

Provided for non-commercial research and educational use. Not for reproduction, distribution or commercial use.
--

Serdica

Mathematical Journal

Сердика

Математическо списание

The attached copy is furnished for non-commercial research and education use only.

Authors are permitted to post this version of the article to their personal websites or institutional repositories and to share with other researchers in the form of electronic reprints.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to third party websites are prohibited.

For further information on
Serdica Mathematical Journal
which is the new series of
Serdica Bulgaricae Mathematicae Publicationes
visit the website of the journal <http://serdica.math.bas.bg>
or contact: Editorial Office
Serdica Mathematical Journal
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: serdica@math.bas.bg

RELATIONS BETWEEN RISK-AVERSE MODELS IN EXTENDED TWO-STAGE STOCHASTIC OPTIMIZATION*

Darinka Dentcheva[†]

Communicated by M. Krastanov

ABSTRACT. We consider an extended two-stage risk-averse stochastic optimization problems in several formulations. Risk-aversion is reflected by risk constraints in form of stochastic-order relations, which are imposed in a time-consistent manner. The problem is analyzed under convexity assumptions. The main goal of this study is to establishing relations between the extended two-stage problem with stochastic-order constraints on the recourse function on the one hand and the two-stage problems with alternative models of risk such as utility functions, distortions, or coherent measures of risk on the other hand.

2020 *Mathematics Subject Classification:* 90C15, 90C25, 91B05, 91B16.

Key words: stochastic dominance, coherent measures of risk, dual utility, distortions, stochastic programming.

*This paper is dedicated to Asen Dontchev, with gratitude for his mentorship and in appreciation of his fundamental contributions to the theory of optimization and control.

[†]The support of the Office of Naval Research under grant N00014-21-1-216 is gratefully acknowledged.

1. Introduction. The two-stage stochastic optimization problem formulation is one of the most popular model in the area of optimization under uncertainty. In its risk-neutral (expected-value) formulation, it has been thoroughly investigated under various assumptions. Here, we consider an extended version of the two-stage problem, which involves two decision stages and two observation stages. This type of model is a version of a sequential decision problem presented in [12]. In that paper, the emphasis is put on the time-consistency of the relation and the numerical method for solving the decision problem. Here, we present more thorough analysis of the model and establish its connection to sequential preferences reflected via utility functions, distortion (also called rank-dependent or dual utilities), and coherent measures of risk. The relation to utility functions was also announced in [12] for a similar problem formulation.

The problem in its risk-neutral formulation is constructed as follows. At the first decision period, we take a “here and now” decision x in an appropriate decision space space. In the next stage, we observe the uncertain data involved and take a decision y , depending on the observations. That process concludes the risk-neutral version of the problem formulation. When risk-aversion is involved, then another stage of the process is pertinent. At that final stage, we make a new observation and evaluate the performance of our decision policy without further decisions. More, precisely, let the random data involved in the model be comprised in the random sequence ξ_1, ξ_2 , where ξ_1 is a random vector defined on a probability space $(\Omega, \mathcal{F}_1, P)$ and ξ_2 is a random vector defined on the space $(\Omega, \mathcal{F}_2, P)$ with $\mathcal{F}_1 \subset \mathcal{F}_2$. The support of the distribution of ξ_t is denoted by Ξ_t , $t = 1, 2$, and $\xi_{[2]}$ denotes a path of the process $\{\xi_1, \xi_2\}$. We consider the following risk-neutral extended two-stage problem:

$$(1) \quad \min_{x \in \mathcal{X}_1} f_1(x) + \mathbb{E} \left[\inf_{y \in \mathcal{X}_2(x, \xi_1)} \left(f_2(y, \xi_1) + \mathbb{E}_{\mathcal{F}_1} [f_3(B(\xi_2)y, \xi_2)] \right) \right].$$

Here $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ is a convex function expressing the known cost of the first-stage decision $x \in \mathbb{R}^n$, $f_2 : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a convex function of its first argument and $\mathcal{F}_1$ -measurable with respect to its second argument, while $f_3 : \mathbb{R}^{m'} \times \mathbb{R}^{d'} \rightarrow \mathbb{R}$ is a convex function of the first argument and $\mathcal{F}_2$ -superpositionally measurable. The random matrix $B(\xi_2)$ has $m' \times m$ dimensional realizations. Further, $\mathcal{X}_1$ is a closed subset of $\mathbb{R}^n$ and $\mathcal{X}_2 : \mathbb{R}^n \times \mathbb{R}^d \rightrightarrows \mathbb{R}^m$ is a measurable closed-valued multifunction. They are defined as follows:

$$\begin{aligned} \mathcal{X}_1 &= \{x \in \mathcal{D} \subset \mathbb{R}^n : Ax = b\}, \\ \mathcal{X}_2(x, \xi_1) &= \{y \in \mathcal{Y}(\xi_1) \subset \mathbb{R}^m : T(\xi_1)x + Wy = h(\xi_1)\}. \end{aligned}$$

The set $\mathcal{D}$ is closed and convex, A is $s \times n$ the matrix and $b \in \mathbb{R}^s$. The realizations of the random elements are as follow: T is $m \times n$ matrix and h is m -dimensional, $\mathcal{Y}(\xi_1) \subset \mathbb{R}^m$ is a closed convex set. The feasible second-stage decisions are measurable selections: $y(\xi_1) \in \mathcal{X}_2(x, \xi_1)$.

The risk-neutral problem (1) has the following equivalent formulation

$$(2) \quad \min_x f(x) + \mathcal{Q}_1(x, \xi_1) \quad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{D}.$$

The function $\mathcal{Q}_1(x, \xi_1)$ is the expectation of $Q_1(x, \xi_1)$, where the latter denotes the optimal values of the following second-stage problem:

$$(3) \quad \begin{aligned} \min_y & f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]}) \\ \text{s.t.} & T(\xi_1)x + Wy = h(\xi_1), \quad y \in \mathcal{Y}(\xi_1). \end{aligned}$$

The notation $\mathcal{Q}_2(x, \xi_{[2]})$ stands for the $\mathcal{F}_1$ - conditional expectation of $Q_2(y, \xi_{[2]}) = f_3(B(\xi_2)y, \xi_2)$. We shall also use the shorthand notation $\mathbb{E}_1[Z] = \mathbb{E}[Z | \mathcal{F}_1]$ when needed. Throughout the paper, we shall assume that $\mathbb{E}_1[\|T(\xi_1)\|]$ and $\mathbb{E}_1[\|h(\xi_1)\|]$ are finite.

If for some x and $\xi_1 \in \Xi_1$ the second stage problem is infeasible, then $Q(x, \xi_1) = +\infty$ by definition. We assume that the second stage problem is always bounded from below and hence $Q(x, \xi_1) > -\infty$. This is a natural assumption because otherwise the value of the second stage problem can be improved indefinitely for some value of the first stage decision and some realization of the random data, which does not occur in any practical situation.

We are concerned with risk and that is why we are interested not only in minimizing the total expected cost resulting from the first- and the second-stage decisions but we also would like to prevent possible very large cost realizations, which may occur with relatively small probability. That is, we associate risk with small probability events, which may result in high cost. Various models could be employed to reflect risk-aversion of this type. One possibility is to use the classical model based on von Neumann and Morgenstern expected utility functions. In this case, we may choose convex non-decreasing functions $u_0 : \mathbb{R} \rightarrow \mathbb{R}$, $u_{\xi_1} : \mathbb{R} \rightarrow \mathbb{R}$ and formulate the following decision problem:

$$(4) \quad \begin{aligned} \min_x & f_1(x) + \mathbb{E}[u_0(Q_1^u(x, \xi_1))] \quad \text{s.t.} \quad x \in \mathcal{X}_1, \\ Q_1^u(x, \xi_1) = \min_y & f_2(y, \xi_1) + \mathbb{E}_1[u_{\xi_1}(Q_2(y, \xi_{[2]}))] \quad \text{s.t.} \quad y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

Next to the expected utility theory, another *dual utility theory* or *rank dependent expected utility theory* has attracted much attention. The ideas of this theory

are presented first in [24] but their mathematical foundation is also present in Choquet integral theory ([3, 31, 32]). Rank dependent utilities are used in actuarial mathematics (see [34, 35]), where they are also referred to as distortions. For a comprehensive treatment, we refer to [25]; see also [38] for a special case discussion. From a different system of axioms than those of von Neumann and Morgenstern, one derives that every decision maker has a certain distortion function $w : [0, 1] \rightarrow \mathbb{R}$. The decision maker prefers a nonnegative outcome X over a nonnegative outcome Y , if and only if

$$(5) \quad \int_0^1 w(p) dF_X^{-1}(p) \leq \int_0^1 w(p) dF_Y^{-1}(p),$$

where F_X^{-1} and F_Y^{-1} are the respective inverse function of the random variables X and Y , to be defined more precisely in the next section. Hence another risk-averse model would be obtained by choosing convex non-decreasing distortion functions $w_1 : [0, 1] \rightarrow \mathbb{R}$, and $w_{\xi_1} : [0, 1] \rightarrow \mathbb{R}$. Denoting the quantile functions of $Q_1(x, \xi_1)$ and $Q_2(y, \xi_{[2]})$ by F_x^{-1} and F_{y, ξ_1}^{-1} , respectively, we formulate the following decision problem with distortions:

$$(6) \quad \begin{aligned} \min_x f_1(x) + \int_0^1 F_x^{-1}(t) dw_1(t) \quad & \text{s.t } x \in \mathcal{X}_1. \\ \min_y f_2(y, \xi_1) + \int_0^1 F_{y, \xi_1}^{-1}(t) dw_{\xi_1}(t) \quad & \text{s.t } y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

