w^{+}}[v(\xi - \hat{x} - \epsilon)] \leq E_{w^{-}, w^{+}}[v(\xi - \rho_{(v, w^{-}, w^{+})}(\xi) - \epsilon)] \leq 0, \forall (v, w^{-}, w^{+}) \in V \times W \times W. \]
The latter implies $\hat{x} + \epsilon$ is a feasible solution to the problem (2.1), a contradiction.

Conversely, consider the case that $\Pi_{V \times W}(\xi) = +\infty$. Assume for a contradiction that $x^* := \rho_{V \times W}(\xi) < +\infty$. Then for any positive number $\epsilon$, $x^* + \epsilon$ is a feasible solution to problem (2.1), that is,
\[ \sup_{(v, w^{-}, w^{+}) \in V \times W \times W} E_{w^{-}, w^{+}}[v(\xi - x^* - \epsilon)] \leq 0, \]
which implies
\[ E_{w^{-}, w^{+}}[v(\xi - x^* - \epsilon)] \leq 0, \forall (v, w^{-}, w^{+}) \in V \times W \times W. \]
The latter guarantees that $\rho_{(v, w^{-}, w^{+})}(\xi) \leq x^* + \epsilon, \forall (v, w^{-}, w^{+}) \in V \times W \times W$ and hence $\Pi_{V \times W}(\xi) \leq x^* + \epsilon < +\infty$, a contradiction.

We now turn to discuss the case that the optimal values of the programs in (2.1) and (2.3) are both finite. Let $x$ be any feasible solution of problem (2.1). Then $x$ is also a feasible solution of problem (2.2) for any $(v, w^{-}, w^{+}) \in V \times W \times W$, which implies $x^* \geq \hat{x}$.

Conversely, for any fixed $(v, w^{-}, w^{+}) \in V \times W \times W$, let $x_{(v, w^{-}, w^{+})}$ be a feasible solution to problem (2.2). Then by definition $\hat{x} \geq x_{(v, w^{-}, w^{+})}$. It follows by Lemma A.2 (i) that
\[ E_{w^{-}, w^{+}}[v(\xi - \hat{x})] \leq E_{w^{-}, w^{+}}[v(\xi - x_{(v, w^{-}, w^{+})})] \leq 0. \]
This shows $\hat{x}$ is a feasible solution of problem (2.1). This shows $\hat{x} \geq x^*$.

Next, we discuss the properties of the newly introduced PRGSR-CPT. First, it is known that the GSR-CPT is a law invariant monetary risk measure (satisfying monotonicity and translation invariance), see [29]. Since these properties are preserved under “sup” operation, then $\Pi_{V \times W}(\cdot)$ is also a law invariant monetary risk measure. We can therefore conclude, via Proposition 2.1, that PRGSR-CPT is a law invariant monetary risk measure. Second, it might be interesting to discuss whether PRGSR-CPT is a convex risk measure. Recall that Mao and Cai [29] derive conditions under which GSR-CPT is a convex risk measure. The next proposition states that if we use the type of convex value functions and convex/concave weighting functions specified by Mao and Cai [29] to define the ambiguity sets $V$ and $W$, then the resulting preference robust SR is also a convex risk measure.

\textbf{Proposition 2.2} Let $R_{\text{con}} := \{\rho_{(v, w^{-}, w^{+})} : \rho_{(v, w^{-}, w^{+})} \text{ is a convex risk measure}\};$
\[ \rho_{R_{\text{con}}} (\xi) = \sup_{\rho_{(v, w^{-}, w^{+})} \in R_{\text{con}}} \rho_{(v, w^{-}, w^{+})}(\xi). \]

Define
\[ \rho_{V \times W_{-} \times W_{+}}(\xi) := \inf_{x \in \mathbb{R}} \left\{ x : \sup_{(v, w^{-}, w^{+}) \in V \times W_{-} \times W_{+}} E_{w^{-}, w^{+}}[v(\xi - x)] \leq 0 \right\}. \tag{2.5} \]
If $V_1 \subset V$ is the set of all value functions satisfying that $v$ is convex on both $\mathbb{R}_-$ and $\mathbb{R}_+$, $W_{\text{con}} \subset W$ is the set of all convex weighting functions, $W_{\text{cav}} \subset W$ is the set of all concave weighting functions, and

$$V_{\text{con}} \times W_1 \times W_2 := \left\{(v, w^-, w^+) \in V_1 \times W_{\text{con}} \times W_{\text{cav}} : \frac{v'_+(0)}{v'_-(0)} \geq \sup_{p \in (0,1)} \frac{(w^-)'(p)}{(1 - w^+(1 - \cdot))'_+(p)} \right\}, \quad (2.6)$$

where $v'_-(0)$ and $v'_+(0)$ are the left and right derivatives of value function $v$ at point $0$, $(w^-)'(p)$ and $(1 - w^+(1 - \cdot))'_+(p)$ are the left and right derivatives of $w^-$ and $1 - w^+(1 - \cdot)$ at point $p$ respectively, then

$$\rho_{V_{\text{con}} \times W_1 \times W_2}(\xi) = \rho_{R_{\text{con}}}(\xi).$$

**Proof.** By Proposition 2.1,

$$\rho_{V_{\text{con}} \times W_1 \times W_2}(\xi) = \sup_{(v, w^-, w^+) \in V_{\text{con}} \times W_1 \times W_2} \rho_{(v, w^-, w^+)}(\xi).$$

By [29, Corollary 2.8], a GSR-CPT $\rho_{(v, w^-, w^+)}(\xi)$ is a convex risk measure if and only if $w^-(p)$ and $1 - w^+(1 - p)$ are convex in $p \in [0,1]$, and $v$ is convex on both $\mathbb{R}_-$ and $\mathbb{R}_+$, and $(v, w^-, w^+)$ satisfying the inequality in (2.6), which means that for any $\rho_{(v, w^-, w^+)}(\xi) \in R_{\text{con}}$, we must have $v \in V_{\text{con}}, w^- \in W_1, w^+ \in W_2$. Thus $\rho_{(v, w^-, w^+)}(\xi) \leq \rho_{V_{\text{con}} \times W_1 \times W_2}(\xi)$ and hence

$$\rho_{R_{\text{con}}}(\xi) = \sup_{\rho_{(v, w^-, w^+)} \in R_{\text{con}}} \rho_{(v, w^-, w^+)}(\xi) \leq \rho_{V_{\text{con}} \times W_1 \times W_2}(\xi).$$

Conversely for any $v \in V_{\text{con}}, w^- \in W_1, w^+ \in W_2$, $\rho_{(v, w^-, w^+)} \in R_{\text{con}}$, and hence

$$\rho_{V_{\text{con}} \times W_1 \times W_2}(\xi) = \sup_{(v, w^-, w^+) \in V_{\text{con}} \times W_1 \times W_2} \rho_{(v, w^-, w^+)}(\xi) \leq \rho_{R_{\text{con}}}(\xi).$$

The proof is complete.

We note that the PRGSR-CPT defined in (2.5) is slightly different from PRGSR-CPT in (2.1) by allowing $w^-$ and $w^+$ to have different ambiguity set. This is because in the setup of the GSR-CPT models, $w^-(p)$ and $1 - w^+(1 - p)$ are convex in $p \in [0,1]$ and consequently $w^+(p)$ is a concave function.

Mao and Cai [29] also derive conditions under which the GSR-CPT is a coherent risk measure. We can easily use their result for establishing coherence of PRGSR-CPT (2.5).

**Proposition 2.3** Let $\rho_{V \times W_1 \times W_2}(\xi)$ be defined as in (2.5). Define

$$R_{\text{coh}} := \{\rho_{(v, w^-, w^+)} : \rho_{(v, w^-, w^+)} \text{ is a coherent risk measure}\},$$

$$\rho_{R_{\text{coh}}}(\xi) := \sup_{\rho_{(v, w^-, w^+)} \in R_{\text{coh}}} \rho_{(v, w^-, w^+)}(\xi),$$

$$V_{\text{coh}} := \{v \in V \mid v(x) = a_1 x \text{ for } x \geq 0, v(x) = a_2 x \text{ for } x < 0, a_1, a_2 > 0\}.$$  

If $W_1 \subset W$ is the set of all convex weighting functions, $W_2 \subset W$ is the set of all concave weighting functions, then

$$\rho_{V_{\text{coh}} \times W_1 \times W_2}(\xi) = \rho_{R_{\text{coh}}}(\xi).$$

8
Proof. Observe first that since \( w^-, w^+ \) are increasing, then \( a_1 w^-(p) - a_2 (1 - w^+(1 - p)) \) is also increasing over \([0, 1]\) for all \( a_1, a_2 > 0 \). By [29, Theorem 2.10], \( \rho_{(v,w^-,w^+)} \) is coherent if and only if \( w^-(p) \) and \( 1 - w^+(1 - p) \) are convex and \( v \in V_{\text{coh}} \). Moreover, it follows by Proposition 2.1 that \( \rho \) can be rewritten as \( \sup_{(v,w^-,w^+) \in V_{\text{coh}}} \rho_{(v,w^-,w^+)}(\xi) \). For any \( \rho_{(v,w^-,w^+)} \in R_{\text{coh}} \), we must have \((v,w^-,w^+) \in V_{\text{coh}} \times W_1 \times W_2 \) and consequently

\[
\rho_{(v,w^-,w^+)}(\xi) \leq \rho_{v_{\text{coh}} \times W_1 \times W_2}(\xi) \quad \text{and} \quad \rho_{R_{\text{coh}}}(\xi) \leq \rho_{v_{\text{coh}} \times W_1 \times W_2}(\xi).
\]

Conversely, for any \((v,w^-,w^+) \in V_{\text{coh}} \times W_1 \times W_2 \), \( \rho_{(v,w^-,w^+)} \in R_{\text{coh}} \). Thus

\[
\rho_{(v,w^-,w^+)} \leq \rho_{R_{\text{coh}}} \quad \text{and} \quad \rho_{v_{\text{coh}} \times W_1 \times W_2}(\xi) \leq \rho_{R_{\text{coh}}}(\xi).
\]

The proof is complete. \[ \blacksquare \]

In the literature of prospect theory [24, 38], the value functions are often S-shaped, which means that they are not convex if they are strictly S-shaped. In this case, the GSR-CPT defined in (2.2) is not convex and consequently the PRGSR-CPT defined in (2.5) is not a convex risk measure. However, if the value function \( v \) is not strictly S-shaped, then the value function may still satisfy the conditions specified in the definition of \( V_{\text{con}} \) (see Proposition 2.2). The next example is provided by Mao in a private communication.

Example 2.1 Let \( w^-(p) = 1 - w^+(1 - p) \), \( \forall p \in [0, 1] \) be differentiable and convex, let the value function \( v \) be convex over \( \mathbb{R}_- \) and linear over \( \mathbb{R}_+ \), i.e.,

\[
v(x) = \begin{cases} v^-(x) & \text{for } x < 0, \\ 0 & \text{for } x = 0, \\ kx & \text{for } x > 0, \end{cases}
\]

where \( v^- \) is convex over \( \mathbb{R}_- \) and the slope \( k \) satisfies that \( \frac{k}{v^-(0)} \geq \sup_{p \in (0,1)} \frac{(w^-)'(p)}{(1-w^+(1-p))_+(p)} \). Since

\[
\sup_{p \in (0,1)} \frac{(w^-)'(p)}{(1-w^+(1-p))_+(p)} = 1,
\]

then \( v \) is actually convex over the whole real line \( \mathbb{R} \). This shows a special S-shaped (actually convex) value function satisfying the conditions of \( V_{\text{con}} \).

The PRGSR-CPT model covers two important preference robust risk measure as special cases.

1. **Preference robust utility-based shortfall risk measures.** If the ambiguity set of weighting functions only contains single linear function, i.e., \( w(p) = p \), \( \forall p \in [0, 1] \), then \( \Pi_{\nu}(\xi) := \sup_{v \in V} \rho_{(v,w^-,w^+)}(\xi) \). It follows from Proposition 2.1 that the above risk measure is equivalent to the preference robust utility-based shortfall risk measures

\[
\rho_{\nu}(\xi) := \inf_{x \in \mathbb{R}} \left\{ x : \sup_{v \in V} E[v(\xi - x)] \leq 0 \right\},
\]

which has been fully investigated by Delage et al. [11].

2. **Preference robust distortion risk measure.** If the ambiguity set of value functions is a singleton, i.e., \( v(t) = t \), \( \forall t \in \mathbb{R} \), then

\[
\Pi_{\nu}(\xi) := \sup_{(w^-,w^+) \in W \times W} \rho_{(v,w^-,w^+)}(\xi).
\]
By Proposition 2.1, the above problem is equivalent to

\[
\rho_W(\xi) := \inf_{x \in \mathbb{R}} \left\{ x : \sup_{(w^-,w^+) \in W \times W} \mathbb{E}_{w^-w^+}[\xi - x] \leq 0 \right\}.
\]

(2.8)

Since \(\mathbb{E}_{w^-w^+}[\xi]\) is essentially about distortion/modification of the CDF of \(\xi\), \(\rho_W(\xi)\) may be viewed as a shortfall risk measure based on Yaari’s dual theory of choice. Moreover, it is well-known that \(\mathbb{E}_{w^-w^+}[\xi]\) is a coherent risk measure when \(w^-(p) = 1 - w^+(1 - p)\) and \(w^-(p)\) is convex, which means \(\mathbb{E}_{w^-w^+}[\xi - x] = \mathbb{E}_{w^-}[\xi] - x\). In this case, \(\Pi_W(\xi)\) coincides with the preference robust distortion risk measure \(\sup_{w^- \in W} \int_{\mathbb{R}} t d\mathbb{E}(F_\xi(t))\) [39].

Before concluding this section, we note that the reason why we define the PRGSR via (2.1) rather than (2.3) is that the former can be solved by tractable reformulations whereas the latter cannot. However, formulation (2.3) connects more directly to decision maker’s preferences. To see this, consider the case that the decision maker prefers prospect \(A\) over prospect \(B\). Assuming that the DM’s preference can be represented by the CPT based GSR \(\rho(v, w^-, w^+)(\xi)\), then we have

\[
\rho(v, w^-, w^+)(A) < \rho(v, w^-, w^+)(B).
\]

(2.9)

The inequality enables us to identify a class of value/weighting functions \(v, w^-, w^+\), i.e., the ambiguity sets \(V\) and \(W\) and subsequently define PRGSR \(\Pi_{V \times W}(\xi)\). To illustrate the idea, consider a simple case that the DM’s preference may be represented by three pairs of value/weighting functions \(\{(v_i, w_i^-, w_i^+) : i = 1, 2, 3\}\). It is found that \((v_1, w_1^-, w_1^+)\) and \((v_2, w_2^-, w_2^+)\) satisfy (2.9) but \((v_3, w_3^-, w_3^+)\) does not. That is, \(V \times W = \{(v_i, w_i^-, w_i^+) : i = 1, 2\}\). Let \(\xi\) be a new prospect. Then

\[
\Pi_{V \times W}(\xi) := \max_{i=1,2} \rho(v_i, w_i^-, w_i^+) \rho(\xi)
\]

(2.10)

and

\[
\rho_{V \times W}(\xi) := \inf_{x \in \mathbb{R}} \left\{ x : \sup_{i=1,2} \mathbb{E}_{w_i^-w_i^+}[v_i(\xi - x)] \leq 0 \right\}.
\]

(2.11)

In this case, we can easily calculate the PRGSR via (2.10). In practice, however, the ambiguity sets \(V \times W \times W\) is not a finite set and consequently it is very difficult to calculate \(\Pi_{V \times W}(\xi)\). In contrast, (2.1) can be reformulated as a linear program when the ambiguity sets have some specific structure. We will come back to this in Section 4. From this simple example, we can see how we may use the equivalence between \(\Pi_{V \times W}(\xi)\) and \(\rho_{V \times W}(\xi)\) to define and calculate the PRGSR: using the DM’s information represented by GSR-CPT to construct the ambiguity sets \(V \times W \times W\) and then use (2.1) to calculate the PRGSR. The flow chart in Figure 1 illustrates the procedures for which we will give rise to details in the next two sections.