A natural risk-averse decision model based on the most recent developments in risk theory would involve a dynamic measure of risk, which is composed of two risk measures: $\varrho_1 : \mathcal{L}_1(\Omega, \mathcal{F}_1, P) \rightarrow \mathbb{R} \cup \{+\infty\}$ and $\varrho_2 : \mathcal{L}_1(\Omega, \mathcal{F}_2, P) \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}_1, P)$. Here $\mathcal{L}_1(\Omega, \mathcal{F}_t, P)$, $t = 1, 2$, stands for the random variables with finite first moments, indistinguishable if differing on sets of P -measure zero. The decision problem would have a form:

$$(7) \quad \min_{x \in \mathcal{X}_1} f_1(x) + \varrho_1 \left[\inf_{y \in \mathcal{X}_2(x, \xi_1)} \left(f_2(y, \xi_1) + \varrho_2(\xi_1) [f_3(B(\xi_2)y, \xi_2)] \right) \right].$$

Finally, we may adopt the point of view, that a random benchmark sequence V_1, V_2 , adapted to $\mathcal{F}_1, \mathcal{F}_2$ is available, whose distribution is acceptable for the random costs resulting from the decision policy. We formulate the first stage problem requiring that $Q_1(x, \cdot)$ at the optimal solution x is stochastically smaller than V_1 plus the projected future cost $\mathbb{E}_1[V_2]$. This relation is denoted $\mathcal{Q}(x, \xi_1) \preceq$

$V_1 + \mathbb{E}_1[V_2]$ for a suitably chosen stochastic order. The notation $V_2|\xi_1$ stands for the realizations of V_2 given the history ξ_1 . At the second stage, we request that the cost $f_2(x, \xi_1) + Q_2(y, \xi_{[2]})|\xi_1 \preceq V_1(\xi_1) + V_2|\xi_1$ for $\mathcal{F}_1$ a.a. realizations of ξ_1 .

Second-order stochastic dominance is the most popular stochastic order when large outcomes are preferred due to its consistency with risk-averse preferences. In our context, the increasing convex order is an appropriate choice of a stochastic ordering relation because it is a counterpart of the second-order stochastic dominance when small outcomes are preferred. The problem takes on the form

$$(8) \quad \begin{aligned} & \min_x f_1(x) + Q_1(x, \xi_1) \\ & \text{s.t. } Q_1(x_1, z_1) \preceq V_1 + \mathbb{E}_1[V_2], \\ & \quad x \in \mathcal{X}. \end{aligned}$$

with $Q_1(x, \xi_1)$ being the optimal value of the problem

$$(9) \quad \begin{aligned} & \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) \\ & \text{s.t. } f_2(y, \xi_1) + Q_2(y, \xi_{[2]})|\xi_1 \preceq V_1(\xi_1) + V_2|\xi_1, \\ & \quad y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

and $Q_2(y, \xi_{[2]}) = f_3(B(\xi_2)y, \xi_2)$.

The goal of this paper is to analyze problem (8)–(9) and to show its relations to models (4), (6), and (7).

The paper is organized as follows. Section 2 contains the notions and background results necessary for our study, as well as some bibliographical remarks. In Section 3, we present the formulation of the problem, where the order relation is a relaxed second-order dominance expressed in terms of shortfall functions. We show sufficient conditions for the convexity of the problem and the existence of optimal policy. That problem formulation leads to duality with problems using expected utility functions as shown in Section 4. Another formulation of the extended two-stage problem with sequential dominance constraint is given in Section 5, where this formulation is analyzed and its relations to coherent measures of risk are established. Finally, we show that the formulation of the stochastic dominance constraint via integrated quantile functions leads to duality between problem (8)–(9) and (6). This analysis is carried out in Section 6.

2. Stochastic orders and measures of risk. Stochastic orders are widely used in statistics, insurance, and economics, see e.g. [17, 26, 34]. The order

known as the stochastic dominance relation of second order is the most popular comparison between random outcomes in those disciplines due to its consistency with risk-averse preferences. This relation compares scalar random variables; it is a key notion in plenitude of papers. However, the notions serving to compare random sequences are not so well established and the literature on that topic is scarce, despite the prevalence of decision problems for dynamical systems in practice. We refer the reader to the monographs [29, 20] for an overview on univariate stochastic orders and their generalizations to multiple random outcomes and processes; see also [27, 36, 37] for proposals on conditional stochastic orders for random sequences.

Optimization problems with stochastic dominance constraints were introduced in [5]. Optimality and duality conditions for convex optimization problems with stochastic dominance constraints were developed in [5, 6, 11]. We have shown that optimization with stochastic-order constraints relate to many other risk-averse models such as optimization using coherent measures of risk [8], utility functions [6], distortion [7], chance constraints, or Average (Conditional) Value-at-Risk constraints [9]. Classical two-stage problems with stochastic ordering constraints on the recourse function were considered in [15, 4]. In [10], a control problem with a so-called discounted stochastic dominance is analyzed. However, the relation used in that paper does not provide a time-consistent comparison. Order-constraints in dynamic optimization are discussed also in [10, 12, 13, 14, 16]. The proposal in [16] addresses the limiting distribution in an infinite-time horizon average-cost Markov decision problem. In [13, 14], time-consistent constraints are imposed in a manner akin to Average Value-at-Risk constraints for a multi-stage stochastic optimization problem. Average Value-at-Risk can be represented as the expected value of a nonlinear function, which facilitates the application of the theory and methods for risk-neutral optimization. Here, we follow the idea presented in [12] and study the problems in this special case in detail.

The right-continuous cumulative distribution function $F_Z(\eta)$ of a random variable Z is defined as follows $F_Z(\eta) = P(Z \leq \eta)$ and the survival function is given by $\bar{F}_Z(\eta) = P(Z > \eta)$. The integrated distribution function $F_Z^{(2)}(\eta)$ and the integrated survival function $H_Z(\eta)$ are defined as follows:

$$F_Z^{(2)}(\eta) = \int_{-\infty}^{\eta} F_Z(t) dt \text{ for } \eta \in \mathbb{R}$$

$$H_Z(\eta) = \int_{\eta}^{\infty} P(Z > t) dt \text{ for } \eta \in \mathbb{R}.$$

Clearly, the functions $F_Z^{(2)}(\cdot)$ and $H_Z(\cdot)$ are finite everywhere whenever Z is integrable.

Definition 2.1. *A random variable Z is stochastically smaller than a random variable V with respect to the first order stochastic dominance (denoted $Z \succeq_{(1)} V$) if $\bar{F}_Z(\eta) \leq \bar{F}_V(\eta)$ for all $\eta \in \mathbb{R}$.*

This relation also means that “ V is stochastically larger than Z ” because V takes larger values more frequently.

Definition 2.2. *For two integrable random variables V and Z ,*

- *The variable V is stochastically larger than Z with respect to the second order stochastic dominance (denoted $Z \succeq_{(2)} V$) if $F_Z^{(2)}(\eta) \leq F_V^{(2)}(\eta)$ for all $\eta \in \mathbb{R}$.*
- *The variable Z is stochastically smaller than V with respect to the increasing convex order (denoted $Z \preceq_{ic} V$) if $H_Z(\eta) \leq H_V(\eta)$ for all $\eta \in \mathbb{R}$.*

We shall use the notation $a_+ = \max(0, a)$ for $a \in \mathbb{R}$. Changing the order of integration in the definitions of the functions $F_Z^{(2)}(\cdot)$ and $H_Z(\cdot)$, we obtain the following equivalent representations and relation between the two notions (cf. also [20, 23, 4]):

$$(10) \quad V \succeq_{(2)} Z \quad \Leftrightarrow \quad \mathbb{E}(\eta - V)_+ \leq \mathbb{E}(\eta - Z)_+, \quad \text{for all } \eta \in \mathbb{R};$$

$$(11) \quad Z \preceq_{ic} V \quad \Leftrightarrow \quad \mathbb{E}(Z - \eta)_+ \leq \mathbb{E}(V - \eta)_+, \quad \text{for all } \eta \in \mathbb{R};$$

$$(12) \quad Z \preceq_{ic} V \quad \Leftrightarrow \quad -Z \succeq_{(2)} -V.$$

The second order stochastic dominance and the increasing convex ordering relation can also be characterized by quantile functions. Let $F_Z^{-1}(\cdot)$ be the left continuous inverse of the cumulative distribution function $F_Z(\cdot)$ defined by

$$F_Z^{-1}(p) = \inf\{\eta : F_Z(\eta) \geq p\}, \quad \text{for } 0 < p < 1.$$

The absolute Lorenz function $L_Z : [0, 1] \rightarrow \mathbb{R}$, introduced in the seminal work [19], is defined as the cumulative quantile function:

$$L_Z(p) = \int_0^p F_Z^{-1}(t) dt \quad \text{for } 0 < p \leq 1.$$

The definition of the function beyond the interval $(0, 1]$ is extended by setting $L_Z(0) = 0$ and $L_Z(p) = \infty$ for $p \notin [0, 1]$. The Lorenz function is widely used

in economics for comparison of income streams. We define the upper Lorenz function as follows:

$$(13) \quad \bar{L}_Z(p) = \int_p^1 F_Z^{-1}(t) dt \quad \text{for } 0 \leq p < 1,$$

and extend it similarly: $\bar{L}_Z(1) = 0$, and $\bar{L}_Z(p) = -\infty$ for $p \notin [0, 1]$. It is a concave function because its derivative is monotonically non-increasing.

Interestingly the integrated distribution (survival) function and the (upper) Lorenz function are related via conjugate duality. The following relations are known: $L_Z(\cdot)$ and $F_Z^{(2)}(\cdot)$ are Fenchel conjugate functions (see [23, Theorem 3.1]. Similarly, $H_Z(\cdot)$ and $-\bar{L}_Z(\cdot + 1)$ are Fenchel conjugate functions ([4, Theorem 1]). These results imply that relating the Lorenz functions of two integrable random variables provides equivalent characterizations of the respective stochastic ordering relations, i.e.,

$$(14) \quad V \succeq_{(2)} Z \Leftrightarrow L_V(p) \geq L_Z(p) \quad \text{for all } p \in [0, 1];$$

$$(15) \quad Z \preceq_{ic} V \Leftrightarrow \bar{L}_Z(p) \leq \bar{L}_V(p) \quad \text{for all } p \in [0, 1].$$

A general way of characterizing a stochastic order relation is to fix a family $\mathcal{U}$ of functions $u : \mathbb{R} \rightarrow \mathbb{R}$ and call a scalar random variable V *stochastically larger* than another random variable Z if

$$(16) \quad \mathbb{E}[u(V)] \geq \mathbb{E}[u(Z)] \quad \text{for all } u \in \mathcal{U}.$$

The family $\mathcal{U}$ is called a generator of the order and the functions u are usually called utility functions. It is known that the generator is not uniquely determined and that every generator is a cone in a suitable functional space. The (maximal) generator of the second-order stochastic dominance relation consists of all non-decreasing concave functions, while the generator of the increasing convex order consists of all non-decreasing convex functions (c.f. [20]). Formulae (10) and (11) show the minimal generators of the two relations.

Now we introduce the notion of coherent measures of risk. A *coherent measure of risk* ϱ is a convex, monotonically increasing, and positively homogeneous functional defined on $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$, which may take a real value $\varrho[X]$ or $+\infty$. It also satisfies, a certain translation property. More, precisely, the following holds for random variables expressing cost or losses, i.e, when small outcomes are preferred.

(R1) Convexity For all $X, Y \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and for all $\alpha \in [0, 1]$, we have $\varrho[\alpha X + (1 - \alpha)Y] \leq \alpha \varrho[X] + (1 - \alpha) \varrho[Y]$.

(R2) Monotonicity If $X, Y \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and $X \leq Y$ P -a.s., then $\varrho[X] \leq \varrho[Y]$.

(R3) Translation Equivariance If $a \in \mathbb{R}$ and $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$, then $\varrho[X + a] = \varrho[X] + a$.

(R4) Positive homogeneity If $t > 0$ and $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$, then $\varrho[tX] = t\varrho[X]$.

This concept was introduced in [1] for functionals ϱ defined on $\mathcal{L}_\infty(\Omega, \mathcal{F}, P)$. For more information and further developments, we refer to [30] and the references therein. The probability space in the theory of measures of risk is typically assumed atomless, and we also adopt this assumption.

A key role is played by a measure called the Average Value at Risk at level p , where $p \in (0, 1]$. It is denoted $\text{AVaR}_p[\cdot]$ and it is also called Conditional Value at Risk. The *Value at Risk* of X at level p is defined as $\text{VaR}_p[X] = F_X^{-1}(1 - p)$.

The *Average Value at Risk* of X at level p can be expressed in several equivalent ways. We use

$$\text{AVaR}_p[X] = \frac{1}{p} \int_{1-p}^1 \text{VaR}_{1-p}[X] dt = \frac{1}{p} \bar{L}_X(1 - p) = \inf_{\eta} \left\{ t + \frac{1}{p} \mathbb{E}[(X - \eta)_+] \right\}.$$

A measure of risk ϱ is called *spectral* if a probability measure μ on $(0, 1]$ exists such that for all X

$$\varrho[X] = \int_0^1 \text{AVaR}_p[X] \mu(dp).$$

A coherent measures of risk ϱ is called *law-invariant* if $\varrho[X] = \varrho[Y]$ whenever X and Y have the same distribution. The following fundamental result is first shown in [18] (see [30, Theorem 6.40] for the general setting). For an atomless space Ω , every law invariant, finite-valued coherent measure of risk on $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ is associated with a convex set $\mathfrak{M}$ of probability measures on $(0, 1]$ such that for all $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ we have

$$\varrho[X] = \sup_{\mu \in \mathfrak{M}} \int_0^1 \text{AVaR}_p[X] \mu(dp).$$

It is clear that every finite-valued law-invariant measure of risk is consistent with the increasing convex order, i.e., if $Z \preceq_{\text{ic}} V$, then $\varrho[Z] \leq \varrho[V]$ (cf. [30, Theorem 6.51].)

To address evaluation of risk in a dynamic situation, we use the notion of a conditional risk mapping. A mapping $\rho : \mathcal{L}_1(\Omega, \mathcal{F}_2, P) \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}_1, P)$ is a

coherent *conditional risk mapping* if it satisfies the modified axioms (R1), (R2), and (R4), in which the equations and inequalities are understood in the $\mathcal{F}_1$ -almost sure sense and the axiom (R3) is modified as follows:

(R3') *Predictable Translation Equivariance*: If $Y \in \mathcal{L}_1(\Omega, \mathcal{F}_1, P)$ and

$$Z \in \mathcal{L}_1(\Omega, \mathcal{F}_2, P), \text{ then } \rho[Z + Y] = \rho[Z] + Y \text{ } \mathcal{F}_1\text{-a.s.}$$

If the sigma algebra $\mathcal{F}_1$ is trivial, then the definition space can be identified with $\mathbb{R}$, and the notion of conditional risk mapping coincides with that of coherent risk measure.

3. Primal formulation properties. We shall consider several formulations of the increasing convex order constraints for integrable random vector $V = (V_1, V_2)$. We start with the primal (direct) formulation first.

$$(17) \quad \min_x f_1(x) + Q_1(x, \xi_1)]$$

$$(18) \quad \text{s.t. } \mathbb{E}[(Q_1(x_1, \xi_1) - \eta)_+] \leq \mathbb{E}[(V_1 + \mathbb{E}_1[V_2] - \eta)_+], \quad \eta \in I_0;$$

$$(19) \quad x \in \mathcal{X}_1.$$

with $Q_1(x, \xi_1)$ being the optimal value of the problem

$$(20) \quad \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]})$$

$$(21) \quad \text{s.t. } \mathbb{E}[(f_2(y, \xi_1) + Q_2(y, \xi_{[2]})(\xi_1))_+] \leq \mathbb{E}[(V_1(\xi_1) + V_2(\xi_1))_+] \quad \eta \in I(\xi_1);$$

$$(22) \quad y \in \mathcal{X}_2(x, \xi_1).$$

and $Q_2(y, \xi_{[2]}) = f_3(B(\xi_2)y, \xi_2)$. In this formulation, $I(\cdot)$ is a measurable multi-function with closed images in $\mathbb{R}$.

We make the following assumptions:

(A1) The functions $f_1(\cdot)$, $f_2(\cdot, \xi_1)$, $f_3(\cdot, \xi_2)$ are convex for a.a. $\xi_{[2]}$ and $f_2(x, \cdot)$, $f_3(B(\cdot)y, \cdot)$ are integrable for $\mathcal{F}_1$ -a.a. measurable selections $y(\xi_1) \in \mathcal{X}_2(x, \xi_1)$ for $x \in \mathcal{X}_1$. Furthermore, $f_3(B(\xi_2)\cdot, \xi_2)$ is bounded a.s. by an integrable constant $\kappa(\xi_2)$ for $\mathcal{F}_1$ -a.a. $y \in \mathcal{Y}(\xi_1)$.

(A2) The sets $\mathcal{D}$ and $\mathcal{Y}(\xi_1)$ are nonempty, closed, and compact. The sets $\mathcal{Y}(\xi_1)$ are contained in a compact set $\mathcal{Y}_0$ for a.a. ξ_1 and for any feasible solution x , the feasible set of the second-stage problem is non-empty for a.a. ξ_1 .

- (A3) One of the following is true: (a) The probability space $(\Omega, \mathcal{F}, P)$ is finite, the sets $\mathcal{D}$, $\mathcal{Y}(\xi_1)$ are polyhedral and all functions are random polyhedral functions; (b) The sets I_0 and $I(\xi_1)$ are compact and a decision policy $(\tilde{x}, \tilde{y}) \in \mathcal{X}_1 \times \mathcal{X}_2(\tilde{x}, \xi_1)$ exist such that the following holds

$$\sup_{\eta \in I(\xi_1)} \mathbb{E}_1[(f_2(\tilde{y}, \xi_1) + Q_2(\tilde{y}, \xi_{[2]})|\xi_1 - \eta)_+ - (V_1(\xi_1) + V_2|\xi_1 - \eta)_+] < 0$$

for a.a. ξ_1 ;

$$\sup_{\eta \in I_0} \mathbb{E}[(Q_1(\tilde{x}, \xi_1) - \eta)_+ - (V_1 + \mathbb{E}_1[V_2] - \eta)_+] < 0,$$

We shall see in due course that Assumption (A1) ensures that the optimization problems at the first and the second stage are convex and are well-defined. Assumption (A2) ensures that the problems at the first and the second stage are solvable; the requirement that the feasible set at the second stage is non-empty for every feasible point at the first stage is known as relatively complete recourse. Assumption (A3) plays the role of a constraint qualification condition. In case (a), the optimization problems in both stages have polyhedral constraints. In case (b), the condition is a counterpart to the uniform dominance condition in [6]. In that case, the assumption means that it is possible to obtain a solution that is strictly smaller than the benchmark in the given range.

As the increasing convex order is relaxed in the case of general probability space, the generator of the relaxed order does not contain all convex non-decreasing function.

Lemma 3.1. *Let $I = [a, b]$, where $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{+\infty\}$, and consider the relaxation of the increasing convex order, in which (11) holds for all $\eta \in [a, b]$. The generator of this relaxation, $\mathcal{U}(I)$, consists of all convex nondecreasing functions $u(\cdot)$ satisfying the following conditions 1) $u(t) = 0$ for all $t \leq a$ if $a \in \mathbb{R}$ and 2) for all $t > b$, $u(t) = u(b) + c(t - b)$ with some constant $c > 0$ if $b < \infty$.*

Proof. In [5], a relaxation of the second-order dominance is discussed, in which (10) is required only on a compact interval instead of the entire real line. The form of the generator for that relaxation consists of all concave non-decreasing function, which are constant above b and linear below a . The statement can be proved in exactly the same way and the proof is omitted here. $\square$

The space of continuous functions on a compact set $I \subset \mathbb{R}$ is denoted by $\mathcal{C}(I)$ and the set of regular countably additive non-negative measures on I , which have finite variation, is denoted by $\mathcal{M}(I)$.