3 Construction of the ambiguity sets

A key ingredient of the PRGSR-CPT model (2.1) is construction of the ambiguity sets of the value functions and weighting functions because it concerns collection of a DM’s preference information.
and calculation of the robust risk measure. In general, it is difficult to construct the ambiguity sets of the value functions and weighting functions simultaneously. In this section, we first discuss the construction of the ambiguity set $\mathcal{V}$ of the value functions for a fixed weighting function and then the ambiguity set $\mathcal{W}$ of weighting functions. Throughout this section, we make the following assumption.

**Assumption 3.1** The decision maker’s preference can be described by GSR-CPT $\rho_{(v,w^-,w^+)}(\xi)$ but there is an ambiguity about how to identify the true value function and weighting function such that $\rho_{(v,w^-,w^+)}(\xi)$ precisely captures the DM’s risk preference.

As in the existing works of PRO models (see e.g. [3, 22, 19]), there are two types of information for identifying the true unknown value function. One is generic information about the shape of the value functions such as convexity, concavity, S-shapedness which is shared by a majority of decision makers in most decision making problems. The other is decision maker/problem specific information which has to be elicited through various approaches such as pairwise comparisons, certainty equivalents and past decisions.

In contrast, the ambiguity set of the weighting functions may be constructed in a different manner given the fact that one can usually use empirical data or subjective judgement to construct a nominal weighting function. Due to the limitation of data, the nominal weighting function might not be accurate and hence one can usually construct a ball of weighting functions centered at the nominal as an ambiguity set. The larger the ball, the more likely the true distortion function is included, see Wang and Xu [40] for the distortion risk measure optimization.

We propose to construct the ambiguity sets of value functions and weighting functions in three steps. First, we extend the certainty equivalent approach in terms of utility-based shortfall risk measure to generalized shortfall risk measure based on CPT to translate the preferences relations of GSR-CPT into inequalities of the value function. Second, we use the classical preference elicitation approach, pairwise comparisons, to narrow down the ambiguity set by increasing the number of questionnaires and follow the experimental result in practise to confine the study to the case that the value function is strictly S-shaped. Third, we will consider a nominal weighting function and take into account the $\mathcal{d}_{\mathcal{W}}$-ball (Definition 3.3) centered at the nominal weighting function with radius.

### 3.1 Characterization of the value function via a DM’s certainty equivalent information

Under Assumption 3.1, we may use $\rho_{(v,w^-,w^+)}$ to describe a DM’s preferences. In order to construct an ambiguity set of the value functions, denoted by $\mathcal{V}$, for a fixed pair of weighting functions $(w^-, w^+)$, we begin by exploiting the sufficient and necessary conditions of value functions for the elicited information.
of GSR-CPT, as analyzed in [11], here we use certainty equivalent approach. Let

\[ R := \{\rho_{(v,w^-,w^+)} : \rho_{(v,w^-,w^+)}(\xi) \text{ is defined as in (2.2)}\}. \tag{3.1} \]

**Definition 3.1 (Certainty equivalent)** Let \( \{A_k\}_{k=1}^K \subset \mathcal{L}^0(\mathbb{R}) \) be a list of considered random variables and \( \{a_k^+\}_{k=1}^K, \{a_k^-\}_{k=1}^K \subset \mathbb{R} \) be two lists of non-random variables. The set \( \mathcal{R}_{ce}^K \subset \mathcal{R} \) denotes the set of GSR-CPT \( \rho_{(v,w^-,w^+) \in \mathcal{R}} \) satisfying an associated set of confidence intervals \([a_k^-, a_k^+] \subset [\text{ess inf } A_k, \text{ess sup } A_k]\) for the “certainty equivalent” of each \( A_k \) for \( k = 1, \ldots, K \), i.e.,

\[ \mathcal{R}_{ce}^K := \{\rho_{(v,w^-,w^+) : \mathcal{L}^0(\mathbb{R}) \rightarrow \mathbb{R} | a_k^- \leq \rho_{(v,w^-,w^+)}(A_k) \leq a_k^+, \forall k = 1, \ldots, K\}. \tag{3.2} \]

Note that (3.2) describes the case that the risk of each \( A_k \) to be lower than \( a_k^+ \) and bigger than \( a_k^- \). This motivates us to identify such two bounds for the riskiness of random variable \( A_k \) by taking the forms of questions such as [11, 39]:

- **Upper bound \( a_k^+ \):** “What is the smallest amount of money that you would decline to pay instead of being exposed to the risk of \( A_k \)?”
- **Lower bound \( a_k^- \):** “What is the largest amount of money that you would be willing to pay instead of being exposed to the risk of \( A_k \)?”

Note that if the answers to both questions give rise to an identical number, i.e., \( a_k^- = a_k^+ = \bar{a}_k \), then the certainty equivalent of \( A_k \) is identified with \( \rho_{(v,w^-,w^+)}(A_k) = \bar{a}_k \). In practice, it is more likely to have an interval \([a_k^-, a_k^+]\) for the equivalent value of random variable \( A_k \). Moreover, we assume that \([a_k^-, a_k^+] \subset (\text{ess inf } A_k, \text{ess sup } A_k)\) which is not restrictive. This is because we can substitute \( a_k^- > \max\{a_k^-, \text{ess inf } A_k\} \) and \( a_k^+ < \min\{a_k^+, \text{ess sup } A_k\} \) otherwise. The above certainty equivalent approach is an extension of Delag et al. [11] under VNM’s expected utility theory or Wang and Xu [39] under the distortion risk measure to CPT based shortfall risk measures.

\( \mathcal{R}_{ce}^K \) defines a set of GSRs each of which can be used to describe the DM’s risk preference expressed via certainty equivalent questionnaires. Our interest here is to use this information to identify a class of value functions \( \mathcal{V}_{ce}^K \subset \mathcal{V} \) such that \( \mathcal{V}_{ce}^K \) can be used for computation of PRGSR \( \rho_{\mathcal{V} \times \mathcal{W}}(\xi) \). The next proposition states this.

**Proposition 3.1 (Characterization of \( \mathcal{V} \) for \( \rho_{\mathcal{V}_{ce}^K}(\xi) \))** Let

\[ \rho_{\mathcal{V}_{ce}^K}(\xi) := \sup_{\rho_{(v,w^-,w^+) \in \mathcal{R}_{ce}^K}} \rho_{(v,w^-,w^+)}(\xi). \tag{3.3} \]

Define

\[ \mathcal{V}_{ce}^K := \left\{ v \in \mathcal{V} : \begin{array}{c} \mathbb{E}_{w^-,w^+}[v(A_k - a_k^+)] \leq 0, \\
\mathbb{E}_{w^-,w^+}[v(A_k - a_k^-)] \geq 0, \end{array} \text{ for } k = 1, \ldots, K \right\}. \tag{3.4} \]

Then

\[ \rho_{(v,w^-,w^+) \in \mathcal{R}_{ce}^K} \iff v \in \mathcal{V}_{ce}^K \tag{3.5} \]
and

\[ \rho_{\mathcal{R}_{ce}^K}(\xi) = \rho_{\mathcal{V}_{ce}^K}(\xi). \]  

(3.6)

**Proof.** The equality in (3.6) follows from the equivalence in (3.5) and Proposition 2.1. So it is enough to prove (3.5). Let

\[ \mathcal{V}_{ce}^K := \{ v \in \mathcal{V} : \rho_{(v,w^-,w^+)} \in \mathcal{R}_{ce}^K \}. \]  

(3.7)

We want to show \( \mathcal{V}_{ce}^K = \mathcal{V}_{ce}^K \). First, we show that \( \mathcal{V}_{ce}^K \subseteq \mathcal{V}_{ce}^K \). Let \( \hat{v} \in \mathcal{V}_{ce}^K \). Then \( \hat{v} \) satisfies the first inequality in (3.4), i.e.,

\[ \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^+)] \leq 0. \]  

(3.8)

We show that this inequality implies the second inequality in (3.2), i.e.,

\[ \rho_{(\hat{v},w^-,w^+)}(A_k) \leq a_k^+. \]  

(3.9)

This is evidenced by the fact that

\[
\begin{align*}
\rho_{(\hat{v},w^-,w^+)}(A_k) &= \inf \{ x \in X : \mathcal{E}_{w^-w^+}[\hat{v}(A_k - x)] \leq 0 \} \\
&= a_k^+ + \inf \{ x' \in \mathbb{R} : \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^+ - x')] \leq 0 \} \\
&\leq a_k^+,
\end{align*}
\]

(3.10)

where the last inequality is due to (3.8). Next, since \( \hat{v} \in \mathcal{V}_{ce}^K \) also implies

\[ \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^-)] \geq 0, \forall k = 1, \cdots, K, \]  

(3.11)

we will show that inequality (3.11) implies the first inequality in (3.2), i.e.,

\[ \rho_{(\hat{v},w^-,w^+)}(A_k) \geq a_k^- \].  

(3.12)

Let \( \Omega_k := \{ \omega \in \Omega : A_k(\omega) \leq a_k^- \} \). Since \( a_k^- \) is the essential infimum of \( A_k \), then \( \mathbb{P}(\Omega_k) > 0 \). Let \( \varepsilon > 0 \) be a positive number. By taking \( a = a_k^- - \varepsilon \) and \( b = a_k^- \), we have from Lemma A.2 (ii) that

\[ \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^- + \varepsilon)] = \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a)] > \mathcal{E}_{w^-w^+}[\hat{v}(A_k - b)] = \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^-)]. \]

Combining this with (3.11), we have

\[ \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^- + \varepsilon)] > 0. \]  

(3.13)

By exploiting (3.13), we have that for \( k = 1, \cdots, K, \)

\[
\begin{align*}
\rho_{(\hat{v},w^-,w^+)}(A_k) &= \inf \{ x \in \mathbb{R} : \mathcal{E}_{w^-w^+}[\hat{v}(A_k - x)] \leq 0 \} \\
&= a_k^- + \inf \{ x' \in \mathbb{R} : \mathcal{E}_{w^-w^+}[\hat{v}(A_k - a_k^- - x')] \leq 0 \} \\
&\geq a_k^-,
\end{align*}
\]

(3.14)

which proves (3.12) as desired. Combining (3.9), (3.12) and the definition of \( \mathcal{R}_{ce}^K \) in (3.2), we conclude that \( \rho_{(\hat{v},w^-,w^+)} \in \mathcal{R}_{ce}^K \) and hence \( \hat{v} \in \mathcal{V}_{ce}^K \). This completes the first part of the proof.
Next, we will show that $\mathcal{V}_{\text{ce}}^K \supseteq \mathcal{V}_{\text{R}}^K$. Let $\hat{v} \in \mathcal{V}_{\text{R}}^K$, which means $\rho(\hat{v}, w^-, w^+) \in \mathcal{R}_{\text{ce}}^K$ by (3.7), that is, (3.12) and (3.9) hold. By exploiting (3.9), i.e., $\rho(\hat{v}, w^-, w^+)(A_k) \leq a_k^+$, we will show (3.8) hold, i.e., $E_{w^-w^+}[\hat{v}(A_k - a_k^+)] \leq 0$. Since $\rho(\hat{v}, w^-, w^+)(A_k)$ is the solution of problem (2.2) with $\xi = A_k$, we have $\rho(\hat{v}, w^-, w^+)(A_k)$ satisfies the constraint, that is,

$$E_{w^-w^+}[\hat{v}(A_k - \rho(\hat{v}, w^-, w^+)(A_k))] \leq 0. \tag{3.15}$$

Let $\hat{\Omega}_k := \{\omega \in \Omega : A_k(\omega) \leq \rho(\hat{v}, w^-, w^+)(A_k)\}$, $a := \rho(\hat{v}, w^-, w^+)(A_k)$ and $b := a_k^+$. Then $a \leq b$. Moreover, since $\rho(\hat{v}, w^-, w^+)(A_k) \in [a_k^-, a_k^+] \subseteq \text{ess inf } A_k, \text{ess sup } A_k$ where the inclusion “$\subseteq$” is strict, then $P(\hat{\Omega}_k) > 0$. It follows from Lemma A.2 (i) that

$$E_{w^-w^+}[\hat{v}(A_k - \rho(\hat{v}, w^-, w^+)(A_k))] = E_{w^-w^+}[\hat{v}(A_k - a)] \geq E_{w^-w^+}[\hat{v}(A_k - b)] = E_{w^-w^+}[\hat{v}(A_k - a_k^+)].$$

Likewise, by exploiting (3.12) i.e., $\rho(\hat{v}, w^-, w^+)(A_k) \geq a_k^-$, we prove (3.11) holds, i.e., $E_{w^-w^+}[\hat{v}(A_k - a_k^-)] \geq 0$, for $k = 1, \cdots, K$. Since $\rho(\hat{v}, w^-, w^+)(A_k)$ is the solution of problem (2.2) with $\xi = A_k$, we have $\rho(\hat{v}, w^-, w^+)(A_k)$ satisfies the indifference equation, that is,

$$E_{w^-w^+}[\hat{v}(A_k - \rho(\hat{v}, w^-, w^+)(A_k))] = 0. \tag{3.16}$$

Moreover, set $a = a_k^-$ and $b = \rho(\hat{v}, w^-, w^+)(A_k)$. Since $a \leq b$, it follows by Lemma A.2 (i) that

$$E_{w^-w^+}[\hat{v}(A_k - \rho(\hat{v}, w^-, w^+)(A_k))] = E_{w^-w^+}[\hat{v}(A_k - a)] \leq E_{w^-w^+}[\hat{v}(A_k - b)] = E_{w^-w^+}[\hat{v}(A_k - a_k^-)].$$

Combining with (3.16), we have (3.11) hold.

Thus by using (3.12) and (3.9), we derive (3.8) and (3.11), which means that for every $\hat{v} \in \mathcal{V}_{\text{R}}^K$, we have $\hat{v} \in \mathcal{V}_{\text{ce}}^K$. Thus $\mathcal{V}_{\text{R}}^K \subseteq \mathcal{V}_{\text{ce}}^K$. The proof is completed.

Representing $\rho(\mathcal{V}_{\text{ce}}^K)(\xi)$ as $\rho_{\mathcal{V}_{\text{ce}}}(\xi)$ is useful for solving (PRGR-CPT) since it reduces the evaluation of the PRGR-CPT to finding the optimal value of a stochastic programming problem with semi-infinite stochastic constraints:

$$\rho_{\mathcal{V}_{\text{ce}}}(\xi) = \inf_{x \in \mathbb{R}} \{x : E_{w^-w^+}[v(\xi - x)] \leq 0, \forall v \in \mathcal{V}_{\text{ce}}^K\}. \tag{3.17}$$

In Section 4, we will address the computational challenges arising from this reformulation for the case that $\xi$ has a finite discrete distribution.