Theorem 3.2. *Assume (A1) – (A3). Then problem (17)–(19) and problems (20)–(22) (for each fixed ξ_1) are convex problems. Furthermore, $Q_1(\cdot, \xi_1)$, $Q_1(\cdot, \xi_1)$, $Q_2(\cdot, \xi_{[2]})$, $Q_2(\cdot, \xi_{[2]})$ are convex and finite-valued for any measurable selection $(x, y(\xi_1)) \in \mathcal{X}_1 \times \mathcal{X}_2(x, \xi_1)$.*

Proof. We observe that given ξ_1 , $Q_2(y, \xi_{[2]}) = f_3(B(\xi_2)y, \xi_2)$ is a finite-valued convex function with respect to y for a.a. ξ_2 by (A1). Furthermore, $|Q_2(y, \xi_{[2]})| \leq \kappa(\xi_2)$ by (A2) and, hence, $Q_2(\cdot, \xi_{[2]})$ is finite-valued and convex. The function $y \mapsto (f_2(y, \xi_1) + Q_2(y, \xi_{[2]})|\xi_1 - \eta|)_+$ is convex for $\mathcal{F}_2$ -a.a. ξ_2 and any fixed η . Thus, the mapping $y \mapsto \mathbb{E}_1[(f_2(y, \xi_1) + Q_2(y, \xi_{[2]})|\xi_1 - \eta|)_+]$ is convex for any fixed η , which entails that the set

$$Y(\xi_1) = \left\{ y \in \mathbb{R}^m : \mathbb{E}_1[(f_2(y, \xi_1) + Q_2(y, \xi_{[2]})|\xi_1 - \eta|)_+] \leq \mathbb{E}_1[(V_1(\xi_1) + V_2|\xi_1 - \eta|)_+] \right\}$$

is convex. It is also closed due to the continuity of the constraint function. The set $\mathcal{Y}(\xi_1)$ is a compact convex set, which together with assumption (A2) implies that the feasible set of the second-stage problem is non-empty, convex, and compact. Thus, the second-stage problem is a convex optimization problem, which is solvable. This implies that $Q_1(\cdot, \xi_1)$ is finite-valued.

We define the function

$$\varphi(y, \xi_1) = f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \mathbf{1}_{\mathcal{B}}(y, \xi_1),$$

where $\mathbf{1}_{\mathcal{B}}$ is the indicator function of convex analysis and the set $\mathcal{B}$ is determined by the constraints (21) and $y \in \mathcal{Y}(\xi_1)$. Observe that $\varphi(\cdot, \xi_1)$ is convex. We assign Lagrange multipliers π to the constraint $Wy = h(\xi_1) - T(\xi_1)x$ and obtain the dual function to this partial Lagrangian relaxation of problem (20)–(22). The dual function has the following form:

$$\begin{aligned} L_D(\pi, \xi_1) &= \min_y \left(\varphi(y, \xi_1) + \pi^\top (h(\xi_1) - T(\xi_1)x - Wy) \right) \\ &= \pi^\top (h(\xi_1) - T(\xi_1)x) - \max_y \left(\pi^\top Wy - \varphi(y, \xi_1) \right) = \pi^\top (h(\xi_1) - T(\xi_1)x) - \varphi^*(W^\top \pi) \end{aligned}$$

At the right-hand side, φ^* is the Fenchel conjugate of φ . The dual problem is

$$\max_{\pi} L_D(\pi, \xi_1).$$

Duality relation holds due to the convexity of the problems. Hence,

$$(23) \quad Q_1(x, \xi_1) = \max_{\pi} \left(\pi^\top (h(\xi_1) - T(\xi_1)x) - \varphi^*(W^\top \pi) \right).$$

This implies, that $Q_1(\cdot, \xi_1)$ is a convex function. It remains to show that $Q_1(x, \cdot)$ is integrable. This will imply that $Q_1(\cdot)$ is convex and finite valued. Under assumption (A3) in case (a), the integrability is obvious. In the case (b), $Q_1(\tilde{x}, \cdot)$ is integrable and due to the compactness of $\mathcal{D}$ and the convexity of $Q_1(\cdot, \xi_1)$, we have

$$Q_1(x, \xi_1) \leq Q_1(\tilde{x}, \xi_1) + L\|x - \tilde{x}\| \leq Q_1(\tilde{x}, \xi_1) + L_0,$$

where L is the Lipschitz constant for $Q_1(\cdot, \xi_1)$ and the constant L_0 results from the compactness of $\mathcal{D}$. Thus, $Q_1(x, \xi_1)$ has an integrable upper bound. On the other hand, for a point $x \in \mathcal{X}_1$, a multiplier $\pi(\xi_1)$ exists such that

$$\begin{aligned} Q_1(x, \xi_1) &\geq \varphi(\tilde{y}(\xi_1), \xi_1) + \pi^\top (h(\xi_1) - T(\xi_1)x - W\tilde{y}(\xi_1)) \\ &\geq f_2(\tilde{y}(\xi_1), \xi_1) + Q_2(\tilde{y}(\xi_1), \xi_{[2]}) + \pi^\top (h(\xi_1) - T(\xi_1)x - W\tilde{y}(\xi_1)) \\ &\geq f_2(\tilde{y}(\xi_1), \xi_1) - \mathbb{E}_1[\kappa(\xi_2)] + \pi^\top T(\xi_1)(\tilde{x} - x) \\ &\geq f_2(\tilde{y}(\xi_1), \xi_1) - \mathbb{E}_1[\kappa(\xi_2)] - \|\pi\| \|\tilde{x} - x\| \|T(\xi_1)\|. \end{aligned}$$

Due to the integrability requirement in (A1) and the compactness assumptions in (A2), we conclude that $Q_1(x, \xi_1)$ has an integrable lower bound. Hence, $Q_1(x, \xi_1)$ is integrable. This completes the proof. $\square$

4. Relations between order-constraint problems and expected utility problems. The following lemma is a counterpart of a result shown in [5]; we also prove slightly more general form of the statement.

Lemma 4.1. *Let $I = [a, b] \subseteq [-\infty, \infty]$. There is an one-to-one correspondence between measures $\nu \in \mathcal{M}(I)$ and utility functions $u \in \mathcal{U}(I)$ so that*

$$(24) \quad \int_I \mathbb{E}[(X - \eta)_+] \nu(d\eta) = \mathbb{E}[u(X)]$$

holds for all integrable random variables X .

Proof. Given a measure $\mu \in \mathcal{M}(I)$, we define a function $u : \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$u(t) = \begin{cases} \int_a^t \mu([a, \tau]) d\tau & t > a, \\ 0 & t \leq a. \end{cases}$$

Since $\mu \geq 0$, the function $\mu([a, \cdot])$ is nonnegative and non-decreasing, which implies that $u(\cdot)$ is nondecreasing and convex.

For an integrable random variable X and a positive constant M , we have

$$\begin{aligned}
 \int_a^b H_X(\eta) d\mu(\eta) &= \int_a^b \int_\eta^M \bar{F}_X(t) dt d\mu(\eta) + H_X(M)\mu([a, b]) \\
 &= \int_a^M \int_a^t d\mu(\eta) \bar{F}_X(t) dt + H_X(M)\mu([a, b]) \\
 (25) \qquad &= \int_a^M \mu([a, t]) \bar{F}_X(t) dt + H_X(M)\mu([a, b]).
 \end{aligned}$$

We can rewrite (25) as

$$\int_a^b H_X(\eta) d\mu(\eta) = \int_a^M \bar{F}_X(t) du(t) + H_X(M)\mu([a, b]).$$

Notice that $u(\cdot)$ is absolutely continuous and $\bar{F}_X(t)$ is monotonic. Thus, we can integrate by parts to obtain

$$\begin{aligned}
 \int_a^M \bar{F}_X(t) du(t) &= \int_a^M 1 - F_X(t) du(t) = u(M) - u(a) - \int_a^M F_X(t) du(t) \\
 &= u(M) - F_X(M)u(M) + F_X(a)u(a) + \int_a^M u(t) dF_X(t) \\
 &= P(X \geq M)u(M) + \mathbb{E}[u(X)] - \mathbb{E}[u(X)\mathbf{1}_{\{X \geq M\}}].
 \end{aligned}$$

The last two equations together yield

$$\begin{aligned}
 (26) \qquad \int_a^b H_X(\eta) d\mu(\eta) &= \mathbb{E}[u(X)] + u(M) - P(X \leq M)u(M) \\
 &\quad - \mathbb{E}[u(X)\mathbf{1}_{\{X \geq M\}}] + H_X(M)\mu([a, b]).
 \end{aligned}$$

Let $M \rightarrow \infty$. Since $\mathbb{E}|X| < \infty$ and the derivative of $u(t)$ is bounded as $t \rightarrow \infty$, we obtain

$$\lim_{M \rightarrow \infty} \mathbb{E}[u(X)\mathbf{1}_{\{X \geq M\}}] = 0.$$

Using the monotonicity of u we also have

$$u(M)P[X \geq M] \leq \mathbb{E}[u(X)\mathbf{1}_{\{X \geq M\}}] \rightarrow 0, \text{ as } M \rightarrow \infty.$$

Consequently, passing to the limit in (26), we obtain the representation in (24).

Now, let a function $u \in \mathcal{U}(I)$ be given. Then the right derivative of u ,

$$u'_+(t) = \lim_{\tau \downarrow t} [u(t) - u(\tau)] / (t - \tau),$$

is well-defined, nondecreasing and continuous from the right. Theorem 12.4 of [2] implies that a unique regular nonnegative measure μ exists such that

$$\mu([a, t]) = u'_+(t).$$

Hence, the stated correspondence is established. $\square$

We now show optimality conditions for the two-stage problem, which exhibit its relation to models with utility functions.