### 3.2 Construction of $\mathcal{V}$ via pairwise comparison approach

In the literature of behavioural economics and preference robust optimization, a widely used approach for eliciting a decision maker’s preference is pairwise comparison, that is, the decision maker is given a pair of prospects for choice and the outcome of the choice is used to infer the decision maker’s utility/risk preference. Here we use the same approach for constructing an ambiguity set of the value functions assuming the weighting functions are known and fixed. Differing from the discussions in the preceding subsection, here we will not use CPT-based shortfall risk measure to represent the DM’s preference choice, rather we will use the distorted expected values of prospects, that is, $E_{w^-w^+}[v(\cdot)]$ to describe the DM’s choice. For example, if the DM’s is found to prefer prospect $B$ over $A$, then we claim
that $E_{w-w^+}[v(A)] \leq E_{w-w^+}[v(B)]$. This is because we are unable to use $\rho_{(v,w-w^+)}(A) \leq \rho_{(v,w-w^+)}(B)$ to infer the corresponding properties of the unknown true value function $v(\cdot)$, see [11] for the same issues in the preference robust SR models. To justify this approach, we note that when DM prefers $B$ to $A$, he/she would also prefer $B - t$ to $A - t$ for any constant $t$. Consequently we have

$$E_{w-w^+}[v(A-t)] \leq E_{w-w^+}[v(B-t)], \forall t \in \mathbb{R}.$$  

This inequality implies $\rho_{(v,w-w^+)}(A) \leq \rho_{(v,w-w^+)}(B)$. The converse argument might not hold. So we may regard the preference represented by functional $E_{w-w^+}[v(\cdot)]$ as CPT-based preference representation, in other words, the DM’s preference in this case can be represented by a “stronger” quantitative indicator.

Let $\{\gamma_m, \eta_m\}_{m=1}^M$ be a set of random variable. Suppose that the DM is found to prefer $\eta_m$ over $\gamma_m$ for $m = 1, \cdots, M$. Based on the discussions above, we can use the preference information to define an ambiguity set of the value functions

$$\mathcal{V}^M_{pair} := \{v \in \mathcal{V} : E_{w-w^+}[v(\gamma_m)] \leq E_{w-w^+}[v(\eta_m)], m = 1, \cdots, M\}. \quad (3.18)$$

In some cases, it might be possible to know roughly the shape of the value function such as convexity, concavity and $S$-shapedness. In the literature of behavioural economics, $S$-shapedness of a value function is widely adopted. Let

$$\mathcal{V}_S := \{v \in \mathcal{V} : v^+ \text{ is concave, } v^- \text{ is convex}\}. \quad (3.19)$$

This kind of information can usually be more easily elicited. For example, if a DM is found to be risk taking in losses and risk averse in gains, then the value function should be $S$-shaped.

Summarizing the elicitation approaches that we have discussed, we may define the ambiguity set of the value functions broadly as

$$\mathcal{V}_{EI} := \mathcal{V}^K_{ce} \cap \mathcal{V}^M_{pair} \cap \mathcal{V}_S \quad (3.20)$$

albeit in practice we may only use part of them.

### 3.3 Construction of ambiguity set $\mathcal{W}$ of weighting functions

Next, we discuss how to construct $\mathcal{W}$. The weighting function reflects the DM’s attitudes toward real probabilities of random losses/gains. A lot of discussions in behavioural economics are devoted to the topic. Empirical evidences suggest that the weighting functions in decision making problems are typically inverse-S shaped (initially concave over interval $(0, p^*)$ for some $p^* \in (0, 1)$ and then convex over $(p^*, 1)$). The inverse-S shape of weighting function reflects that the DM tends to overweight probabilities of extreme outcomes (including suffering large losses and small losses). For instance, in insurance pricing problem, this is consistent with individuals’ demand for insuring large losses and the eagerness to insure small losses and insurer should pay more attention on these two extreme outcomes.

In this section, we assume that the true weighting function of a decision maker is roughly known either based on empirical data or based on a subjective judgement. For example, Escobar and Pflug [14] propose an effective approach to construct a step-like distortion function with empirical data and use it
to approximate the true unknown distortion/weighting function in an insurance premium calculation model. We call such an estimated weighting function as nominal weighting function. Instead of using the nominal weighting function directly for the calculation of the preference robust shortfall risk in (2.1), we take a precaution to construct an ambiguity set of weighting functions near the nominal for (2.1). Here we consider a “ball” of weighting functions centered at the nominal weighting function under some pseudo-metric.

**Definition 3.2 (Pseudo-metric $d_{\tilde{g}}$)** Let

$$\mathcal{G} := \left\{ g : [0, 1] \rightarrow \mathbb{R}, g \text{ is measurable, } |g(t)| \leq \tilde{g}(t) \text{ for } t \in [0, 1] \text{ with } \int_0^1 \tilde{g}(t)dt < +\infty \right\}.$$  

The pseudo-metric $d_{\tilde{g}}$ of weighting functions in the space of $W$ is defined as

$$d_{\tilde{g}}(w_1, w_2) := \sup_{g \in \mathcal{G}} \left| \int_0^1 g(t)\psi_1(t)dt - \int_0^1 g(t)\psi_2(t)dt \right|,$$

where $\psi_1$ and $\psi_2$ are derivatives of the weighting functions $w_1$ and $w_2$ respectively.

Note that in this definition $d_{\tilde{g}}(w_1, w_2) = 0$ if and only if $\int_0^1 g(t)\psi_1(t)dt = \int_0^1 g(t)\psi_2(t)dt$ for all $g \in \mathcal{G}$, but $\psi_1$ is not necessarily equal to $\psi_2$ unless the set $\mathcal{G}$ is sufficiently large. If we interpret $\mathcal{G}$ as a set of test functions, then $d_{\tilde{g}}(w_1, w_2) = 0$ means that there is no difference between the two weighting functions for the specified set of test functions/cases. Moreover, we implicitly assume that the weighting functions are differentiable with at most a finite number of nondifferentiable points over $(0, 1)$. The reason why we use the pseudo-metric as opposed to other metrics/distance is because it is easy to use particularly in the derivation of tractable formulations of PRGSR-CPT in the next section.

Let $W_T \subset W_{inS}$ be the set of all piecewise linear inverse S-shaped functions $w : [0, 1] \rightarrow [0, 1]$ with breakpoints $0 = t_1 < \cdots < t_{T+1} = 1$ with $t_{p^*} = p^*$, where $l_{p^*} \in \{1, \cdots, T + 1\}$. Let

$$\Psi_T := \{ \psi : [0, 1] \rightarrow \mathbb{R}_+ : \psi \text{ is decreasing over } (0, p^*), \text{ increasing over } (p^*, 1) \}$$

be the set of the derivative functions of $W_T$. Obviously $\Psi_T$ is a set of left continuous step-like functions defined over $[0, 1]$ with jumps at the breakpoints $t_2, \cdots, t_T$. Then we can write $\Psi_T$ as:

$$\Psi_T = \left\{ \psi : [0, 1] \rightarrow \mathbb{R}_+ : \psi(t) = \sum_{i=1}^T \phi \mathbb{I}_{[t_i, t_{i+1})}(t), \sum_{i=1}^T \psi(t_{i+1} - t_i) = 1, \psi_{l+1} \leq \psi_1, l = 1, \cdots, l_{p^*} - 2, \psi_l \leq \psi_{l+1}, l = l_{p^*}, \cdots, T \right\},$$

where $\phi, l = 1, \cdots, T$ are the slopes of linear pieces of $w \in W_T$ and $\mathbb{I}_S$ is the indicator function of set $S$ with $\mathbb{I}_S(t) = 1$ for $t \in S$ and 0 otherwise.

**Definition 3.3 ($\tilde{g}$-ball of weighting functions)** Let $w_0 \in W_T$ be a weighting function which is piecewise linear and inverse S-shaped. The $\tilde{g}$-ball of inverse S-shaped weighting functions centered at $w_0$ with radius $r$ is defined as

$$B(w_0, r) := \{ w \in W_{inS} : d_{\tilde{g}}(w, w_0) \leq r \},$$

where the radius is related to the DM’s confidence about $w_0$.  

16
Note that $B(w_0, r)$ may contain inverse S-shaped weighting functions which are not piecewise linear. However, we will show later on that the worst case weighting function is piecewise linear, see Theorem 4.1. Note also that when radius $r$ is driven to 0, $B(w_0, r)$ shrinks to a singleton $\{w_0\}$ in the case that $d_\tilde{g}$ is a full metric. Based on the above comment, in this paper, for facilitating computation, we consider the ambiguity set of weighting functions as follows

$$W_r = B(w_0, r).$$

(3.23)

Note that in the case when $\tilde{g}$ takes a specific function, we can derive the $d_\tilde{g}$-ball explicitly.

Example 3.1 Let $\tilde{g}(t) = 1$. Then $\tilde{g}$-ball in Definition 3.3 collapses to the following ball in the $L_1$-metric sense:

$$B_{L_1}(w_0, r) := \{w \in \mathcal{W}_{inS} : d_{L_1}(w, w_0) \leq r\},$$

(3.24)

where $d_{L_1}(w, w_0) := \int_0^1 |\psi(t) - \psi^0(t)|dt$.

4 Computational schemes

In this section, we discuss how to calculate PRGSR-CPT $\rho_{V \times W}(\xi)$ for a given prospect $\xi$ when the ambiguity sets $V$ and $W$ are given. Specifically, we consider how to solve the following minimization problem efficiently

$$(\text{PRGSR-CPT-S}) \quad \inf_{x \in \mathbb{R}} x \quad \text{s.t.} \quad \sup_{(v, w^-, w^+) \in \mathcal{V}_{EI} \times \mathcal{W}_r \times \mathcal{W}_r} \mathbb{E}_{w^-w^+}[v(\xi - x)] \leq 0,$$

(4.1)

where $\mathcal{V}_{EI}$ and $\mathcal{W}_r$ are defined in (3.20) and (3.23) respectively. We use (PRGSR-CPT-S) to indicate that this is a (PRGSR-CPT) when the ambiguity sets take a specific structure. Throughout this section, we confine our discussions to the case that the random variable $\xi$ is finitely discretely distributed. In the case that the distribution is continuous, we can use the sample average approximation method to discretize, see, e.g. Guo and Xu [19]. Problem (4.1) is mathematical program with semi-infinite constraints. To tackle the problem, we propose to derive a dual formulation of the constraint function. To this end, we need to introduce the following notation.

Let $\Xi := \{\xi_1, \ldots, \xi_N\} \subset \mathbb{R}$ denote the support set of $\xi$ with $P(\xi = \xi_i) = p_i$ for $i = 1, \ldots, N$. By sorting out the elements in $\Xi$ in an increasing order ($\xi_{j_1} \leq \xi_{j_2}$, $j_1 < j_2$), we obtain $\xi_{j_1} - x \leq \xi_{j_2} - x$ for $j_1 < j_2$. By convention, for the outcomes $\xi_j - x < 0$, we label them with indexes $j \in \{-m, -m + 1, \ldots, -1\}$ and for those $\xi_j - x \geq 0$, we label them by $j \in \{0, \ldots, n\}$. Consequently we can rewrite the integral $\mathbb{E}_{w^-w^+}[v(\xi - x)] = \int_{-\infty}^{0} v^-(t)dw^-(F_{\xi-x}(t)) + \int_{0}^{\infty} v^+(t)d(1-w^+(1-F_{\xi-x}(t)))$ as

$$\mathbb{E}_{w^-w^+}[v(\xi - x)] = \sum_{i=-m}^{n} \pi_i v(\xi_i - x),$$

(4.2)
where \( v(\xi - x) = v^-(\xi - x) \) for \( i = -m, \ldots, -1 \), and \( v(\xi - x) = v^+(\xi - x) \) for \( i = 0, \ldots, n \), and

\[
\pi_i := \begin{cases} 
  w^-(p_{-m}) & \text{for } i = -m \\
  w^-(\sum_{j=-m}^{i} p_j) - w^-(\sum_{j=-m}^{i-1} p_j) & \text{for } i = -m + 1, \ldots, -1 \\
  w^+(\sum_{j=i}^{n} p_j) - w^+(\sum_{j=i+1}^{n} p_j) & \text{for } i = 0, 1, \ldots, n - 1 \\
  w^+(p_n) & \text{for } i = n.
\end{cases} \tag{4.3}
\]

In the case that \( w^+(p) = 1 - w^-(1 - p) \), \( \forall p \in [0, 1] \), we have \( \sum_{i=-m}^{n} \pi_i = 1 \), and for \( i = 0, \ldots, n - 1 \),

\[
\pi_i = 1 - w^-\left(1 - \sum_{j=i}^{n} p_j\right) - \left(1 - w^-\left(1 - \sum_{j=i+1}^{n} p_j\right)\right) = w^-\left(\sum_{j=-m}^{i} p_j\right) - w^-\left(\sum_{j=-m}^{i-1} p_j\right).
\]

Consequently, \( \pi_i \) can be expressed \( \pi_i = w^-\left(\sum_{j=-m}^{i} p_j\right) - w^-\left(\sum_{j=-m}^{i-1} p_j\right) \), for \( i = -m + 1, \ldots, n \).

### 4.1 Reformulation of (PRGSR-CPT-S)

With the notation introduced above, we can recast (PRGSR-CPT-S) as (PRGSR-CPT-S-Dis) (where “Dis” is used to indicate the case that \( \xi \) is discretely distributed):

\[
\begin{align*}
\text{(PRGSR-CPT-S-Dis)} & \quad \rho_{V_{EI} \times W_r}(\xi) = \inf_{x \in \mathbb{R}} x \quad \text{s.t.} \quad -\sup_{(v,w^-,w^+) \in V_{EI} \times W_r \times W_r} \sum_{i=-m}^{n} \pi_i v(x - \xi_i) \leq 0. \tag{4.4}
\end{align*}
\]

In this section, we confine our study to the case that value function is defined over \([\alpha, \beta]\) and take fixed values at points \( \alpha, \beta \) with \( \alpha < 0 < \beta \). The next theorem states that the constraint in program (PRGSR-CPT-S-Dis) can be computed by solving a finite dimensional biconvex program of reasonable size.

**Theorem 4.1** For any fixed \( x \), the constraint function

\[
-\sup_{(v,w^-,w^+) \in V_{EI} \times W_r \times W_r} \sum_{i=-m}^{n} \pi_i v(\xi_i - x)
\]
in (4.4) is the optimal value of the following biconvex program:

\[
\sup_{a,b,\psi^-\psi^+,\eta^-\eta^+,\theta^+} \left\{ \sum_{i=-m}^{1} \left( (a_1(\xi_i - x) + b_1) \mathbb{I}_{(-\infty,0] \cap [y_1,y_2]}(\xi_i - x) \\
+ \sum_{j=2}^{J-1} (a_j(\xi_i - x) + b_j) \mathbb{I}_{(-\infty,0] \cap (y_j,y_{j+1}]}(\xi_i - x) \right) \sum_{t=1}^{T} \psi^-_t \int_{q_{l-1}}^{q_{l}} \mathbb{I}_{[t_t_{l+1}]}(t)dt \\
+ \sum_{i=0}^{n} (a_1(\xi_i - x) + b_1) \mathbb{I}_{(-\infty,0] \cap [y_1,y_2]}(\xi_i - x) \\
+ \sum_{j=2}^{J-1} (a_j(\xi_i - x) + b_j) \mathbb{I}_{(-\infty,0] \cap (y_j,y_{j+1}]}(\xi_i - x) \right) \sum_{t=1}^{T} \psi^+_t \int_{1-q_{l}}^{1-q_{l-1}} \mathbb{I}_{[t_{l+1}]}(t)dt \right\}
\]

\text{s.t.}
\begin{align*}
\psi^-_{l+1} & \leq \psi^-_l, l = 1, \cdots, l_p - 2, \quad (4.5a) \\
\psi^+_{l+1} & \leq \psi^+_l, l = 1, \cdots, l_p - 2, \quad (4.5b) \\
\psi^-_l & \leq \psi^-_{l+1}, l = l_p, \cdots, T - 1, \quad (4.5c) \\
\psi^+_l & \leq \psi^+_{l+1}, l = l_p, \cdots, T - 1, \quad (4.5d) \\
\sum_{t=1}^{T} \psi^-_t (t_{l+1} - t_l) & = 1, \quad (4.5e) \\
\sum_{t=1}^{T} \psi^+_t (t_{l+1} - t_l) & = 1, \quad (4.5f) \\
\sum_{t=1}^{T} (\eta^-_t + \theta^-_t) \int_{t_l}^{t_{l+1}} \tilde{g}(t)dt & \leq r, \quad (4.5g) \\
\sum_{t=1}^{T} (\eta^+_t + \theta^+_t) \int_{t_l}^{t_{l+1}} \tilde{g}(t)dt & \leq r, \quad (4.5h) \\
\psi^-_l - \psi^-_0 + \eta^-_l - \theta^-_l & = 0, l = 1, \cdots, T, \quad (4.5i) \\
\psi^+_l - \psi^+_0 + \eta^+_l - \theta^+_l & = 0, l = 1, \cdots, T, \quad (4.5j) \\
\psi^-_l > 0, \psi^+_l > 0, \eta^-_l, \theta^-_l, \eta^+_l, \theta^+_l & \geq 0, l = 1, \cdots, T, \quad (4.5k) \\
\sum_{j=1}^{J-1} a_j \int_{y_j}^{y_{j+1}} \phi_m(t)dt & \leq 0, m = 1, \cdots, M, \quad (4.5l) \\
a_{j-1}y_j + b_{j-1} & = a_j y_j + b_j, j = 2, \cdots, J - 1, \quad (4.5m) \\
a_j & \leq a_{j+1}, j = 1, \cdots, j_0 - 2, \quad (4.5n) \\
a_{j+1} & \leq a_j, j = j_0, \cdots, J - 2, \quad (4.5o) \\
a_{j_0} + b_{j_0} & = 0, a_1 y_1 + b_1 = b_t, a_{J-1} y_J + b_{J-1} = b_r, \quad (4.5p) \\
c^k & \leq 0, \ c^k & \geq 0, k = 1, \cdots, K, \quad (4.5q) \\
a_j & > 0, j = 1, \cdots, J, \quad (4.5r)
\end{align*}
where \( j_0 := \{ j \in \{1, \cdots, J \} : y_{j_0} = 0 \} \),

\[
\hat{c}^k = \left( (a_1(A_k^{1} - a_k^{+}) + b_1)\mathbb{1}_{(-\infty,0) \cap [y_1, y_2]}(A_k^{1} - a_k^{+}) \right.
\]

\[
+ \sum_{j=2}^{J-1} (a_j(A_k^{1} - a_k^{+}) + b_j)\mathbb{1}_{(-\infty,0) \cap (y_j, y_{j+1})}(A_k^{1} - a_k^{+}) \right) \left( w_+^*(q_k^1) - w_+^*(q_k^0) \right)
\]

\[
\left. + \left( (a_1(A_k^{2} - a_k^{+}) + b_1)\mathbb{1}_{(0,\infty) \cap [y_1, y_2]}(A_k^{2} - a_k^{+}) \right) \left( w_+^*(1-q_k^1) - w_+^*(1-q_k^2) \right) \right). \]

\[
\hat{c}^k = \left( (a_1(A_k^{1} - a_k^{−}) + b_1)\mathbb{1}_{(-\infty,0) \cap [y_1, y_2]}(A_k^{1} - a_k^{−}) \right.
\]

\[
+ \sum_{j=2}^{J-1} (a_j(A_k^{1} - a_k^{−}) + b_j)\mathbb{1}_{(-\infty,0) \cap (y_j, y_{j+1})}(A_k^{1} - a_k^{−}) \right) \left( w_−^*(q_k^1) - w_−^*(q_k^0) \right)
\]

\[
\left. + \left( (a_1(A_k^{2} - a_k^{−}) + b_1)\mathbb{1}_{(0,\infty) \cap [y_1, y_2]}(A_k^{2} - a_k^{−}) \right) \left( w_−^*(1-q_k^1) - w_−^*(1-q_k^2) \right) \right). \]

\( q_{m-1} := 0, q_i := \sum_{j=-m}^{i} p_j \) for \( i = -m, \cdots, n \), \( q_0^0 := 0, q_1^1 := 1 - p_k, q_2^2 := 1, A_1^1 := 0, A_1^2 := r_k \) (\( p_k \) and \( r_k \) are defined as in Section 5.3, Q2) and \( \{y_j\}_{j=1}^J \) is the \( j \)-th entry of ranked support set

\[
\mathcal{S} := \bigcup_{m=1}^{M} (\text{supp}(\gamma_m) \cup \text{supp}(\eta_m)) \bigcup \{\alpha, \beta\} \bigcup \{0\}
\]

\[
\bigcup_{i=1}^{N} (\xi_i - x) \bigcup_{k=1}^{K} \text{supp}(A_k - a_k^+) \bigcup_{k=1}^{K} \text{supp}(A_k - a_k^-), \tag{4.6}
\]

and \( J := |\mathcal{S}| = 3M + 3 + N + 4K. \phi_m \in \text{condition (4.5l)} \) is defined as in Section 5.3, Q1. Condition (4.5l) characterizes the DM’s preference between variables \( \gamma_m \) and \( \eta_m \), \( m = 1, \cdots, M \), conditions (4.5m), (4.5n) and (4.5o) represents the increasing of value function \( v \) and convexity of \( v^- \) and concavity of \( v^+ \), conditions (4.5p) represents three fixed values of \( v \).

The proof is a bit long, we defer it to Appendix A.2.

In general, it is difficult to solve the biconvex program (4.5). Haskell et al. [20] consider a preference robust optimization model where the ambiguity lies not only in the DM’ risk preferences but also in the probability distribution of exogenous uncertainties. The resulting minmax optimization model is a bilinear programming problem. The authors show how to solve their problem directly in a special case when the ambiguity set of the probability distributions can be represented by the mixtures of some given distributions. However, the bilinear programming problem is generally difficult to solve. They propose two techniques to address the difficulty: a reformulation-linearization technique (RLT) and an semidefinite programming (SDP) relaxation approach. Both are approximation schemes and may introduce many variables if we use them to deal with the biconvex structure in (4.5) directly. This is undesirable particularly when \( N \) is large (\( \xi \) has many scenarios).
Here, we take a different approach. We can reformulate program (4.5) as a new linear program by changing some variables as follows:

\[
\begin{align*}
\tilde{a}_{l,j}^- &:= a_j \psi_i^-, \tilde{a}_{l,j}^+ := a_j \psi_i^+ , \tilde{b}_{l,j}^- &:= b_j \psi_i^- , \tilde{b}_{l,j}^+ &:= b_j \psi_i^+ , \text{for } j = 1, \ldots , J , \ l = 1, \ldots , T . \\
\end{align*}
\]

that is,

\[
\begin{align*}
\sup_{\tilde{a}^- , \tilde{b}^- , \tilde{a}^+ , \tilde{b}^+ , \psi^- , \psi^+, \eta^- , \eta^+, \theta^- , \theta^+} & \left\{ \sum_{i=-m}^{-1} \sum_{l=1}^{T} \left( (\tilde{a}_{l,j}^- (\xi_i - x) + \tilde{b}_{l,j}^-) \mathbb{I}_{(-\infty ,0) \cap [y_i,y_{i+1}]} (\xi_i - x) \right) \\
+ \sum_{j=2}^{J-1} (\tilde{a}_{j}^- (\xi_i - x) + \tilde{b}_{j}^-) \mathbb{I}_{(-\infty ,0) \cap [y_j,y_{j+1}]} (\xi_i - x) \right) \int_{q_{i-1}}^{q_i} \mathbb{I}_{[t_i,t_{i+1}]} (t) dt \\
+ \sum_{i=0}^{n} \sum_{l=1}^{T} \left( (\tilde{a}_{l,j}^+ (\xi_i - x) + \tilde{b}_{l,j}^+) \mathbb{I}_{(0,\infty) \cap [y_i,y_{i+1}]} (\xi_i - x) \right) \int_{1-q_{i-1}}^{1-q_i} \mathbb{I}_{[t_i,t_{i+1}]} (t) dt \right\} \\
\text{s.t.} & \ (4.5a) - (4.5k) , \\
\sum_{j=1}^{J-1} \tilde{a}_{j}^- \int_{y_{j-1}}^{y_j} \phi_m(t) dt \leq 0 , \ m = 1 , \ldots , M , \ l = 1 , \ldots , T , \quad (4.7a) \\
\sum_{j=1}^{J-1} \tilde{a}_{j}^+ \int_{y_{j-1}}^{y_j} \phi_m(t) dt \leq 0 , \ m = 1 , \ldots , M , \ l = 1 , \ldots , T , \quad (4.7b) \\
\tilde{a}_{j-1}^- y_j + \tilde{b}_{j-1}^- = \tilde{a}_{j}^- y_j + \tilde{b}_{j}^- , \ j = 2 , \ldots , J , \ l = 1 , \ldots , T , \quad (4.7c) \\
\tilde{a}_{j-1}^+ y_j + \tilde{b}_{j-1}^+ = \tilde{a}_{j}^+ y_j + \tilde{b}_{j}^+ , \ j = 2 , \ldots , J , \ l = 1 , \ldots , T , \quad (4.7d) \\
\tilde{a}_{j}^- \leq \tilde{a}_{j+1}^- , \ j = 1 , \ldots , j_0 - 1 , \ l = 1 , \ldots , T , \quad (4.7e) \\
\tilde{a}_{j}^+ \leq \tilde{a}_{j+1}^+ , \ j = 1 , \ldots , j_0 - 1 , \ l = 1 , \ldots , T , \quad (4.7f) \\
\tilde{a}_{j+1}^- \leq \tilde{a}_{j}^- , \ j = j_0 , \ldots , J - 2 , \ l = 1 , \ldots , T , \quad (4.7g) \\
\tilde{a}_{j+1}^+ \leq \tilde{a}_{j}^+ , \ j = j_0 , \ldots , J - 2 , \ l = 1 , \ldots , T , \quad (4.7h) \\
\tilde{a}_{j_0}^- 0 + \tilde{b}_{j_0}^- = 0 , \ l = 1 , \ldots , T , \quad (4.7i) \\
\tilde{a}_{j_0}^+ 0 + \tilde{b}_{j_0}^+ = 0 , \ l = 1 , \ldots , T , \quad (4.7j) \\
\tilde{a}_{1}^- y_1 + \tilde{b}_{1}^- = \psi_i^- b_l , \ l = 1 , \ldots , T , \quad (4.7k) \\
\tilde{a}_{1}^+ y_1 + \tilde{b}_{1}^+ = \psi_i^+ b_l , \ l = 1 , \ldots , T , \quad (4.7l) \\
\tilde{a}_{j-1}^- y_j + \tilde{b}_{j-1}^- = \psi_i^- b_l , \ l = 1 , \ldots , T , \quad (4.7m) \\
\tilde{a}_{j-1}^+ y_j + \tilde{b}_{j-1}^+ = \psi_i^+ b_l , \ l = 1 , \ldots , T , \quad (4.7n) \\
\tilde{c}_{k}^- \geq 0 , \quad \tilde{c}_{k}^- \leq 0 , \ k = 1 , \ldots , K , \quad (4.7o) \\
\tilde{c}_{k}^+ \geq 0 , \quad \tilde{c}_{k}^+ \leq 0 , \ k = 1 , \ldots , K , \quad (4.7p) \\
\tilde{a}_{j}^- > 0 , \tilde{a}_{j}^+ > 0 , \ j = 1 , \ldots , J , \ l = 1 , \ldots , T , \quad (4.7q)
\end{align*}
\]

where \( q_{m+1} = 0 , q_i = \sum_{j=-m}^{n} p_j \) for \( i = -m , \ldots , n , \tilde{a}^- := (\tilde{a}_{-i}^-)_{i \times T} , \tilde{b}^- := (\tilde{b}_{-i}^-)_{i \times T} , \tilde{a}^+ := (\tilde{a}_{i}^+)_{i \times T} , \tilde{b}^+ := (\tilde{b}_{i}^+)_{i \times T} , \psi^- := (\psi_i^- , \ldots , \psi_T^-) , \psi^+ := (\psi_i^+ , \ldots , \psi_T^+) , \psi^0 := (\psi_i^0 , \ldots , \psi_T^0) , q_k^0 = 0 , q_k^1 = 1 - p_k , \)
\[ q_k^2 = 1, \quad A_k^1 = 0, \quad A_k^2 = r_k, \]
\[
\tilde{c}_k^- = \left( (\tilde{a}_1^-(A_k^1 - a_k^+) + \tilde{b}_1^-) 1_{(-\infty, 0) \cap [y, y_2]}(A_k^1 - a_k^+) \right. \\
+ \sum_{j=2}^{J-1} (\tilde{a}_j^- (A_k^1 - a_k^+) + \tilde{b}_j^-) 1_{(-\infty, 0) \cap (y_j, y_{j+1})} (A_k^1 - a_k^+) \left. \right) \left( w^- (q_k^1) - w^- (q_k^0) \right) \\
+ \left( (\tilde{a}_1^-(A_k^2 - a_k^+) + \tilde{b}_1^-) 1_{(0, \infty) \cap [y_1, y_2]} (A_k^2 - a_k^+) \right. \\
+ \sum_{j=2}^{J-1} (\tilde{a}_j^- (A_k^2 - a_k^+) + \tilde{b}_j^-) 1_{(0, \infty) \cap (y_j, y_{j+1})} (A_k^2 - a_k^+) \left. \right) \left( w^+ (1 - q_k^1) - w^+ (1 - q_k^2) \right),
\]
\[
\tilde{c}_k^- = \left( (\tilde{a}_1^-(A_k^1 - a_k^-) + \tilde{b}_1^-) 1_{(-\infty, 0) \cap [y_1, y_2]} (A_k^1 - a_k^-) \right. \\
+ \sum_{j=2}^{J-1} (\tilde{a}_j^- (A_k^1 - a_k^-) + \tilde{b}_j^-) 1_{(-\infty, 0) \cap (y_j, y_{j+1})} (A_k^1 - a_k^-) \left. \right) \left( w^- (q_k^1) - w^- (q_k^0) \right) \\
+ \left( (\tilde{a}_1^-(A_k^2 - a_k^-) + \tilde{b}_1^+) 1_{(0, \infty) \cap [y_1, y_2]} (A_k^2 - a_k^-) \right. \\
+ \sum_{j=2}^{J-1} (\tilde{a}_j^- (A_k^2 - a_k^-) + \tilde{b}_j^+) 1_{(0, \infty) \cap (y_j, y_{j+1})} (A_k^2 - a_k^-) \left. \right) \left( w^+ (1 - q_k^1) - w^+ (1 - q_k^2) \right),
\]

and \( \tilde{c}_k^+ \) in (4.7p) denotes the counterpart of \( \tilde{c}_k^- \) with \( \tilde{a}_j^- \) and \( \tilde{b}_j^- \) being replaced by \( \tilde{a}_j^+ \) and \( \tilde{b}_j^+ \), \( \tilde{c}_k^+ \) in (4.7p) denotes the counterpart of \( \tilde{c}_k^- \) with \( \tilde{a}_j^- \) and \( \tilde{b}_j^- \) being replaced by \( \tilde{a}_j^+ \) and \( \tilde{b}_j^+ \).