Theorem 4.2. *Assume conditions (A1) – (A3). A feasible policy consisting of a point $\hat{x} \in \mathcal{X}_1$ and a measurable selection $\hat{y} \in \mathcal{X}_2(\hat{x}, \xi_1)$ is optimal for problem (17)–(22) if and only if functions $\hat{u} \in \mathcal{U}(I)$ and $\hat{u}_{\xi_1} \in \mathcal{U}(I(\xi_1))$ exist such that the policy is also optimal for the following problem*

$$(27) \quad \begin{aligned} & \min_x f_1(x) + \mathbb{E}[\tilde{Q}_1(x, \xi_1)] + \mathbb{E}[\hat{u}(\tilde{Q}_1(x, \xi_1))] - \mathbb{E}[\hat{u}(V_1 + \mathbb{E}_1[V_2])] \\ & \text{s.t. } x \in \mathcal{X}_1, \end{aligned}$$

where $\tilde{Q}_1(x, \xi_1)$ is the optimal value of the problem

$$(28) \quad \begin{aligned} & \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \mathbb{E}_1[\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]} | \xi_1))] \\ & - \mathbb{E}_1[\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)] \quad \text{s.t. } y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

and, additionally,

$$(29) \quad \begin{aligned} \mathbb{E}_1[\hat{u}_{\xi_1}(f_2(\hat{x}, \xi_1) + Q_2(\hat{y}, \xi_{[2]} | \xi_1))] &= \mathbb{E}_1[\hat{u}_{\xi_1}((V_1 + V_2) | \xi_1)] \quad \mathcal{F}_1 - a.s. \\ \mathbb{E}[\hat{u}(\tilde{Q}_1(\hat{x}, \xi_1))] &= \mathbb{E}[\hat{u}(V_1 + \mathbb{E}_1[V_2])]. \end{aligned}$$

Conversely, if an optimal policy $(\hat{x}, \hat{y})$ for problem (27)–(28) is feasible for problem (17)–(22) and satisfies (29), then $(\hat{x}, \hat{y})$ is optimal for problem (17)–(22).

Proof. We define the operators $\Gamma(y, \xi_1) : \mathcal{X}_1 \rightarrow \mathcal{C}(I_{\xi_1})$ by setting for any $\eta \in I_{\xi_1}$

$$\Gamma(y, \xi_1)(\eta) = \mathbb{E}_1[(f_2(y, \xi_1) + Q_2(y, \xi_{[2]} | \xi_1 - \eta)_+ - ((V_1 + V_2) | \xi_1 - \eta)_+].$$

The second-stage problem is reformulated in the following abstract way:

$$(30) \quad \begin{aligned} & \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) \\ & \text{s.t. } \Gamma(y, \xi_1) \in \mathcal{K}(\xi_1), \\ & y \in \mathcal{X}_2(x, \xi_1) \end{aligned}$$

Here $\mathcal{K}(\xi_1)$ denotes the cone of the non-positive continuous functions defined on I_{ξ_1} .

It follows as in the proof of Theorem 3.2 that the operator Γ is convex with respect to the cone $\mathcal{K}(\xi_1)$, that is, for any y^1, y^2 in $\mathcal{X}_2(x, \xi_1)$ and all $\lambda \in [0, 1]$,

$$\Gamma(\lambda y^1 + (1 - \lambda)y^2, \xi_1) - [\lambda \Gamma(y^1, \xi_1) + (1 - \lambda)\Gamma(y^2, \xi_1)] \in \mathcal{K}(\xi_1).$$

Hence, for every ξ_1 , problem (30) is a convex optimization problem with an operator constraint. The dual space to the space of continuous functions is the space of all countably additive regular measures on $I(\xi_1)$ due to the Riesz representation theorem. The following partial Lagrangian is formulated:

$$(31) \quad \Lambda(y, \mu, \xi_1) = f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]}) + \int_{I(\xi_1)} \Gamma(y, \xi_1)(\eta) \mu(d\eta, \xi_1)$$

with $\mu(\xi_1) \in \mathcal{M}[I(\xi_1)]$. Note that our constraint qualification condition is equivalent to Robinson's condition as shown in [6]. Using the optimality conditions in Banach spaces, we infer that at the optimal solution $\hat{y}$, measures $\hat{\mu}(\xi_1)$ exist such that

$$(32) \quad \Lambda(\hat{y}, \hat{\mu}(\xi_1), \xi_1) = \min_{y \in \mathcal{X}_2(x, \xi_1)} \Lambda(y, \hat{\mu}(\xi_1), \xi_1)$$

$$(33) \quad \int_{I_{\xi_1}} \Gamma(\hat{y}, \xi_1)(\eta) \hat{\mu}(d\eta, \xi_1) = 0.$$

Now, we apply Lemma 4.1 to the measure $\hat{\mu}(\xi_1)$ and transform the term

$$\int_{I_{\xi_1}} \Gamma(\hat{y}, \xi_1)(\eta) \hat{\mu}(d\eta, \xi_1) = \mathbb{E}_1[\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}|\xi_1) - \hat{u}_{\xi_1}((V_1 + V_2)|\xi_1))]$$

We obtain that problem (28) is a Lagrangian-like relaxation of problem (20)–(22). Let $\tilde{Q}_1(x, \xi_1)$ be the optimal value of (28). Due to the convexity assumptions, we obtain that problem (20)–(22) and problem (28) are equivalent and their optimal value coincide. We consider now the first-stage problem (17)–(19). In exactly the same way, we show that a utility function $\hat{u} \in \mathcal{U}(I_0)$ exists, so that that problem (17)–(19) is equivalent to

$$(34) \quad \min_x f_1(x) + \mathbb{E}[Q_1(x, \xi_1)] + \mathbb{E}[\hat{u}(Q_1(x_1, \xi_1))] - \mathbb{E}[\hat{u}(V_1 + \mathbb{E}_1[V_2])]$$

$$(35) \quad \text{s.t. } x \in \mathcal{X}_1.$$

This completes the proof. $\square$

We could substitute $\tilde{Q}_1(x, \xi_1)$ for $Q_1(x, \xi_1)$ in (34) to obtain a more explicit form of the objective function of problem (27). We obtain,

$$\begin{aligned} f_1(x) + \mathcal{Q}_1(y, \xi_{[2]}) + \mathbb{E} \left[\mathbb{E}_1 [\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) | \xi_1)] - \mathbb{E}_1 [\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)] \right] \\ + \mathbb{E} \left[\hat{u}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \mathbb{E}_1 [\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) | \xi_1)] \right. \\ \left. - \mathbb{E}_1 [\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)]) \right] - \mathbb{E} [\hat{u}(V_1 + \mathbb{E}_1[V_2])] \end{aligned}$$

The optimal solution of the first-stage problem will not change, if we optimize the following objective at the first stage:

$$\begin{aligned} f_1(x) + \mathcal{Q}_1(y, \xi_{[2]}) + \mathbb{E} \left[\mathbb{E}_1 [\hat{u}_{\xi_1}(Q_1(x, \xi_1) | \xi_1)] - \mathbb{E}_1 [\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)] \right] \\ + \mathbb{E} \left[\hat{u} \left(Q_1(x, \xi_1) + \mathbb{E}_1 [\hat{u}_{\xi_1}(Q_1(x, \xi_1) | \xi_1)] - \mathbb{E}_1 [\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)] \right) \right]. \end{aligned}$$

Similarly, we could drop the term $\mathbb{E}_1 [\hat{u}_{\xi_1}(V_1 + V_2 | \xi_1)]$ from the second-stage problem without changing the optimal solution $\hat{y}(\xi)$. Hence, using the modified objective of the second stage and with slight abuse of notation, we could consider the following objective at the first stage:

$$\begin{aligned} f_1(x) + \mathcal{Q}_1(y, \xi_{[2]}) + \mathbb{E} \left[\mathbb{E}_1 [\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) | \xi_1)] \right] \\ + \mathbb{E} \left[\hat{u}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \mathbb{E}_1 [\hat{u}_{\xi_1}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) | \xi_1)]) \right]. \end{aligned}$$

As we see, this modification brings us closer to the utility formulation problem (4), although the two problems are not equivalent. Notice that we have a composition of utility functions, which is in line with the theory on sequential preferences but it is impossible to construct upfront; in our case it is implied by the benchmark process.

5. Inverse formulation and relations to measures of risk.

Now we consider the inverse formulation of the increasing convex order for the random sequence $V = (V_1, V_2)$, which uses the upper Lorenz function. For clarity, we shall denote its value at $p \in [0, 1]$ for a random variable Z as $\bar{L}(Z, p)$. Let now $J_0 \subset (0, 1)$ be a closed interval and $J(\xi_1)$ be a $\mathcal{F}_1$ -measurable multifunction with closed images in $(0, 1)$.

$$(36) \quad \min_x f_1(x) + \mathbb{E}[\Phi(x, \xi_1)]$$

$$(37) \quad \text{s.t. } \bar{L}(\Phi(x_1, \xi_1), p) \leq \bar{L}(V_1 + \mathbb{E}_1[V_2], p), \quad p \in J_0;$$

$$(38) \quad x \in \mathcal{X}_1.$$

with $\Phi(x, \xi_1)$ being the optimal value of the problem

$$(39) \quad \min_y f_2(y, \xi_1) + \mathbb{E}[Q_2(y, \xi_{[2]}|\xi_1]$$

$$(40) \quad \text{s.t. } \bar{L}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}|\xi_1, p) \leq \bar{L}(V_1(\xi_1) + V_2|\xi_1, p) \quad p \in J(\xi_1);$$

$$(41) \quad y \in \mathcal{X}_2(x, \xi_1).$$

and $Q_2(y, \xi_{[2]}) = f_3(B(\xi_2)y, \xi_2)$.