Moreover, for the optimal solution of program (4.5), we can solve (4.7) to obtain the optimal solution
\[
\tilde{a}_j^{sl}, \tilde{a}_j^{sl+}, \tilde{b}_j^{sl}, \tilde{b}_j^{sl+}, \psi_1^s+, \psi_1^s-, \psi_2^s+, \psi_2^s-, \theta^s-, \theta^s+, \theta^s+ + 1, \ldots, J, l = 1, \ldots, T,
\]
then we can get the optimal solution of (4.5)
\[
a_j^* = \frac{\tilde{a}_j^{sl}}{\psi_j^s}, \quad b_j^* = \frac{\tilde{b}_j^{sl}}{\psi_j^s}, \quad \psi_1^s+, \psi_2^s+, \theta^s-, \theta^s+, \theta^s+, j = 1, \ldots, J, \text{ for some } l \in \{1, \ldots, T\}.
\]

This method will be used in Section 5 for numerical tests.

From Proposition 2.1, we claim that the proposed PRGSR-CPT (4.4) can be solved by using the following two-stage procedures:

**Algorithm 4.1 (Two-stage procedure for solving the proposed PRGSR-CPT (4.4))**
Step 1. For each fixed $x$, solve linear program:

$$h(x) := \sup_{\hat{a} - \hat{b} \in \mathbb{R}^n, \psi, \psi^+} \left\{ \sum_{i=0}^{n-1} \sum_{l=1}^{m} \left( (\hat{a}_i^l - (\xi_i - x) + \hat{b}_i^l) \mathbb{I}_{(-\infty,0] \cap (y_i,y_{i+1}]}(\xi_i - x) \\ + \sum_{j=2}^{J-1} (\hat{a}_j - (\xi_i - x) + \hat{b}_j^l) \mathbb{I}_{(-\infty,0] \cap (y_j,y_{j+1}]}(\xi_i - x) \right) \right\}$$

\[ \begin{aligned}
&+ \sum_{i=0}^{n} \sum_{l=1}^{m} \left( (\hat{a}_i^l - (\xi_i - x) + \hat{b}_i^l) \mathbb{I}_{[0,\infty] \cap (y_i,y_{i+1}]}(\xi_i - x) \\ &+ \sum_{j=2}^{J-1} (\hat{a}_j - (\xi_i - x) + \hat{b}_j^l) \mathbb{I}_{[0,\infty] \cap (y_j,y_{j+1}]}(\xi_i - x) \right) \int_{q_{i-1}}^{q_i} \mathbb{I}_{[t_i,t_{i+1}]}(t) dt \\
&+ \sum_{i=0}^{n} \sum_{l=1}^{m} \left( (\hat{a}_i^l - (\xi_i - x) + \hat{b}_i^l) \mathbb{I}_{[0,\infty] \cap (y_i,y_{i+1}]}(\xi_i - x) \right) \int_{q_{i-1}}^{1} \mathbb{I}_{[t_i,t_{i+1}]}(t) dt \\
\end{aligned} \]  

s.t. \quad (4.5a) - (4.5k), (4.7a) - (4.7q),  

(4.8)

Step 2. Solve the following minimization problem:

$$\inf_{x \in \mathbb{R}} \{ x : h(x) \leq 0 \}.$$  

(4.9)

Note that when the random variable $\xi$ is essentially bounded, we may solve problem (4.9) by bisection method. Consider, for instance, that $\Xi = [0, 0.5]$. We begin with an initial feasible solution $x_0$ and set $x_1 := \frac{x_0}{2}$. If $x_1$ is feasible, then update the feasible set by $[0, x_1]$, otherwise set $[x_1, x_0]$, repeat the process. We will come back to this in Section 5.

5 Numerical tests

To examine the performance of the PRGSR-CPT model and the computational schemes, we carry out some numerical tests on problem (4.4) with an academic example where the true value function and weighting functions are known and how elicited preference information may affect the convergence of the worst case value function and weighting functions as more and more information is available. Since this is artificially made example, we use the well known random utility split approach [3] to generate pairwise comparison questionnaires (prospects) and then use the GSR-CPT based on the tuple of true value function and weighting functions to select a preferred prospect (act like the DM). The preference information is then used to construct the ambiguity set of the value functions.

All of the numerical experiments are carried out with Matlab 2020b installed on a PC (16GB RAM, CPU 2.3 GHz) with Intel Core i7 processor. In this section, we report the details of the test and the outcomes.

5.1 Setup

Let $\xi$ be a random variable which is uniformly distributed over a discrete set $\Xi = \{ \xi_1, \cdots, \xi_N \}$ with $N = 10$. We sort out $\{ \xi_i \}_{i=1}^{N}$ in non-decreasing order, $\xi_1 < \xi_2 < \cdots < \xi_N$. Specifically let

$$\Xi = \{ \xi_1, \cdots, \xi_{10} \} = \{ 0.4074, 0.4529, 0.0635, 0.4567, 0.3162, 0.0488, 0.1392, 0.2734, 0.4788, 0.4824 \}.$$  

(5.1)
We consider a specific situation where the true value function is

\[ v_r(x) = \begin{cases} v_r^+(x) = x^{1/3} & \text{for } x \geq 0, \\ v_r^-(x) = -1.5(-x)^{0.2} & \text{for } x < 0, \end{cases} \]

and the true weighting functions \( w_r^- (p) = w_r^+(p) \) are piecewise functions with \( T \) breakpoints of \( w(p) \):

\[ w(p) = \frac{p^\gamma}{p^\gamma + (1 - p)^\gamma}, \quad \gamma = 0.6. \]

It is easy to verify that \( 1 - w(1 - p) = w(p) \). Then we consider the case that \( \mathbb{E}_{w^- w^+}[v(\xi)] = \int_0^{0.5} v(t) dw^-(F_\xi(t)) \) and \( 1 - w(1 - p) = w(p) \). We consider problem (PRGSR-CPT-S-Dis) defined in (4.4):

\[ \rho_{\mathcal{V}_El \times \mathcal{W}_r}(\xi) = \inf \left\{ x : \sup_{(v,w^-) \in \mathcal{V}_{El} \times \mathcal{W}_r} \sum_{i=-m}^{n} \pi_i v(\xi_i - x) \leq 0 \right\}, \tag{5.2} \]

where \( v(x - \xi_i) = v^- (x - \xi_i) \) for \( i = -m, \cdots, -1 \), \( v(x - \xi_i) = v^+ (x - \xi_i) \) for \( i = 0, \cdots, n \),

\[ \pi_i := \begin{cases} w^- \left( \frac{i + m + 1}{10} \right) - w^- \left( \frac{i + m}{10} \right) & \text{for } i = -m, \cdots, -1, \\ w^+ \left( \frac{n - i + 1}{10} \right) - w^+ \left( \frac{n - i}{10} \right) & \text{for } i = 0, 1, \cdots, n, \end{cases} \]

and \( m + n + 1 = 10 \), \( \mathcal{V}_{El} = \mathcal{V}_{\text{ke}}^K \cap \mathcal{V}_{\text{pair}}^M \cap \mathcal{V}_S \) and \( \mathcal{W}_r = B(w_0, r) \). In this experiment, we assume the shape of the true value function is known, that is, it is S-shaped with \( v(0) = 0 \). Based on this information and the concrete positive values of \( \xi_i \) in the support set \( \Xi \), we deduce that \( \rho_{(v,w^-,w^+)}(\xi) \in [0,0.5] \) for any value function and weighting functions. This explains why we write the feasible set of \( x \) as \([0,0.5]\) instead of \( \mathbb{R} \) in problem (5.2).

In the next subsections, we will describe in detail how the ambiguity sets \( \mathcal{W}_r \) and \( \mathcal{V}_{El} \) are constructed.

### 5.2 Construction of \( \mathcal{W}_r \)

As discussed earlier, we use a ball centered at a nominal weighting function to define the ambiguity set \( \mathcal{W}_r \) under the pseudo-metric, see Definition 3.3. We take a specific \( \tilde{g}(t) = 1 \) in the test, i.e., \( \mathcal{W}_r := B_{L_1}(w_0, r) \), see (3.24). It follows by Theorem 4.1 that

\[ \sup_{(v,w^-) \in \mathcal{V}_{El} \times \mathcal{W}_r} \sum_{i=-m}^{n} \pi_i v(\xi_i - x) = \sup_{(v,w^-) \in \mathcal{V}_{El} \times \tilde{B}_{L_1}(w_0,r)} \sum_{i=-m}^{n} \pi_i v(\xi_i - x), \]

where

\[ \tilde{B}_{L_1}(w_0, r) = \left\{ w \in \mathcal{W}_T : w \text{ is increasing and inverse-S shaped} \right\}, \tag{5.3} \]

where \( \mathcal{W}_T := \mathcal{W} \), \( \tilde{B}_{L_1}(w_0, r) := \{ w \in \mathcal{W}_T : w(0) = 0, w(1) = 1, d\mathcal{W}_1(w, w_0) \leq r \} \), and the set of derivatives of \( w \in \tilde{B}_{L_1}(w_0, r) \),

\[ \tilde{\Psi}_{L_1}(\psi^0, r) := \left\{ \psi \in \Psi : w \in \tilde{B}_{L_1}(w_0, r) \right\} = \left\{ \psi \in \Psi_{T} : \sum_{l=1}^{T} \left| \psi_l - \psi^0_l \right| (t_{l+1} - t_l) \leq r \right\}, \tag{5.4} \]

where \( \Psi \) is the set of derivative functions of the weighting functions in \( \mathcal{W}_{\text{inS}} \).
5.3 Construction of $V_{EI}$

We randomly generate the certainty equivalent questionnaires and pairwise comparison questionnaires. These questions may be phrased as follows.

- Q1. For each pair of random prospects $(\gamma_m, \eta_m)$, $m = 1, \cdots, M$, which prospect does the DM prefer? As we discussed earlier in Section 3, we represent by $E_{w^{-w^{+}}}[v(\gamma_m)] \geq E_{w^{-w^{+}}}[v(\eta_m)]$ if random prospect $\gamma_m$ is preferred to random prospect $\eta_m$. Note that in this case, the DM's true weighting functions $w^{-}, w^{+}$ are known.

- Q2. What’s the smallest amount of cash, denoted by $a^+_k$, that the DM would decline to commit instead of being exposed to the random prospect $A_k$ and what is the largest amount of cash, denoted by $a^-_k$, that DM would be willing to commit instead of being exposed to the random prospect $A_k$ for $k = 1, \cdots, K$.

It remains to explain how the random prospects are generated. For Q1, we use the basic idea of the utility split approach (see e.g. [3, 11]) to generate pairs of random prospects $(\gamma_m, \eta_m)$, $m = 1, \cdots, M$. First, we explain how to generate random prospects. Let

$$Z_1 = \begin{cases} r_1 & \text{with probability } 1 - p, \\ r_3 & \text{with probability } p, \end{cases} \text{ and } Z_2 = r_2,$$

where $r_1$ and $r_3$ are generated randomly over $[-0.5, 0.5]$ with $r_1 < r_3$, let $r_2 = (r_1 + r_3)/2$. Next, we need to set an appropriate value for $p$. Observe that

$$E_{w^{-}w^{+}}[v(Z_1)] = (1 - w^{+}(p))v(r_1) + w^{+}(p)v(r_3) \quad \text{and} \quad E_{w^{-}w^{+}}[v(Z_2)] = v\left(\frac{r_1 + r_3}{2}\right).$$

Note that here the true weighting functions are known whereas the value function $v$ is unknown. Consider the case

$$(1 - w^{+}(p))v(r_1) + w^{+}(p)v(r_3) < v\left(\frac{r_1 + r_3}{2}\right). \quad (5.5)$$

Let $\bar{v}(x) = \frac{v(x) - v(r_1)}{v(r_3) - v(r_1)}$ for $x \in [r_1, r_3]$. Then $\bar{v}(r_1) = 0, \bar{v}(r_3) = 1$ and consequently (5.5) can be written as

$$w^{+}(p) < \bar{v}\left(\frac{r_1 + r_3}{2}\right). \quad (5.6)$$

In other words, the characterization of the value function $v$ is down to the characterization of the normalized value function $\bar{v}$ at point $r_2$. Let $V_{pair}^k$ denote the set of value functions which are consistent to pairwise comparison questionnaires that we have already generated. Define

$$I_1 := \min_{v \in V_{pair}^k \cap V_{S}} \bar{v}(r_2) \quad \text{and} \quad I_2 := \max_{v \in V_{pair}^k \cap V_{S}} \bar{v}(r_2). \quad (5.7)$$

Since $\bar{v}(r_2) \in [0, 1]$, then $I_1, I_2 \in [0, 1]$. We set the value of $p$ such that

$$w^{+}(p) = \frac{I_1 + I_2}{2}, \quad (5.8)$$

25
and then use the true value function \( v_\ast \) (acting as the DM) to check whether \( \mathbb{E}_{w_\ast^-} [\tilde{v}_\ast(Z_1)] < \mathbb{E}_{w_\ast^-} [\tilde{v}_\ast(Z_2)] \) or not, that is, \( w_\ast^+(p) < \tilde{v}_\ast(r_2) \) or not. Let

\[
\phi_{k+1}(y) = -\left( w_\ast^- \left( (1 - p_{k+1}) I_{y \geq r_1} + p_{k+1} I_{y < r_3} \right) - w_\ast^- \left( I_{y \geq r_2} \right) \right)
\]

\[
= -I_{y \geq r_3} + I_{y \geq r_2} - (1 - w_\ast^+(p_{k+1})) I_{y \geq r_1, y < r_3},
\]

and \( S \) be the set of all \( r_1, r_2, r_3 \) generated from \( M \) questionnaires and points \(-0.5, 0.5, 0, \) and \( \tilde{M} = 3M + 3 \). The following algorithm describes the procedures for constructing \( \psi_{M}^{\ast} \cap V_{S} \).

**Algorithm 5.1** Initialization: set \( k := 0 \) and the set of questionnaires is empty.

**Step 1.** Choose \( r_1 \) and \( r_3 \) randomly from \([-0.5, 0.5]\) with equal probability with \( r_1 < r_3 \) and then set \( r_2 := \frac{r_1 + r_3}{2} \).

**Step 2.** Let \( j_0 := \max \{ j \in \{1, \ldots, \tilde{M} \} : t_j = 0 \} \) and solve the following problem:

\[
\min_{a, b} \left( (a_1 r_2 + b_1) I_{[t_1, t_2]}(r_2) + \sum_{j=2}^{\tilde{M}-1} (a_j r_2 + b_j) I_{[t_j, t_{j+1}]}(r_2) + 1.5 \cdot (0.5)^{0.2} \right) / ((0.5)^{1/3} + 1.5 \cdot (0.5)^{0.2})
\]

\[
(5.9a)
\]

s.t. \( a_{j-1}y_j + b_{j-1} = a_jy_j + b_j, \forall j = 2, \ldots, \tilde{M} - 1, \)

\[
a_1r_1 + b_1 = -1.5 \cdot (0.5)^{0.2},
\]

\[
a_{\tilde{M}-1}r_3 + b_{\tilde{M}-1} = (0.5)^{1/3},
\]

\[
\sum_{j=1}^{\tilde{M}-1} a_j \int_{y_{j-1}}^{y_j} \phi_l(y) dy \leq 0, l = 1, \ldots, k,
\]

\[
a_j \leq a_{j+1}, j = 1, \ldots, j_0 - 1,
\]

\[
a_{j+1} \leq a_j, j = j_0, \ldots, \tilde{M} - 2,
\]

\[
a_j \geq 0, j = 1, \ldots, \tilde{M} - 1.
\]

Let \( I_1 \) be the optimal value of the program. Replace the “min” with “max” and solve the maximization problem. Let \( I_2 \) be the optimal value of the maximization problem. Let \( p_{k+1} \) be such that \( w_\ast^+(p_{k+1}) = \frac{I_1 + I_2}{2} \), and then use the true value function \( v_\ast \) to check whether

\[
\mathbb{E}_{w_\ast^-} [\tilde{v}_\ast(Z_2)] > \mathbb{E}_{w_\ast^-} [\tilde{v}_\ast(Z_1)]
\]

\[
(5.10)
\]

or not, that is, \( \tilde{v}_\ast(r_2) > w_\ast^+(p_{k+1}) \) or not. If \( \tilde{v}_\ast(r_2) > w_\ast^+(p_{k+1}) \), then \( Z_2 \) is preferred over \( Z_1 \). Let

\[
\phi_{k+1}(y) = \begin{cases} 
-1 + w_\ast^+(p_{k+1}) & \text{if } r_1 \leq y < r_2, \\
 w_\ast^+(p_{k+1}) & \text{if } r_2 \leq y < r_3, \\
 0 & \text{otherwise}.
\end{cases}
\]

\[
(5.11)
\]

Add the inequality \( \int_{0.5}^{0.5} \phi_{k+1}(y) dv(y) \leq 0 \) to the constraints of problem (5.9). Otherwise, \( \tilde{v}_\ast(r_2) \leq w_\ast^+(p_{k+1}) \), \( Z_1 \) is preferred over \( Z_2 \) and add inequality \( \int_{0.5}^{0.5} -\phi_{k+1}(y) dv(y) \leq 0 \) to the constraints of problem (5.9).
Step 3. Set \( k := k + 1 \) and return to Step 1.