We augment assumption (A3) to the following:

- (A4) One of the following is true: (a) The probability space $(\Omega, \mathcal{F}, P)$ is finite and the sets $\mathcal{D}, \mathcal{Y}(\xi_1)$ are polyhedral; (b) A decision policy $(\tilde{x}, \tilde{y}) \in \mathcal{X}_1 \times \mathcal{X}_2(\tilde{x}, \xi_1)$ exist such that the following holds

$$\begin{aligned} \sup_{p \in J(\xi_1)} \bar{L}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}|\xi_1, p) - \bar{L}(V_1(\xi_1) + V_2|\xi_1, p) &< 0 \quad \text{for a.a. } \xi_1; \\ \sup_{\eta \in J_0} \bar{L}(Q_1(x_1, \xi_1), p) - \bar{L}(V_1 + \mathbb{E}_1[V_2], p) &< 0, \end{aligned}$$

Lemma 5.1. *For every $p \in (0, 1)$, the function $\bar{L}(\cdot, p)$ is convex, monotonic (w.r. to the almost sure order), and positively homogeneous on $\mathcal{L}_1(\Omega, \mathcal{F}, P)$.*

Proof. It is proved in [7, Lemma 2] that for every $p \in (0, 1)$ the mapping $X \mapsto L_X(p)$ is concave and positively homogeneous on $\mathcal{L}_1(\Omega, \mathcal{F}, P)$. Thus, $X \mapsto \bar{L}(X, p) = \mathbb{E}[X] - L_X(p)$ is convex and positively homogeneous on $\mathcal{L}_1(\Omega, \mathcal{F}, P)$ for every $p \in (0, 1)$. It follows directly from the definition that $\bar{L}(X, p) \leq \bar{L}(Y, p)$ whenever $X \leq Y$ a.s. $\square$

Lemma 5.2. *Assume (A1), (A2) and (A4). Then problem (36)–(38) and problems (39)–(41) (for each fixed ξ_1) are convex problems. Furthermore, the functions $\Phi(\cdot, \xi_1)$ and $\mathbb{E}[\Phi(\cdot, \xi_1)]$, $Q_2(\cdot, \xi_{[2]})$, $\bar{Q}_2(\cdot, \xi_{[2]})$ are convex and finite-valued for all feasible points.*

Proof. The functions $Q_2(\cdot, \xi_{[2]})$, $\bar{Q}_2(\cdot, \xi_{[2]})$ are the same as in the direct formulation and, therefore, they are convex and finite-valued by virtue of Theorem 3.2. Due to Lemma 5.1, the feasible set determined by constraint (37) is convex and so are the sets determined by constraint (40) for each ξ_1 . With this observations in mind, the proof of the statement follows exactly the same line of arguments as in Theorem 3.2. $\square$

We now establish a form of necessary conditions of optimality for this formulation of the extended two-stage problem, which demonstrate the relation of the problem to models with measures of risk.

Theorem 5.3. *Assume conditions (A1), (A2), and (A4). If feasible policy consisting of a point $\hat{x} \in \mathcal{X}_1$ and a measurable selection $\hat{y} \in \mathcal{X}_2(\hat{x}, \xi_1)$ is optimal for problem (36)–(41), then $\kappa_0 \geq 0$, and $\kappa(\xi_1) \geq 0$ ($\mathcal{F}_1$ -a.s.) exist, along with a coherent measure of risk $\hat{\varrho} : \mathcal{L}_1(\Omega, \mathcal{F}_1, P) \rightarrow \mathbb{R}$ and a conditional risk mapping $\varrho_2 : \mathcal{L}_1(\Omega, \mathcal{F}_2, P) \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}_1, P)$ such that the policy is also optimal for the following problem:*

$$(42) \quad \begin{aligned} & \min_x f_1(x) + \mathbb{E}[\Phi_1(x, \xi_1)] + \kappa_0 \hat{\varrho}[\Phi_1(x_1, \xi_1)] - \kappa_0 \hat{\varrho}[V_1 + \mathbb{E}_1[V_2]] \\ & \text{s.t. } x \in \mathcal{X}_1. \end{aligned}$$

where $\Phi_1(x, \xi_1)$ is the optimal value of the problem

$$(43) \quad \begin{aligned} & \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]} + \kappa(\xi_1) \varrho_{\xi_1} [f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] \\ & \quad - \kappa(\xi_1) \varrho_{\xi_1} [V_1(\xi_1) + V_2|\xi_1] \text{ s.t. } y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

Additionally,

$$(44) \quad \begin{aligned} & \kappa(\xi_1) \varrho_{\xi_1} [f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] = \kappa(\xi_1) \varrho_{\xi_1} [V_1(\xi_1) + V_2|\xi_1] \quad \mathcal{F}_1\text{-a.s.} \\ & \kappa_0 \hat{\varrho}[\Phi_1(x_1, \xi_1)] = \kappa_0 \hat{\varrho}[V_1 + \mathbb{E}_1[V_2]]. \end{aligned}$$

Proof. We define now the operators $\tilde{\Gamma}(y, \xi_1) : \mathcal{X}_2(x, \xi_1) \rightarrow \mathcal{C}(J_{\xi_1})$ by setting

$$\tilde{\Gamma}(y, \xi_1)(p) = \bar{L}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}|\xi_1, p) - \bar{L}(V_1(\xi_1) + V_2|\xi_1, p))$$

The second-stage problem takes on the form:

$$(45) \quad \begin{aligned} & \min_y f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) \\ & \text{s.t. } \tilde{\Gamma}(y, \xi_1) \in \tilde{\mathcal{K}}(\xi_1), \\ & y \in \mathcal{X}_2(x, \xi_1) \end{aligned}$$

Here $\tilde{\mathcal{K}}(\xi_1)$ is the cone of the non-positive continuous functions defined on J_{ξ_1} .

It follows from Lemma 5.1 that the operator $\tilde{\Gamma}$ is convex with respect to the cone $\tilde{\mathcal{K}}(\xi_1)$. Hence, for every realization of ξ_1 , problem (45) is a convex

optimization problem with an operator constraint. Using the Riesz representation theorem again, we formulate the following partial Lagrangian:

$$(46) \quad \tilde{\Lambda}(y, \nu, \xi_1) = f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]}) + \int_{J(\xi_1)} \tilde{\Gamma}(y, \xi_1)(p) \nu(dp, \xi_1)$$

with $\nu(\xi_1) \in \mathcal{M}[J(\xi_1)]$. Assumption (A4) provides a constraint qualification condition, which is equivalent to Robinson's condition by the same argument as before. The optimality conditions in Banach spaces imply that at the optimal solution $\hat{y}$, measures $\hat{\nu}(\xi_1)$ exist such that

$$(47) \quad \hat{\Lambda}(\hat{y}, \hat{\nu}(\xi_1), \xi_1) = \min_{y \in \mathcal{X}_2(x, \xi_1)} \Lambda(y, \hat{\nu}(\xi_1), \xi_1)$$

$$(48) \quad \int_{J_{\xi_1}} \tilde{\Gamma}(\hat{y}, \xi_1)(\eta) \hat{\nu}(d\eta, \xi_1) = 0.$$

We extend the measure $\hat{\nu}(\xi_1)$ to the entire interval $[0, 1]$ by setting the measure of any measurable subset $\mathcal{A}$ of $[0, 1]$ to be equal to $\hat{\nu}(\mathcal{A} \cap J(\xi_1), \xi_1)$ or zero if the intersection is empty. Further, we define $\kappa(\xi_1) = \int_0^1 t \hat{\nu}(dt, \xi_1)$ and a measure $\hat{\mu}(\xi_1)$ such that

$$\hat{\mu}(dp, \xi_1) = \begin{cases} \kappa^{-1} p \hat{\nu}(dp, \xi_1) & \text{if } \kappa > 0 \\ dp & \text{if } \kappa = 0. \end{cases}$$

Observe that μ is a probability measure. Indeed, it is non-negative because $\hat{\nu}(\xi_1)$ is nonnegative and

$$\int_0^1 \kappa^{-1}(\xi_1) p \hat{\nu}(dp, \xi_1) = \kappa^{-1}(\xi_1) \int_0^1 p \hat{\nu}(dp, \xi_1) = 1.$$

For any random variable $Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$, we define the spectral measure of risk

$$\varrho_{\xi_1}[Z] = \int_0^1 \text{AVaR}_p[Z] \hat{\mu}(dp, \xi_1).$$

The following transformation holds:

$$(49) \quad \begin{aligned} \int_0^1 \bar{L}(Z, p) \hat{\nu}(dp, \xi_1) &= \int_0^1 \frac{1}{p} \bar{L}(Z, p) p \hat{\nu}(dp, \xi_1) \\ &= \kappa(\xi_1) \int_0^1 \text{AVaR}_p[Z] \hat{\mu}(dp, \xi_1) = \kappa(\xi_1) \varrho_{\xi_1}[Z]. \end{aligned}$$

Substituting this formula into (48), we obtain that the complementarity condition can be re-written as

$$\kappa(\xi_1)\varrho_{\xi_1}[f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] = \kappa(\xi_1)\varrho_{\xi_1}[V_1(\xi_1) + V_2|\xi_1]$$

Substituting (49) into (46) yields that $\hat{y}$ is also an optimal solution to the problem with the following objective

$$f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \kappa(\xi_1)\left(\varrho_{\xi_1}[f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] - \varrho_{\xi_1}[V_1(\xi_1) + V_2|\xi_1]\right)$$

subject to the constraint (41). Due to the convexity assumptions, we obtain that problem (39)-(41) and problem (43) are equivalent and their optimal value satisfy

$$\tilde{\Phi}(x, \xi_1) = \Phi(x, \xi_1) + \kappa(\xi_1)\varrho_{\xi_1}[V_1(\xi_1) + V_2|\xi_1]$$

We turn now to the first-stage problem (36)–(38). In exactly the same way, we show that a constant κ_0 and a spectral risk measure $\hat{\varrho}$ exist such that that problem is equivalent to

$$\min_{x \in \mathcal{X}_1} f_1(x) + \mathbb{E}[\Phi_1(x, \xi_1)] + \kappa_0\left(\hat{\varrho}[\Phi_1(x, \xi_1)] - \hat{\varrho}[V_1 + \mathbb{E}_1[V_2]]\right)$$

This completes the proof. $\square$

Substituting the equivalent representation of Φ , we obtain that problem (36)–(38) is equivalent to

$$\begin{aligned} \min_x \quad & f_1(x) + \mathbb{E}[\Phi_1(x, \xi_1)] + \kappa_0\hat{\varrho}\left[f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) \right. \\ & \left. + \kappa(\xi_1)\varrho_{\xi_1}[f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] - \kappa(\xi_1)\varrho_{\xi_1}[V_1(\xi_1) + V_2|\xi_1]\right) \\ & \left. - \kappa_0\hat{\varrho}[V_1 + \mathbb{E}_1[V_2]]\right) \\ \text{s.t. } & x \in \mathcal{X}_1, \quad y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

However, we may observe that $\hat{y}$ is also an optimal solution to the following problem:

$$(50) \quad \begin{aligned} \tilde{\Phi}(x, \xi_1) = \min_y \quad & f_2(y, \xi_1) + Q_2(y, \xi_{[2]}) + \kappa(\xi_1)\varrho_{\xi_1}[f_2(y, \xi_1) + Q_2(y, \xi_{[2]})] \\ \text{s.t. } & y \in \mathcal{X}_2(x, \xi_1). \end{aligned}$$

Substituting now the expression of $\tilde{\Phi}$ as an argument of $\hat{\varrho}$ and dropping the term $\kappa_0\hat{\varrho}[V_1 + \mathbb{E}_1[V_2]]$, which does not depend on x , we obtain a problem that has a similar structure to that of problem (7) formulated in the introduction. The two problems are, however, not equivalent.

6. Relations to distortions. Let us denote $\mathcal{W}$ the set of all convex nondecreasing functions $w : [0, 1] \rightarrow \mathbb{R}$, which are subdifferentiable at 1. The following characterization is shown in [7]. For two random variables $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ the relation $X \succeq_{(1)} Y$ holds if and only if

$$(51) \quad \int_0^1 F_X^{-1}(p) dw(p) \geq \int_0^1 F_Y^{-1}(p) dw(p),$$

for all continuous nondecreasing functions $w : [0, 1] \rightarrow \mathbb{R}$. The relation $X \succeq_{(2)} Y$ holds if and only if (51) is satisfied for all functions $w \in \mathcal{W}$. Additionally, a one-to-one correspondence between non-negative measures on $[0, 1]$ and functions $w \in \mathcal{W}$ is established.

Consider $[\alpha, \beta] \subseteq [0, 1], \alpha < \beta$, and the set $\mathcal{W}([\alpha, \beta])$ of convex and non-decreasing functions $w : [0, 1] \rightarrow \mathbb{R}$ such that $w(p) = 0$ for all $p \in [0, \alpha]$ and $w(p) = w(\beta) + c(p - \beta)$, with some $c > 0$, for all $p \in [\beta, 1]$. It is evident that $\mathcal{W}([\alpha, \beta])$ is a convex cone in $\mathcal{C}([0, 1])$ and $\mathcal{W}([\alpha, \beta]) \subset \mathcal{W}$.

In a way analogous to [7], we can show the following lemma, which is a counterpart of Lemma 4.1. We shall provide the proof for completeness and for the sake of convenience of the reader.

Lemma 6.1. *Let $J = [\alpha, \beta] \subseteq [0, 1]$ be given. There is a one-to-one correspondence between measures $\nu \in \mathcal{M}(J)$ and distortion functions $w \in \mathcal{W}(J)$ so that*

$$(52) \quad \int_{\alpha}^{\beta} \bar{L}_X(p) d\nu(p) = \int_0^1 F_X^{-1}(p) dw(p)$$

holds for all integrable random variables X .

Proof. Given a measure $\nu \in \mathcal{M}(J)$, we can extend it to the interval $[0, 1]$ by assigning zero mass to sets not interesting $[\alpha, \beta]$. The extension is denoted $\bar{\nu}$. The function $\xi_{\nu}(\tau) = \bar{\nu}([0, 1 - \tau])$ and the function $w : [0, 1] \rightarrow \mathbb{R}$ is defined as follows:

$$w(p) = \int_{1-p}^1 \xi_{\nu}(\tau) d\tau = \int_0^p \xi_{\nu}(1-t) dt.$$

The derivative $w'(p) = \xi_{\nu}(1-p) = \bar{\nu}([0, p])$. The property $\nu \geq 0$ entails that the function $t \rightarrow \xi_{\nu}(1-t)$ is nonnegative and nondecreasing, which implies that $w(\cdot)$ is nondecreasing and convex. The left derivative $w'_-(1) = \bar{\nu}([0, 1])$ is finite,

implying that w is subdifferentiable at $p = 1$ and, thus, $w \in \mathcal{W}$. If $\beta < 1$, then the derivative $w'(p) = \nu(J)$ for all $p \in [\beta, 1)$. If $\alpha > 0$, then the derivative $w'(p) = 0$ for all $p \in (0, \alpha)$ and $w(\alpha) = \int_0^\alpha \bar{\nu}([0, t]) d\tau = 0$. Hence, $w \in \mathcal{W}(J)$.

For the opposite direction, let $w \in \mathcal{W}(J)$. Hence, the right derivative of w ,

$$w'_+(p) = \lim_{\tau \downarrow p} [w(\tau) - w(p)] / (\tau - p), \quad p \in [0, 1),$$

is well-defined, nondecreasing and right-continuous. Using [2, Theorem 12.4] again, we infer the existence of a unique regular nonnegative measure ν on $[0, 1]$ satisfying

$$\nu([0, p]) = \begin{cases} w'_+(p) & \text{for } p \in [0, 1), \\ w'_-(1) & \text{for } p = 1. \end{cases}$$

If $w \in \mathcal{W}(J)$, $J = [\alpha, \beta]$ and $\alpha > 0$, then $w'_+(p) = 0$ for $p \in [0, \alpha)$. If $\beta < 1$, then $\nu([0, p]) = c$ remains constant for $p \in [\beta, 1]$, which implies that the measure ν has support on $J = [\alpha, \beta]$.

We now change the order of integration in the integrals in (52). We obtain

$$\begin{aligned} \int_\alpha^\beta \bar{L}_X(p) d\nu(p) &= \int_0^1 \bar{L}_X(p) d\bar{\nu}(p) = \int_0^1 \int_t^1 F_X^{-1}(p) dp d\bar{\nu}(t) = \int_0^1 \int_0^p d\nu(t) F_X^{-1}(p) dp \\ &= \int_0^1 \nu([0, p]) F_X^{-1}(p) dp = \int_0^1 F_X^{-1}(p) dw(p). \end{aligned}$$

□

Lemma 6.1 implies the following characterization of the (relaxed) increasing convex order via distortions.

Corollary 6.2. *For two integrable random variables X and Y , the relation $X \preceq_{ic} Y$ holds on $J \subset [0, 1]$ if and only if*

$$(53) \quad \int_0^1 F_X^{-1}(p) dw(p) \leq \int_0^1 F_Y^{-1}(p) dw(p),$$

for all distortion functions $w \in \mathcal{W}(J)$.

Let us return to problem (36)-(41). We have a new equivalent problem featuring distortion functions.

Theorem 6.3. *Assume conditions (A1), (A2), and (A4). If feasible policy consisting of a point $\hat{x} \in \mathcal{X}_1$ and a measurable selection $\hat{y} \in \mathcal{X}_2(\hat{x}, \xi_1)$ is optimal for problem (36)–(41), then distortion functions $\hat{w}_0 \in \mathcal{W}(J)$, and $\hat{w}(\xi_1) \in \mathcal{W}(J_{\xi_1})$ for $\mathcal{F}_1$ -a.a. ξ_1 exist, such that the policy is also optimal for the following problem*

$$(54) \quad \min_x f_1(x) + \mathbb{E}[\Psi_1(x, \xi_1)] + \int_0^1 F_{\Psi_1(x, \xi_1)}^{-1}(p) d\hat{w}_0(p) - \int_0^1 F_{V_1 + \mathbb{E}_1[V_2]}^{-1}(p) d\hat{w}_0(p) \quad \text{s.t. } x \in \mathcal{X}_1.$$

where $\Psi_1(x, \xi_1)$ is the optimal value of the problem

$$(55) \quad \min_y f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]}) + \int_0^1 F_{\Psi_2(y, \xi_{[2]})|\xi_1}^{-1}(p) d\hat{w}(\xi_1, p) - \int_0^1 F_{V_1(\xi_1) + V_2|\xi_1}^{-1}(p) d\hat{w}(\xi_1, p) \quad \text{s.t. } y \in \mathcal{X}_2(x, \xi_1).$$

with $\Psi_2(y, \xi_{[2]}) = f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]})$. Additionally,

$$(56) \quad \int_0^1 F_{\Psi_1(\hat{x}, \xi_1)}^{-1}(p) d\hat{w}_0(p) = \int_0^1 F_{V_1 + \mathbb{E}_1[V_2]}^{-1}(p) d\hat{w}_0(p), \\ \int_0^1 F_{\Psi_2(y, \xi_{[2]})}^{-1}(p) d\hat{w}(\xi_1, p) = \int_0^1 F_{V_1(\xi_1) + V_2|\xi_1}^{-1}(p) d\hat{w}(\xi_1, p) \quad \mathcal{F}_1\text{-a.s.}$$

Conversely, if a policy $\hat{x} \in \mathcal{X}_1$ and a measurable selection $\hat{y} \in \mathcal{X}_2(\hat{x}, \xi_1)$ is optimal for problem (54)–(55) for some distortion functions $\hat{w}_0 \in \mathcal{W}(J)$, and $\hat{w}(\xi_1) \in \mathcal{W}(J_{\xi_1})$ for $\mathcal{F}_1$ -a.a. ξ_1 , conditions (56), (37), and (40) are satisfied, then that policy is also optimal for problem (36)–(41).