For Q2, we let

\[
A_k := \begin{cases} 
     r_k & \text{with probability } p_k, \\
     0 & \text{with probability } 1 - p_k,
\end{cases}
\]

where \( r_k \) is randomly generated with uniform distribution over \([0, 0.5]\) and \( p_k \) is also randomly generated with the uniform distribution over \([0.1, 0.2, \cdots , 0.9]\). The randomness will enable us to pick up any value in the domain of the value function and with any likelihood. The upper and lower bounds \( a_k^+, a_k^- \) are set as follows:

\[
a_k^+ = (1 + \tau)\rho_{(v_*, w_-, w_+)}(A_k), \quad a_k^- = (1 - \tau)\rho_{(v_*, w_-, w_+)}(A_k),
\]

where \( \rho_{(v_*, w_-, w_+)}(A_k) \) is the true GSR-CPT value of \( A_k \), i.e.,

\[
\rho_{(v_*, w_-, w_+)}(A_k) := \inf_{x \in [0, r_k]} \left\{ x : (r_k - x)^{1/3}w_+(p_k) - 1.5(x^{0.2}(1 - w_+(p_k)) \leq 0 \right\},
\]

and \( \tau \) is randomly generated with a uniform distribution over \([0, 0.05]\).

The following algorithm describes the procedures for constructing \( V_{ce}^k \).

**Algorithm 5.2** Initialization: set \( k := 0 \) and the set of questionnaires is empty.

Step 1. Choose randomly \( r_k \in [0, 0.5] \) and \( p_k \in \{0.1, 0.2, \cdots , 0.9\} \).

Step 2. Compute \( a_k^+, a_k^- \) via formulae (5.12) and construct

\[
V_{ce}^k = \{ v \in V : \mathbb{E}_{w_-, w_+} [v(A_i - a_i^+)] \leq 0, \mathbb{E}_{w_-, w_+} [v(A_i - a_i^-)] \geq 0, i = 1, \cdots , k \}.
\]

Step 3. If problem \( \rho_{V_{pair} \cap V_{ce}^k \times W_r} (\xi) \) in (5.2) is feasible, then set \( k := k + 1 \). The loop goes to step 1.
5.4 Numerical experiments

We consider Problem (5.2) and use Algorithm 4.1 to solve it. Since $\tilde{g}(t) = 1$, we can rewrite (4.8) as

$$h(x) = \inf_{\bar{a}^{-}, \bar{b}^{-}, \bar{w}^{-}} \left\{ \sum_{i=-m}^{n} \sum_{l=1}^{T} \left( (\bar{a}^{-}_{l} - (\xi_{i} - x) + \bar{b}^{-}_{l}) 1_{[y_{l}, y_{l+1}]}(\xi_{i} - x) \right) \right. $$

$$+ \left. \sum_{j=2}^{J-1} (\bar{a}^{-}_{j}(\xi_{i} - x) + \bar{b}^{-}_{j}) 1_{(y_{j}, y_{j+1})}(\xi_{i} - x) \right] \right\}$$

s.t.

$$(4.5a), (4.5c), (4.5e),$$

$$\sum_{i=1}^{y_{j+1}} |\psi_{i}^{-} - \psi_{i}^{0}|(t_{i+1} - t_{i}) \leq r,$$

$$\psi_{i}^{-} > 0, l = 1, \ldots, T,$$

$$\sum_{j=1}^{y_{j}} (\bar{a}^{-}_{j} + \bar{b}^{-}_{j}) \int_{y_{j}}^{y_{j+1}} \phi_{m}(y)dy \leq 0, m = 1, \ldots, M,$$

$$(4.7a), (4.7c), (4.7e), (4.7g), (4.7i), (4.7k), (4.7m), (4.7o),$$

$$a_{l}^{-} > 0, a_{l}^{+} > 0, j = 1, \ldots, J, l = 1, \ldots, T,$$

$$q^{0}_{k} = 0, q^{1}_{k} = 1 - p_{k}, q^{2}_{k} = 1, A^{1}_{k} = 0, A^{2}_{k} = r_{k},$$

$$(5.13)$$

where $\bar{\psi}^{-} := (\psi_{1}^{-}, \ldots, \psi_{T}^{-}), \bar{\psi}^{0} := (\psi_{1}^{0}, \ldots, \psi_{T}^{0}), q^{0}_{k} = 0, q^{1}_{k} = 1 - p_{k}, q^{2}_{k} = 1, A^{1}_{k} = 0, A^{2}_{k} = r_{k},$

$$\tilde{c}_{k}^{-} = \sum_{i=1}^{2} \left( (\bar{a}^{-}_{i} - (A^{i}_{k} - a^{+}_{k}) + \bar{b}^{-}_{i}) 1_{[y_{l}, y_{l+1}]}(A^{i}_{k} - a^{+}_{k}) \right) \left( w^{-}_{*}(q^{i}_{k}) - w^{-}_{*}(q^{i-1}_{k}) \right),$$

$$\tilde{c}_{k}^{+} = \sum_{i=1}^{2} \left( (\bar{a}^{-}_{i} - (A^{i}_{k} - a^{-}_{k}) + \bar{b}^{-}_{i}) 1_{[y_{l}, y_{l+1}]}(A^{i}_{k} - a^{-}_{k}) \right) \left( w^{-}_{*}(q^{i}_{k}) - w^{-}_{*}(q^{i-1}_{k}) \right).$$

In this experiment, we examine how increase of the number of the pairwise comparisons $M$ affects the worst case value function. We also investigate the change of the worst case weighting function as the radius of the $\tilde{g}$-ball is driven to zero. The breakpoints of the piecewise linear weighting function are chosen with $t_{i} = 0.1i$ for $i = 1, \ldots, 10$ and $T = 9$. The cdf function of $\xi$ takes a value of $q_{k} = k/10$ at $\xi_{k}$ for $k = 1, \ldots, 10$ with $q_{0} = 0$, see (5.1) for the definition of support set of $\xi$. Note that the optimal value of the true problem is 0.2044.

**Impact of the pairwise comparisons and certainty equivalents**

Figure 2 (a) depicts changes of the worst-case value functions when the number of pairwise comparisons $M$ increases from 5 to 20, 40 and 100. The number of certainty equivalents $K$ is fixed at 5. The radius of the $\tilde{g}$-ball is also fixed at $r = 0.01$. From the figure, we can see the worst-case value
function calculated with each $M \in \{5, 20, 40, 100\}$ converges gradually to the true one $M$ increases. This is because more and more information of the DM’s preference CPT is elicited via Algorithm 5.1 and the ambiguity set $V_{pair}^M \cap V_S$ shrinks. Figure 2 (b) depicts changes of the optimal values of problem (5.2) for $M \in \{5, 20, 40, 60, 80, 100\}$.

Figure 3 (a) depicts changes of the worst-case value functions when the number of certainty equivalents $K$ increases from 5 to 20, 40 and 100. The number of pairwise comparisons $M$ is fixed at 5. The radius of the $\tilde{g}$-ball is also fixed at $r = 0.01$. From the figure, we can see the worst-case value function calculated with each $K \in \{5, 20, 40, 100\}$ converges gradually to the true one as $K$ increases. This is because more and more information of the DM’s preference CPT is elicited via Algorithm 5.1 and the ambiguity set $V_{ce}^K \cap V_S$ shrinks. Figure 2 (b) depicts changes of the optimal values of problem (5.2) for $K \in \{5, 20, 40, 60, 80, 100\}$.

![Figure 2: (a) Worst case value function vs increase of $M$. (b) Change of $\rho_{V_{EI} \times W_r}(\xi)$ in problem (5.2) vs increase of $M$.](image)

**Impact of ambiguity of the weighting functions**

Figure 4 (a) depicts changes of the worst-case weighting function w.r.t. decrease of the radius while the number of pairwise comparisons is fixed at $M = 5$ and the number of certainty equivalents is fixed at $K = 5$. From the figure, we can see the convergence of the worst-case weighting functions to the true as $r$ reduces to 0.

Figure 4 (b) displays the optimal values of PRGSR-CPT for the ambiguity sets $(V_{EI}, W_r)$. We can see that for the fixed number of pairwise comparisons, the PRGSR-CPT becomes smaller as the radius of ambiguity set decreases.

The test results are consistent with our expectation that as more and more information about the DM’s preferences is obtained, the ambiguity of the true value function and weighting function reduces, the preference robust generalized shortfall risk converges to the true GSR-CPT.
Figure 3: (a) Worst case value function vs increase of $K$. (b) Change of $\rho_{V_{EI} \times W}(\xi)$ in problem (5.2) vs increase of $K$.

Figure 4: (a) Worst case weighting function vs decrease of $r$. (b) Change of $\rho_{V_{EI} \times W}(\xi)$ in problem (5.2) vs decrease of $r$. 
6 Concluding remarks

In this paper, we explore Mao and Cai’s generalized shortfall risk measure when the value function and/or the weighting function are ambiguous. We introduce the concept of preference robust generalized shortfall risk measure which is based on a pair of the worst value function and weighting function from a joint ambiguity set value/weighting functions, investigate the properties of the robust generalized shortfall risk measure and develop computational algorithms for calculating the robust risk. The numerical tests show that the algorithms perform well w.r.t. increase of information about the DM’s risk preferences.

There are some limitations in the developed models and preference elicitation approaches. One is that the ambiguity set of the weighting functions is constructed through a ball centered at a nominal weighting function, it might be interesting to combine it with the pairwise comparison approach to reduce the ambiguity. This approach is expected to work well for each fixed value function but it seems to be challenging to elicit both the value function and the weighting function simultaneously. The other is that we are unable to use the CPT-based shortfall risk measure to represent the DM’s preference choice, rather we use the distorted expected values of prospects, that is, $E[w_u w_+ v(\cdot)]$ to describe the DM’s pairwise choice. As discussed in Section 3.2, there could be a gap between the two approaches. Note also that in the numerical tests, instances of pairwise comparisons and certainty equivalents are generated randomly without any costs. In practice, eliciting a DM’s preferences in this manner may incur financial costs, see Vayanos et al. [32]. Thus it might be interesting to explore the idea as to how to optimally design questionnaires to elicit the true value/weighting functions using approaches such as [4]. We leave all these for future research.

References

[1] C. Acerbi, Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26: 1505-1518, 2002.

[2] M. Allais, Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école américaine. *Econometrica: Journal of the Econometric Society*, 21: 503-546, 1953.

[3] B. Armbruster and E. Delage, Decision making under uncertainty when preference information is incomplete. *Management science*, 61: 111-128, 2015.

[4] D. Bertsimas and A. O’Hair, Learning preferences under noise and loss aversion: An optimization approach, *Operations Research*, 61: 1190–1199, 2013.

[5] H. Bühlmann, *Mathematical Methods in Risk Theory*. Springer, Berlin, 1970.

[6] H. Bühlmann, Premium calculation from top down. *ASTIN Bulletin: The Journal of the IAA*, 15: 89-101, 1985.

[7] J. Cai, J. Li and T. Mao, Distributionally robust optimization under distorted expectations. Available at SSRN 3566708, 2020.
[8] R.T. Clemen and T. Reilly, *Making Hard Decisions with Decision Tools Suite*. 2nd ed. Duxbury, Pacific Grove, CA, 2001.

[9] D.R. Cavagnaro, M.A. Pitt, R. Gonzalez and J.I. Myung, Discriminating among probability weighting functions using adaptive design optimization. *Journal of risk and uncertainty*, 47: 255-289, 2013.

[10] G.B. Davies and S.E. Satchell, The behavioural components of risk aversion. *Journal of Mathematical Psychology*, 51: 1-13, 2007.

[11] E. Delage, S. Guo and H. Xu, Shortfall risk models when information of loss function is incomplete, 2018.

[12] D. Denneberg, Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin: The Journal of the IAA*, 20: 181-190, 1990.

[13] D. Ellsberg, Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, 75: 643-669, 1961.

[14] D.D. Escobar and G.C. Pflug, The distortion principle for insurance pricing: properties, identification and robustness. *Annals of Operations Research*, 292: 771-794, 2018.

[15] H. Föllmer and A. Schied, Convex measures of risk and trading constraints. *Finance and stochastic*, 6: 429-447, 2002.

[16] H. Föllmer and A. Schied, *Stochastic Finance. An Introduction in Discrete Time*. 2nd ed., Berlin: De Gruyter, 2004.

[17] K. Giesecke, T. Schmidt and S. Weber, Measuring the risk of large losses. *Journal of Investment Management*, 6: 1-15, 2008.

[18] S. Guo and H. Xu, Distributionally robust shortfall risk optimization model and its approximation. *Mathematical Programming*, 174: 473-498, 2019.

[19] S. Guo and H. Xu, Utility Preference Robust Optimization with Moment-Type Information Structure. Optimization Online, 2021.

[20] B. Haskell, L. Fu and M. Dessouky, Ambiguity in risk preferences in robust stochastic optimization. *European Journal of Operational Research*, 254: 214-225, 2016.

[21] S. Heilpern, A rank-dependent generalization of zero utility principle. *Insurance: Mathematics and Economics*, 33: 67–73, 2003.

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

[23] J. Hu and G. Stepanyan, Optimization with reference-based robust preference constraints. *SIAM Journal on Optimization*, 27: 2230-2257, 2017.

[24] D. Kahneman and A. Tversky, Prospect theory: an analysis of decision under risk, *Econometrica*, 47: 263–291, 1979.
[25] M. Kaluszka and M. Krzeszowiec, Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*, 50: 159-166, 2012.

[26] U.S. Karmarkar, Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational behavior and human performance*, 21: 61-72, 1978.

[27] U.S. Karmarkar, Subjectively weighted utility and the Allais paradox. *Organizational Behavior and Human Performance*, 24: 67-72, 1979.

[28] A. Kothiyal, V. Spinu and P.P. Wakker, Prospect theory for continuous distributions: A preference foundation. *Journal of risk and uncertainty*, 42: 195, 2011.

[29] T. Mao and J. Cai, Risk measures based on behavioural economics theory. *Finance and Stochastics*, 22: 367-393, 2018.

[30] M. Nardon and P. Pianca, Insurance premium calculation under continuous cumulative prospect theory. *University Ca’Foscari of Venice, Dept. of Economics Research Paper Series No. 3*, 2019.

[31] J. von Neumann and O. Morgenstern, Theory of games and economic behavior. *Bulletin American Mathematical Society*, 51: 498-504, 1945.

[32] P. Vayanos, Y. Ye, D. McElfresh, J. Dickerson, E. Rice, Robust active preference elicitation, arXiv:2003.01899, 2020.

[33] S.M. Pesenti, Q. Wang and R. Wang, Optimizing distortion riskmetrics with distributional uncertainty. Available at SSRN, 2020.

[34] D. Prelec, The probability weighting function. *Econometrica*, 497-527, 1998.

[35] J. Quiggin, A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3: 323-343, 1982.

[36] J. Quiggin, *Generalized expected utility theory: The rank-dependent model*. Springer Science & Business Media, 2012.

[37] J.W. Tuthill and D.L. Frechette, Non-expected utility theories: Weighted expected, rank dependent, and cumulative prospect theory utility. No. 1266-2016-101906, 2002.

[38] A. Tversky and D. Kahneman, Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5: 297-323, 1992.

[39] W. Wang and H. Xu, Preference robust distortion risk measure and its application. Available at SSRN: [http://dx.doi.org/10.2139/ssrn.3763632](http://dx.doi.org/10.2139/ssrn.3763632)

[40] W. Wang and H. Xu, Robust spectral risk optimization when information on risk spectrum is incomplete. *SIAM Journal on Optimization*, 30: 3198-3229, 2020.

[41] P. Wakker and A. Tversky, An axiomatization of cumulative prospect theory. *Journal of risk and uncertainty*, 7: 147-175, 1993.
Appendix A  Supplementary materials

A.1 Properties of preference functional $E_{w^-}v(\cdot)$

Lemma A.1 Let $E_{w^-}v(\zeta)$ be defined as in (1.3). Then

$$E_{w^-}v(\eta) \leq E_{w^-}v(\xi) \forall \eta, \xi \in L^0(\mathbb{R}) \text{ with } \eta(\omega) \leq \xi(\omega), \forall \omega \in \Omega. \quad (A.1)$$

Proof. By definition (see (1.3)),

$$E_{w^-}v(\xi) = \int_0^\infty w^+(\mathbb{P}(v(\xi) > t))dt + \int_0^\infty -w^-(\mathbb{P}(-v(\xi) > t))dt.$$ 

Since $\eta(\omega) \leq \xi(\omega)$ for all $\omega \in \Omega$ and $v(\cdot)$ is monotonically increasing, then $v(\eta(\omega)) \leq v(\xi(\omega))$ for all $\omega \in \Omega$, and consequently

$$\mathbb{P}(v(\eta) > t) \leq \mathbb{P}(v(\xi) > t) \text{ and } \mathbb{P}(-v(\eta) > t) \geq \mathbb{P}(-v(\xi) > t).$$

Moreover, since $w^-, w^+$ are strictly increasing, we have

$$w^+(\mathbb{P}(v(\eta) > t)) \leq w^+(\mathbb{P}(v(\xi) > t)) \text{ and } -w^-(\mathbb{P}(-v(\eta) > t)) \leq -w^-(\mathbb{P}(-v(\xi) > t)).$$

A combination of the inequalities give rise to

$$E_{w^-}v(\eta) \leq E_{w^-}v(\xi).$$

The proof is complete. \hfill $\blacksquare$

Lemma A.2 Let $E_{w^-}v(\cdot)$ be defined as in (1.3) and $\xi \in L^0(\mathbb{R})$. Then the following assertions hold.

(i) For any $a, b \in \mathbb{R}$ with $a \leq b$,

$$E_{w^-}v(\xi - a) \geq E_{w^-}v(\xi - b). \quad (A.2)$$

(ii) For any $a, b \in \mathbb{R}$ with $a < b$,

$$E_{w^-}v(\xi - a) > E_{w^-}v(\xi - b) \forall a < b. \quad (A.3)$$
**Proof.** Part (i) follows from directly from Lemma A.1.

Part(ii). Let $\Omega_1 := \{\omega \in \Omega : \xi(\omega) \leq b\}$. For real number $a \in \mathbb{R}$, let $(a)_+ := \max\{a, 0\}$ and $(a)_- := \min\{a, 0\}$. For any $a < b$, we have

$$\xi(\omega) - b < [\xi(\omega) - a]_-, \forall \omega \in \Omega_1 \quad \text{and} \quad \xi(\omega) - b < [\xi(\omega) - a]_+, \forall \omega \in \Omega \setminus \Omega_1.$$  

Since the value function $v \in \mathcal{V}$ is strictly increasing, then

$$v^-(\xi(\omega) - b) < v^-([\xi(\omega) - a]_-), \forall \omega \in \Omega_1 \quad \text{and} \quad v^+(\xi(\omega) - b) < v^+([\xi(\omega) - a]_+), \forall \omega \in \Omega \setminus \Omega_1.$$  

Let

$$R_1 := \int_{\Omega_1} v^-((\xi(\omega) - b)\psi^-([\mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\})])d\mathbb{P}(\omega)$$

and

$$R_2 := \int_{\Omega \setminus \Omega_1} v^+(\xi(\omega) - b)\psi^+(1 - \mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\}))d\mathbb{P}(\omega).$$

By the definition of $\mathbb{E}_{(v,w^-)}[v(\cdot)]$, we have

$$\mathbb{E}_{w^-}[v(\xi - b)] = \int_{\Omega_1} v^-((\xi(\omega) - b)\psi^-([\mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\})])d\mathbb{P}(\omega)$$

$$+ \int_{\Omega \setminus \Omega_1} v^+(\xi(\omega) - b)\psi^+(1 - \mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\}))d\mathbb{P}(\omega)$$

$$= R_1 + R_2.$$  

Since we have either $\mathbb{P}(\Omega_1) > 0$ or $\mathbb{P}(\Omega \setminus \Omega_1) > 0$ or both, then

$$\mathbb{E}_{w^-}[v(\xi - b)] < \int_{\Omega_1} v^-([\xi(\omega) - a]_-)\psi^-([\mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\})])d\mathbb{P}(\omega)$$

$$+ \int_{\Omega \setminus \Omega_1} v^+([\xi(\omega) - a]_+)\psi^+(1 - \mathbb{P}(\{\hat{\omega} \in \Omega : \xi(\hat{\omega}) \leq \xi(\omega)\}))d\mathbb{P}(\omega)$$

$$= \int_{-\infty}^b v^-([u - a]_-)\psi^-([F_\xi(u)])dF_\xi(u) + \int_{b}^{\infty} v^+([u - a]_+)\psi^+(1 - F_\xi(u))dF_\xi(u)$$

$$\leq \int_{-\infty}^a v^-([u - a]_-)\psi^-([F_\xi(u)])dF_\xi(u) + \int_{a}^{\infty} v^+([u - a]_+)\psi^+(1 - F_\xi(u))dF_\xi(u)$$

$$= \mathbb{E}_{w^-}[v(\xi - a)].$$  

The proof is complete.

**A.2 Proof of Theorem 4.1**

**Proof.** We divide the proof into four steps.

**Step 1.** By definition, we have

$$\sup_{(v,w^-) \in \mathcal{V}_{\mathbb{E}} \times \mathcal{W}_r \times \mathcal{W}_v} \sum_{i=1-m}^{n} \pi_i v(x - \xi_i) = \sup_{v \in \mathcal{V}_{\mathbb{E}}} \sup_{(w^-, w^+) \in \mathcal{W}_v \times \mathcal{W}_r} \sum_{i=1-m}^{n} \pi_i v(x - \xi_i),$$

35
where $\mathcal{V}_E = \mathcal{V}^{K}_{ce} \cap \mathcal{V}^{M}_{pair} \cap \mathcal{V}_S$ and $\mathcal{W}_v = B(w_0, r)$. For each fixed $v \in \mathcal{V}_E$, let

$$\phi(v) := \sup_{(w^-, w^+) \in \mathcal{W}_v \times \mathcal{W}_v} \sum_{i=-m}^{n} \pi_i v(x - \xi_i).$$

Then $\sup_{(v, w^- , w^+)} \in \mathcal{V}_E \times \mathcal{W}_v \times \mathcal{W}_v \sum_{i=-m}^{n} \pi_i v(x - \xi_i) = \sup_v \in \mathcal{V}_E \phi(v)$. It suffices to derive the tractable reformulation of $\phi(v)$ for each $v \in \mathcal{V}_E$.

**Step 2.** Let $q_{-m-1} = 0$, $q_i = \sum_{j=-m}^{i} p_j$ for $i = -m, \cdots, n$. We choose the set of breakpoints of the piecewise linear weighting functions such that it contains $q_i, 1 - q_i$, i.e.,

$$\{ q_i : i = -m - 1, \cdots, -1 \} \cup \{ 1 - q_i : -1, \cdots, n \} \subset \{ t_1, \cdots, t_{T+1} \}. \quad (A.4)$$

Define the $\tilde{g}$-ball of piecewise linear inverse $S$-shaped weighting functions centered at $w_0$ with radius $r$

$$\tilde{B}(w_0, r) := \{ w \in \mathcal{V}_T \cap \mathcal{W}_{inS} : d_{\tilde{g}}(w, w_0) \leq r \}. \quad (A.5)$$

Denote by $\Psi(\psi^0, r)$ the set of all derivative functions of $w \in B(w_0, r)$, i.e.,

$$\Psi(\psi^0, r) := \{ \psi \in \Psi : w \in B(w_0, r) \} = \left\{ \psi \in \Psi : \sup_{g \in \mathscr{G}} \left| \int_{0}^{1} g(t)(\psi(t) - \psi^0(t))dt \right| \leq r \right\},$$

where $\Psi$ is the set of all derivative functions of $w \in \mathcal{W}_{inS}$ \footnote{Note that $B(w_0, r)$ may consist of weighting functions which are not differentiable at some points of $[0,1]$. Here we assume that $w \in \mathcal{W}_{inS}$ is continuously differentiable over $[0,1]$ except a finite number of points in which case the non-differentiable points do not affect the integral $\int_{0}^{1} g(t)(\psi(t) - \psi^0(t))dt$.}. Note that for $\psi \in \Psi(\psi^0, r)$, it is not necessarily a step-like function. Thus we propose to use step-like functions with jump points at set $\{ t_2, \cdots, t_T \}$ to approximate it. Consequently we define

$$\tilde{\Psi}(\psi^0, r) = \left\{ \psi \in \Psi_T : \sup_{g \in \mathscr{G}} \left| \sum_{i=1}^{T} \int_{t_{i-1}}^{t_i} g(t)(\psi(t) - \psi^0(t))dt \right| \leq r \right\}, \quad (A.6)$$

where $\Psi_T$ is defined as in (3.22). Since $\pi_i = w^- (q_i) - w^- (q_{i-1})$, for $i = -m, \cdots, -1$, and $\pi_i = w^+ (1 - q_{i-1}) - w^+ (1 - q_i)$, for $i = 0, 1, \cdots, n$, we can write $\phi(v)$ as

$$\phi(v) = \sup_{(v^- , v^+) \in \Psi(\psi^0, r) \times \Psi(\psi^0, r)} \left\{ \sum_{i=-m}^{n} v^- (\xi_i - x) \int_{q_{i-1}}^{q_i} \psi^-(t)dt + \sum_{i=0}^{n} v^+ (\xi_i - x) \int_{1-q_{i-1}}^{1-q_i} \psi^+(t)dt \right\}. \quad (A.7)$$

Next, we will prove that $\phi(v)$ is equal to

$$\tilde{\phi}(v) := \sup_{(v^- , v^+) \in \tilde{\Psi}(\psi^0, r) \times \tilde{\Psi}(\psi^0, r)} \left\{ \sum_{i=-m}^{n} v^- (\xi_i - x) \int_{q_{i-1}}^{q_i} \psi^-(t)dt + \sum_{i=0}^{n} v^+ (\xi_i - x) \int_{1-q_{i-1}}^{1-q_i} \psi^+(t)dt \right\}.$$
for any $v \in \mathcal{V}_E$. Since $\bar{B}(w_0, r) \subset B(w_0, r)$, we have $\bar{\phi}(v) \leq \phi(v)$. It suffices to show that $\bar{\phi}(v) \geq \phi(v)$. Let $\varepsilon > 0$, and $w^{*-\varepsilon}, w^{+\varepsilon} \in B(w_0, r)$ be such that

$$\sum_{i=-m}^{-1} \left( w^{*-\varepsilon}(q_i) - w^{*-\varepsilon}(q_{i-1}) \right) v(\xi_i - x) \geq \sup_{(w^-,w^+) \in (B(w_0,r) \times B(w_0,r))} \left\{ \sum_{i=-m}^{-1} \left( w^-(q_i) - w^-(q_{i-1}) \right) v(\xi_i - x) + \sum_{i=0}^{n} (w^+(1 - q_i) - w^+(1 - q_i)) v(\xi_i - x) \right\} - \varepsilon. \quad (A.8)$$

We can find piecewise linear functions $w^{*-\varepsilon}_{\varepsilon,T}, w^{+\varepsilon}_{\varepsilon,T}$ such that $w^{*-\varepsilon}_{\varepsilon,T}(t_i) = w^{*-\varepsilon}(t_i)$ and $w^{+\varepsilon}_{\varepsilon,T}(t_i) = w^{+\varepsilon}(t_i)$, $i = 1, \cdots, T + 1$,

$$w^{*-\varepsilon}_{\varepsilon,T}(t) := w^{*-\varepsilon}_{\varepsilon,T}(t_{i-1}) + \frac{w^{*-\varepsilon}_{\varepsilon,T}(t_i) - w^{*-\varepsilon}_{\varepsilon,T}(t_{i-1})}{t_i - t_{i-1}}(t - t_{i-1}), \quad \text{for } t \in [t_{i-1}, t_i], i = 2, \cdots, T + 1,$n

$$w^{+\varepsilon}_{\varepsilon,T}(t) := w^{+\varepsilon}_{\varepsilon,T}(t_{i-1}) + \frac{w^{+\varepsilon}_{\varepsilon,T}(t_i) - w^{+\varepsilon}_{\varepsilon,T}(t_{i-1})}{t_i - t_{i-1}}(t - t_{i-1}), \quad \text{for } t \in [t_{i-1}, t_i], i = 2, \cdots, T + 1.$$

For $w^{*-\varepsilon}, w^{+\varepsilon} \in B(w_0, r)$, we have

$$d_{\bar{g}}(w^{*-\varepsilon}, w_0) = \sup_{g \in \mathcal{G}} \left| \int_0^1 g(t)(\psi^\varepsilon_{\varepsilon,T}(t) - \psi^0(t))dt \right| \leq r,$n

$$d_{\bar{g}}(w^{+\varepsilon}, w_0) = \sup_{g \in \mathcal{G}} \left| \int_0^1 g(t)(\psi^+_{\varepsilon,T}(t) - \psi^0(t))dt \right| \leq r.$$

Then

$$d_{\bar{g}}(w^{*-\varepsilon}_{\varepsilon,T}, w_0) = \sup_{g \in \mathcal{G}} \left| \sum_{i=2}^{T+1} \int_{t_{i-1}}^{t_i} g(t) \left( \frac{w^{*-\varepsilon}_{\varepsilon,T}(t_i) - w^{*-\varepsilon}_{\varepsilon,T}(t_{i-1})}{t_i - t_{i-1}} - \psi^0(t) \right) dt \right| = d_{\bar{g}}(w^{*-\varepsilon}, w_0) \leq r,$n

$$d_{\bar{g}}(w^{+\varepsilon}_{\varepsilon,T}, w_0) = \sup_{g \in \mathcal{G}} \left| \sum_{i=2}^{T+1} \int_{t_{i-1}}^{t_i} g(t) \left( \frac{w^{+\varepsilon}_{\varepsilon,T}(t_i) - w^{+\varepsilon}_{\varepsilon,T}(t_{i-1})}{t_i - t_{i-1}} - \psi^0(t) \right) dt \right| = d_{\bar{g}}(w^{+\varepsilon}, w_0) \leq r,$$

which means $w^{*-\varepsilon}_{\varepsilon,T}, w^{+\varepsilon}_{\varepsilon,T} \in \bar{B}(w_0, r)$. Since $q_i, 1 - q_i \in \{t_1, \cdots, t_{T+1}\}$ in (A.4), then $w^{*-\varepsilon}_{\varepsilon,T}(q_i) = w^{*-\varepsilon}(q_i), \cdots$.
The Lagrange dual can be written as

\[
\tilde{\phi}(v) = \sup_{(w^-, w^+) \in \bar{B}(w_0, r) \times B(w_0, r)} \left\{ -\sum_{i=-m}^{-1} (w_e^+(q_i) - w_e^-(q_{i-1})) v(\xi_i) - \varepsilon \right\}
\]

where the last inequality is due to (A.8). This means \( \tilde{\phi}(v) \geq \phi(v) \) since \( \varepsilon \) can be arbitrarily small. This shows \( \phi(v) = \tilde{\phi}(v) \) as desired.