Proof. In the proof of Theorem 5.3, we have obtained that problem (39)–(41) satisfy the following abstract necessary and sufficient optimality conditions.

$$\hat{\Lambda}(\hat{y}, \hat{\nu}(\xi_1), \xi_1) = \min_{y \in \mathcal{X}_2(x, \xi_1)} \Lambda(y, \hat{\nu}(\xi_1), \xi_1) \\ \int_{J_{\xi_1}} \tilde{\Gamma}(\hat{y}, \xi_1)(\eta) \hat{\nu}(d\eta, \xi_1) = 0,$$

where the abstract Lagrangian is given by

$$\tilde{\Lambda}(y, \hat{\nu}, \xi_1) = f_2(y, \xi_1) + \mathcal{Q}_2(y, \xi_{[2]}) + \int_{J(\xi_1)} \tilde{\Gamma}(y, \xi_1)(p) \hat{\nu}(dp, \xi_1),$$

$\hat{\nu}(\xi_1) \in \mathcal{M}[J(\xi_1)]$ and $\tilde{\Gamma}(y, \xi_1)(p) = \bar{L}(f_2(y, \xi_1) + Q_2(y, \xi_{[2]}|\xi_1, p) - \bar{L}(V_1(\xi_1) + V_2|\xi_1, p)$.

We use Lemma 6.1 to associate a distortion function $\hat{w}(\xi_1, \cdot)$ with the measure $\hat{\nu}$ and to transform the integral term involving $\tilde{\Gamma}$ as follows:

$$\int_{J_{\xi_1}} \tilde{\Gamma}(\hat{y}, \xi_1)(p) \hat{\nu}(dp, \xi_1) = \int_0^1 F_{\Psi_2(y, \xi_{[2]})}^{-1}(p) d\hat{w}(\xi_1, p) - \int_0^1 F_{V_1(\xi_1) + V_2|\xi_1}^{-1} d\hat{w}(\xi_1, p)$$

Substituting this expression into (47) and (48) and using the convexity assumptions, we obtain that problem (39)–(41) and problem (55) are equivalent and their optimal values coincide. We proceed with the first-stage problem (36)–(38) in exactly the same way. $\square$

In conclusion, we have establish relations between the optimization model with stochastic-order constraints to all popular risk averse decision problems. The characteristic observation is that those models use transformations of the random outcomes, which are incorporated into the objective function. In this way the objective contains an aggregate representation of the distribution. The decision problem with stochastic-order constraints allows for an explicit shaping of the distribution. On the other hand, the disadvantage of this model arises in situations when no natural benchmark distribution is available.

REFERENCES

- [1] P. ARTZNER, F. DELBAEN, J.-M. EBER, D. HEATH. Coherent measures of risk. *Math. Finance* **9**, 3 (1999), 203–228.
- [2] P. BILLINGSLEY. Probability and Measure. New York, John Wiley & Sons, Inc., 1995.
- [3] G. CHOQUET. Theory of capacities. *Annales de l'Institut Fourier Grenoble* **5** (1953), 131–295.
- [4] D. DENTCHEVA, G. MARTINEZ. Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse. *European J. Oper. Res.* **219**, 1 (2012), 1–8.
- [5] D. DENTCHEVA, A. RUSZCZYŃSKI. Optimization with stochastic dominance constraints. *SIAM J. Optim.* **14**, 2 (2003), 548–566.

- [6] D. DENTCHEVA, A. RUSZCZYŃSKI. Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. *Math. Program.* **99**, 2 (2004), 329–350.
- [7] D. DENTCHEVA, A. RUSZCZYŃSKI. Inverse stochastic dominance constraints and rank dependent expected utility theory. *Math. Program. Ser. B* **108**, 2–3 (2006), 297–311.
- [8] D. DENTCHEVA, A. RUSZCZYŃSKI. Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization. *Pac. J. Optim.* **4**, 3 (2008), 433–446.
- [9] D. DENTCHEVA, A. RUSZCZYŃSKI. Risk-averse portfolio optimization via stochastic dominance constraints. In: *Handbook of Quantitative Finance and Risk Management* (eds Cheng-Few Lee, Alice C. Lee, John Lee). New York, Springer, 2010, 247–258.
- [10] D. DENTCHEVA, A. RUSZCZYŃSKI. Stochastic dynamic optimization with discounted stochastic dominance constraints. *SIAM J. Control and Optimization*, **47** (2008), 2540–2556.
- [11] D. DENTCHEVA, A. RUSZCZYŃSKI. Composite semi-infinite optimization. *Control Cybernet.* **36**, 3 (2007), 633–646.
- [12] D. DENTCHEVA, M. YE, Y. YI. Risk-averse sequential decision problems with time-consistent stochastic dominance constraints. 2022 IEEE 61st IEEE Conference on Decision and Control (CDC), Cancun, Mexico, 2022, 3605–3610, doi: 10.1109/CDC51059.2022.9993044.
- [13] L. F. ESCUDERO, M. A. GARÍN, M. MERINO, G. PÉREZ. On time stochastic dominance induced by mixed integer-linear recourse in multistage stochastic programs. *European J. Oper. Res.* **249**, 1 (2016), 164–176.
- [14] L. F. ESCUDERO, J. F. MONGE, D. R. MORALES. On the time-consistent stochastic dominance risk averse measure for tactical supply chain planning under uncertainty. *Comput. Oper. Res.*, **100** (2018), 270–286.
- [15] R. GOLLMER, F. NEISE, R. SCHULTZ. Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. *SIAM J. Optimization* **19**, 2 (2008), 552–571.

- [16] W. B. HASKELL, R. JAIN. Stochastic dominance-constrained Markov decision processes. *SIAM Journal on Control and Optimization*, **51** (2013), 273–303.
- [17] J. HADAR, W.R. RUSSELL. Rules for ordering uncertain prospects. *The American Economic Review* **59**, 1 (1969), 25–34.
- [18] S. KUSUOKA. On law invariant coherent risk measures. *Adv. Math. Econ.* **3** (2001), 83–95.
- [19] M. O. LORENZ. Methods of measuring concentration of wealth. *Publications of the American Statistical Association* **9**, 70 (1905), 209–219, <http://dx.doi.org/10.2307/2276207>.
- [20] A. MÜLLER, D. STOYAN. Comparison methods for stochastic models and risks. Wiley Ser. Prob. Stat., Chichester, John Wiley & Sons, Ltd., 2002.
- [21] A. MÜLLER, M. SCARSINI. Stochastic comparison of random vectors with a common copula. *Math. Oper. Res.* **26**, 4 (2001), 723–740.
- [22] P. MULIERE, M. SCARSINI. A note on stochastic dominance and inequality measures. *J. Econom. Theory* **49**, 2 (1989), 314–323.
- [23] W. OGRYCAK, A. RUSZCZYŃSKI. Dual stochastic dominance and related mean-risk models. *SIAM J. Optim.* **13**, 1 (2002), 60–78.
- [24] J. QUIGGIN. A theory of anticipated utility. *Journal of Economic Behavior and Organization* **3**, 4 (1982), 323–343.
- [25] J. QUIGGIN. Generalized Expected Utility Theory: The Rank-Dependent Expected Utility Model. Dordrecht and London, Kluwer Academic Publishers, 1993.
- [26] J. P. QUIRK, R. SAPOSNIK. Admissibility and measurable utility functions. *Review of Economic Studies* **9**, 2 (1962), 140–146.
- [27] L. RÜSCHENDORF. On conditional stochastic ordering of distributions. *Adv. in Appl. Probab.* **23**, 1 (1991), 46–63.
- [28] A. RUSZCZYŃSKI. Risk-averse dynamic programming for Markov decision processes. *Math. Program.* **125**, 2 (2010), 235–261.
- [29] M. SHAKED, J. G. SHANTHIKUMAR. Stochastic orders and their applications. Probab. Math. Statist., Boston, MA, Academic Press, Inc., 1994.

- [30] A. SHAPIRO, D. DENTCHEVA, A. RUSZCZYŃSKI. Lectures on stochastic programming – modeling and theory. MOS-SIAM Ser. Optim., vol. **28**, 2021.
- [31] D. SCHMEIDLER. Integral representation without additivity. *Proc. Amer. Math. Soc.* **97**, 2 (1986), 255–261.
- [32] D. SCHMEIDLER. Subjective probability and expected utility without additivity. *Econometrica* **57**, 3 (1989), 571–587.
- [33] V. STRASSEN. The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** (1965), 423–439.
- [34] S. S. WANG, V.R. YONG, H. H. PANJER. Axiomatic characterization of insurance prices. *Insurance Math. Econom.* **21**, 2 (1997), 173–183.
- [35] S. S. WANG, V. R. YONG. Ordering risks: expected utility versus Yaari’s dual theory of risk. *Insurance Math. Econom.* **22**, 2 (1998), 145–161.
- [36] W. WHITT. Uniform conditional stochastic order. *J. Appl. Probab.* **17**, 1 (1980), 112–123.
- [37] W. WHITT. Uniform conditional variability ordering of probability distributions. *J. Appl. Probab.* **22**, 3 (1985), 619–633.
- [38] M. E. YAARI. The dual theory of choice under risk. *Econometrica* **55**, 1 (1987), 95–115.

Stevens Institute of Technology

Hoboken, NJ, USA

e-mail: darinka.dentcheva@stevens.edu

Received October 10, 2022

Accepted April 3, 2023