**Step 3.** Let \( y_l := \int_{t_l}^{t_{l+1}} g(t) dt \). Since \( G \) constitutes all measurable functions with \( |g(t)| \leq \tilde{g}(t) \) for \( t \in [0,1] \), then we have \( |y_l| \leq \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \). The inequality in \( \tilde{\Psi}(\psi^0, r) \) (A.6) can be reformulated as follows

\[
\sup_{g \in G} \left| \sum_{l=1}^{T} \int_{t_l}^{t_{l+1}} g(t)(\psi_l - \psi_l^0) dt \right| = \sup_{y_1, \ldots, y_T} \sum_{l=1}^{T} (\psi_l - \psi_l^0) y_l \\
\text{s.t. } |y_l| \leq \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt, \ l = 1, \ldots, T. \tag{A.9}
\]

The right hand side is a linear program. We derive its Lagrange dual. Let \( \eta \in \mathbb{R}^T_{\geq 0} \) and \( \theta \in \mathbb{R}^T_+ \) denote the dual variables corresponding to constraints \(-y_l \leq \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \) for \( l = 1, \ldots, T \) and \( y_l \leq \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \) for \( l = 1, \ldots, T \) respectively. Let

\[
L(y; \eta, \theta) := \sum_{l=1}^{T} (\psi_l - \psi_l^0) y_l - \sum_{l=1}^{T} \eta_l \left( -y_l - \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \right) - \sum_{l=1}^{T} \theta_l \left( y_l - \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \right)
\]

\[
= \sum_{l=1}^{T} \left( \psi_l - \psi_l^0 + \eta_l - \theta_l \right) y_l + \sum_{l=1}^{T} \eta_l \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt + \sum_{l=1}^{T} \theta_l \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt.
\]

The Lagrange dual can be written as

\[
\inf_{\eta, \theta \geq 0} \sum_{l=1}^{T} (\eta_l + \theta_l) \int_{t_l}^{t_{l+1}} \tilde{g}(t) dt \\
\text{s.t. } \psi_l - \psi_l^0 + \eta_l - \theta_l = 0, l = 1, \ldots, T.
\]
Thus $\phi(v)$ in (A.7) can be reformulated as

$$
\phi(v) = \sup_{(\psi^-,\psi^+) \in \Psi(v^0,r) \times \Psi(v^0,\bar{r})} \left\{ \sum_{i=-m}^{-1} v^-(\xi_i - x) \int_{q_{i-1}}^{q_i} \psi^-(t)dt + \sum_{i=0}^{n} v^+(\xi_i - x) \int_{q_i}^{1-q_{i-1}} \psi^+(t)dt \right\}
$$

$$
= \sup_{\psi^-_{\theta},\psi^+_{\theta},\eta,\theta = -m} \sum_{i=-m}^{-1} v^-(\xi_i - x) \sum_{t=1}^{T} \psi^-_{t} \int_{q_{i-1}}^{q_i} \mathbb{I}_{[t,t+t)}(t)dt + \sum_{i=0}^{n} v^+(\xi_i - x) \sum_{t=1}^{T} \psi^+_{t} \int_{1-q_{i-1}}^{1-q_i} \mathbb{I}_{[t,t+t)}(t)dt
$$

s.t. \quad (4.5a) - (4.5k),

where $\tilde{\psi}^- = (\psi^-_1,\cdots,\psi^-_T)$ and $\tilde{\psi}^+ = (\psi^+_1,\cdots,\psi^+_T)$.

**Step 4.** We are now ready to reformulate the constraint function in (4.4)

$$
\sup_{(v,w^-,w^+) \in (V_{EI},W_r,W_r)} \sum_{i=-m}^{n} \pi_i v(\xi_i - x) = \sup_{v \in V_{EI}} \phi(v).
$$

First, we will prove

$$
\sup_{v \in \tilde{V}_{EI}} \phi(v) = \sup_{v \in V_{EI}} \phi(v), \tag{A.11}
$$

where $\tilde{V}_{EI}$ is defined as follows.

Let $\{y_j\}_{j=1}^J$ be the $j$-th entry of ranked support set $\mathcal{S} := \bigcup_{m=1}^{M}(\text{supp}(\gamma_m) \cup \text{supp}(\eta_m)) \cup \{a,b\} \cup \{-1\} \cup \bigcup_{i=1}^{N}(\xi_i - x) \cup \bigcup_{k=1}^{K}\text{supp}(A_k - a^+_k) \cup \bigcup_{k=1}^{K}\text{supp}(A_k - a^-_k)$. We will use $y_j$ to denote the $j$-th smallest entry of $\mathcal{S}$. Let $j_0$ be the index satisfying $y_{j_0} = 0$. Denote

$$
\tilde{\mathcal{Y}} := \left\{ v : [\alpha,\beta] \rightarrow \mathbb{R} : \begin{array}{l}
 v \text{ is negative and convex on } [\alpha,0], \text{ and positive and concave on } (0,\beta) \\
 v \text{ is increasing continuous, } v(\alpha) = b_l, v(\beta) = b_r
\end{array} \right\}.
$$

Let $\tilde{\mathcal{Y}}_j$ denote the set of all piecewise linear functions $v \in \tilde{\mathcal{Y}}$ with breakpoints $a = y_1 < y_2 < \cdots < y_J = b$, $y_1,\cdots,y_J \in \mathcal{S}$. For

$$
\mathcal{V}_{EI} = \mathcal{V}_{ce}^K \cap \mathcal{V}_{pair}^M \cap \mathcal{V}_S = \left\{ v \in \tilde{\mathcal{Y}} : \begin{array}{l}
 \mathbb{E}_{w^-,w^+}[v(\gamma_m)] \leq \mathbb{E}_{w^-,w^+}[v(\eta_m)], m = 1,\cdots,M \\
 \mathbb{E}_{w^-,w^+}[v(A_k - a^+_k)] \leq 0, k = 1,\cdots,K \\
 \mathbb{E}_{w^-,w^+}[v(A_k - a^-_k)] \geq 0, k = 1,\cdots,K
\end{array} \right\},
$$

we define the ambiguity set

$$
\tilde{\mathcal{V}}_{EI} = \mathcal{V}_{ce,J}^K \cap \mathcal{V}_{pair,J}^M \cap \mathcal{V}_S = \left\{ v_j \in \tilde{\mathcal{Y}}_j : \begin{array}{l}
 \mathbb{E}_{w^-,w^+}[v_j(\gamma_m)] \leq \mathbb{E}_{w^-,w^+}[v_j(\eta_m)], m = 1,\cdots,M \\
 \mathbb{E}_{w^-,w^+}[v_j(A_k - a^+_k)] \leq 0, k = 1,\cdots,K \\
 \mathbb{E}_{w^-,w^+}[v_j(A_k - a^-_k)] \geq 0, k = 1,\cdots,K
\end{array} \right\}.
$$

We consider to approximate $\mathcal{V}_{EI}$ by $\tilde{\mathcal{V}}_{EI}$. The approximation function $v_j \in \tilde{\mathcal{Y}}_j$ preserves convexity, concavity and increasing property, and then $\tilde{\mathcal{Y}}_j \subset \tilde{\mathcal{Y}}$. We choose $\gamma_m$ and $\eta_m$, $m = 1,\cdots,M$ with discrete distributions, and the breakpoints cover their support sets. Since

$$
\mathbb{E}_{w^-,w^+}[v(\xi)] = \int_{-\infty}^{\infty} w^+(\mathbb{P}(v(\xi) \geq t))dt - \int_{-\infty}^{0} w^-(\mathbb{P}(v(\xi) \leq t))dt
$$

$$
= \int_{0}^{\infty} w^+(1 - F_\xi(t))dv^+(t) - \int_{-\infty}^{0} w^-(F_\xi(t))dv^-(t),
$$

39
then
\[
\begin{align*}
\mathbb{E}_{w-w^+}[v_J(\gamma_m)] - \mathbb{E}_{w-w^+}[v_J(\eta_m)] &= \int_{y_{i}}^{y_{j}} \phi_m(y) dv_J(y) = \sum_{i=1}^{f-1} \phi_m(y_j) (v_J(y_{i+1}) - v_J(y_i)) \\
&= \sum_{i=1}^{f-1} \phi_m(y_j) (v(y_{i+1}) - v(y_i)) = \mathbb{E}_{w-w^+}[v(\gamma_m)] - \mathbb{E}_{w-w^+}[v(\eta_k)],
\end{align*}
\]

where \( \phi_m \) is defined as in (5.11).

\[
\begin{align*}
\mathbb{E}_{w-w^+}[v_J(A_k - a_k^+)] &= \int_{y_{j0}}^{y_{j}} w^+(1 - F_{A_k-a_k^+}(y)) dv_J^+(y) - \int_{y_{j0}}^{y_{j}} w^-(F_{A_k-a_k^+}(y)) dv_J^-(y) \\
&= \sum_{j=j_0}^{j-1} w^+(1 - F_{A_k-a_k^+}(y_j))(v_J^+(y_{j+1}) - v_J^+(y_j)) - \sum_{j=1}^{j_0} w^-(F_{A_k-a_k^+}(y_j))(v_J^-(y_{j+1}) - v_J^-(y_j)) \\
&= \mathbb{E}_{w-w^+}[v(A_k - a_k^+)]
\end{align*}
\]

and

\[
\begin{align*}
\mathbb{E}_{w-w^+}[v_J(A_k - a_k^-)] &= \int_{y_{j0}}^{y_{j}} w^+(1 - F_{A_k-a_k^-}(y)) dv_J^+(y) - \int_{y_{j0}}^{y_{j}} w^-(F_{A_k-a_k^-}(y)) dv_J^-(y) \\
&= \sum_{j=j_0}^{j-1} w^+(1 - F_{A_k-a_k^-}(y_j))(v_J^+(y_{j+1}) - v_J^+(y_j)) - \sum_{j=1}^{j_0} w^-(F_{A_k-a_k^-}(y_j))(v_J^-(y_{j+1}) - v_J^-(y_j)) \\
&= \mathbb{E}_{w-w^+}[v(A_k - a_k^-)].
\end{align*}
\]

This shows \( \tilde{V}_{EI} \subset V_{EI} \), and hence \( \sup_{v \in \tilde{V}_{EI}} \phi(v) \leq \sup_{v \in V_{EI}} \phi(v) \). Conversely, let \( \varepsilon > 0 \), and \( v_{\varepsilon}^* \in V_{EI} \) such that

\[
\phi(v_{\varepsilon}^*) \geq \sup_{v \in \tilde{V}_{EI}} \phi(v) - \varepsilon.
\]

We can find a piecewise linear function \( v_{\varepsilon,J}^* \in \tilde{V}_{EI} \) such that \( v_{\varepsilon,J}^*(\xi - x) = v_{\varepsilon}^*(\xi - x) \) for \( 1 \leq i \leq N \). Specifically, for any \( v \in V_{EI} \), we can construct such a \( v_J \) as

\[
v_J(y) := v(y_{i-1}) + \frac{v(y_i) - v(y_{i-1})}{y_i - y_{i-1}} (y - y_i), \quad \text{for } y \in [y_{i-1}, y_i], \ i = 2, \cdots, J,
\]

where \( y_1 = a, y_J = b \). Observe that \( v_J(y_i) = v(y_i) \) for each \( i \). Consequently we have

\[
\sup_{v \in \tilde{V}_{EI}} \phi(v) \geq \phi(v_{\varepsilon,J}^*) = \phi(v_{\varepsilon}^*) \geq \sup_{v \in V_{EI}} \phi(v) - \varepsilon.
\]

Then \( \sup_{v \in \tilde{V}_{EI}} \phi(v) \geq \sup_{v \in V_{EI}} \phi(v) \) follows as \( \varepsilon \) can be arbitrarily small, and thus (A.11) holds.
Next, by invoking (A.10), we can reformulate the constraint function in (4.4) as

\[
\sup_{(v,w^-,w^+)\in \mathcal{V}_{El}\times \mathcal{W}_i\times \mathcal{W}_i} \sum_{i=-m}^{n} \pi_i v(\xi_i - x) = \sup_{v \in \mathcal{V}_{El}} \phi(v)
\]

\[
= \sup_{v \in \mathcal{V}_{El}} \sup_{\psi^-,\psi^+,\eta,\theta} \sum_{i=-m}^{-1} v^-(\xi_i - x) \sum_{l=1}^{T} \psi_l^- \int_{q_{i-1}}^{q_i} \mathbb{1}_{[t_l,t_{l+1})}(t) dt + \sum_{i=0}^{n} v^+(\xi_i - x) \sum_{l=1}^{T} \psi_l^+ \int_{1-q_i}^{1-q_{i-1}} \mathbb{1}_{[t_l,t_{l+1})}(t) dt
\]

s.t. (4.5a) - (4.5k),

\[
= \sup_{a,\tilde{a},\psi^-},\psi^+,\eta,\theta} \left\{ \sum_{i=-m}^{-1} \left( (a_1(\xi_i - x) + b_1) \mathbb{1}_{(-\infty,0) \cap [y_1,y_2]}(\xi_i - x) \right) + \sum_{j=2}^{J-1} (a_j(\xi_i - x) + b_j) \mathbb{1}_{(-\infty,0) \cap [y_j,y_{j+1}])(\xi_i - x) \sum_{l=1}^{T} \psi_l^- \int_{q_{i-1}}^{q_i} \mathbb{1}_{[t_l,t_{l+1})}(t) dt \right. \\
+ \left. \sum_{i=0}^{n} \left( (a_1(\xi_i - x) + b_1) \mathbb{1}_{(-\infty,0) \cap [y_1,y_2]}(\xi_i - x) \right) + \sum_{j=2}^{J-1} (a_j(\xi_i - x) + b_j) \mathbb{1}_{(-\infty,0) \cap [y_j,y_{j+1}])(\xi_i - x) \sum_{l=1}^{T} \psi_l^+ \int_{1-q_i}^{1-q_{i-1}} \mathbb{1}_{[t_l,t_{l+1})}(t) dt \right\}
\]

s.t. (4.5a) - (4.5k), (4.5l) - (4.5r).

(A.12a)

The proof is complete. \(\blacksquare\)