Risk Analysis, Vol. , No. , DOI:

Mean-Deviation Analysis in The Theory of Choice

Bogdan Grechuk, Anton Molyboha, Michael Zabarankin

Mean-deviation analysis, along with the existing theories of coherent risk measures and dual utility, is examined in the context of the theory of choice under uncertainty, which studies rational preference relations for random outcomes based on different sets of axioms such as transitivity, monotonicity, continuity, etc. An axiomatic foundation of the theory of coherent risk measures is obtained as a relaxation of the axioms of the dual utility theory, and a further relaxation of the axioms are shown to lead to the mean-deviation analysis. Paradoxes arising from the sets of axioms corresponding to these theories and their possible resolutions are discussed, and application of the mean-deviation analysis to optimal risk sharing and portfolio selection in the context of rational choice is considered.

KEY WORDS: mean-deviation analysis, theory of choice, deviation measures, coherent risk measures

1. INTRODUCTION

Planning and managing of any real-life project requires understanding not only the *risks* involved but also principles of *rational choice* or a coherent system of attitudes toward those risks. It is well-known that investment projects are subjected to various financial risks, the agricultural sector is often affected by natural hazards, e.g. floods and fires, construction projects should foresee the possibility of floods, hurricanes, earthquakes, volcanic eruptions, etc. While estimating potential losses in each particular project is an important issue, does the analysis of those losses in a corresponding decision problem always follow the principles of rational choice?

As an illustrating example, consider the following decision problem in management of natural resources.[(1)] Suppose there are two dam proposals, both with a project life of N years. Dam A with the net value $P_A > 0$ is to be built on geologically stable ground, whereas dam B with the net value $P_B > P_A$ is to be built on land that has nonzero probability of undergoing an earthquake within the next N years. Without risk consideration, dam B is a clear choice, whereas given potential losses due to an earthquake, the net value of dam B can be modeled as a random variable (r.v.), and the principle of *risk aversion* suggests to select dam B only if $E[P_B] > P_A$, where $E[P_B]$ is the expected value of P_B. But for a real-life project with a finite budget, comparing only expected values is not sufficient. Indeed, even if $E[P_B] > P_A$, there may be nonzero probability that P_B falls below P_A or even zero. If this is the case, the *safety first* principle dictates to select dam A regardless of how small the probability of $P_B < P_A$ is. As a less conservative approach to comparing alternatives, the expected net value for each dam can be decreased by a *safety margin* defined as ρ units of the standard deviation σ, so that if $E[P_B] - \rho\,\sigma(P_B) > P_A$, dam B is selected. However, it is well-known that this measure violates the monotonicity of risk preferences. In other words, there may exist dam proposal C with the net value defined by an

[1]Department of Mathematics, University of Leicester, LE1 7RH, UK, bg83@leicester.ac.uk
[2]Department of Mathematical Sciences, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, amolyboh@stevens.edu
[3]Department of Mathematical Sciences, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, mzabaran@stevens.edu

 0272-4332//0100-0001$22.00/1

r.v. P_C such that $P_C > P_B$ with probability 1 but $E[P_B] - \rho\,\sigma(P_B) > E[P_C] - \rho\,\sigma(P_C)$. This inconsistency can be resolved by measuring the safety margin in units of the *standard lower semideviation* σ_-. But does the latter measure conform to the principles of rational choice for any safety margin ρ? Both measures $E[\cdot] - \rho\,\sigma(\cdot)$ and $E[\cdot] - \rho\,\sigma_-(\cdot)$ are particular cases of a *mean-deviation model* $V(m,d)$, where the mean m and deviation d are combined linearly: $V(m,d) = m - \rho\,d$. Examining a general mean-deviation functional $V(m,d)$, increasing in the first argument and decreasing in the second, for decision making in the context of the theory of choice is the subject of this work. The discussion of the dam selection example will be continued in Section 5.1.

In economics and finance, rational choice is modeled by a set of axioms on preference relations often stated for lotteries (*probability mixtures*) and translated into numerical representations or functionals that combine lottery outcomes and corresponding probabilities in a certain way. The well-known *expected utility theory* [2] (EUT) characterizes a rational preference relation by four axioms: *completeness*, *transitivity*, *continuity*, and *independence*, and represents the preference by a functional linear with respect to the lottery's probabilities. Though being generally accepted as a normative model of rational choice, it is inconsistent with some empirical evidence, known as counter-examples or paradoxes (e.g. Allais' paradox [3,4] that exploits the certainty effect). Attempts to resolve certain EUT's empirical violations have inevitably led to changing the set of underlying axioms and resulted in alternative utility theories such as the *prospect theory*, [3] *dual utility theory* [5] (DUT), *anticipated utility theory*, [6] *rank-dependent expected utility theory*, [7] etc. Those theories, however, are not free from their own paradoxes.

The recently emerged theory of coherent risk measures [8,9] has opened a new perspective in modeling of rational choice under risk. In contrast to the EUT, it postulates axioms on acceptance sets for r.v.'s rather than on preference relations for lotteries. Loosely speaking, "*coherency*" means assigning greater penalties to greater losses. Remarkably, this theory is closely related to the DUT, which keeps three EUT's axioms: completeness, transitivity, and continuity, but replaces independence axiom by the *dual independence axiom*. For coherent risk measures, the dual independence axiom should be changed to some weaker assumptions. In this sense, the theory of coherent risk measures generalizes the DUT, and Yaari's dual utility functional [5] corresponds to the class of so-called *comonotonic* coherent risk measures.

General deviation, as another aspect of an r.v. along with utility and risk, was introduced by Rockafellar et al. [10] to address shortcomings of Markowitz's mean-variance approach. Namely, it preserves four main properties of standard deviation (*nonnegativity*, *positive homogeneity*, *subadditivity*, and *insensitivity to constant shift*) but is not necessarily symmetric with respect to ups and downs of the r.v., e.g. *standard lower semideviation* and *conditional value-at-risk (CVaR) deviation.* [11,12,10] The mean-deviation analysis provides necessary flexibility in expressing individual risk preferences and generalizes the one-fund theorem and capital asset pricing model (CAPM) well-known in the portfolio theory. [13,14,15,16] Rockafellar et al. [10] showed that deviation measures correspond one-to-one with *averse measures of risk* [10] and are required to be *lower range dominated* to reconcile with coherent risk measures.

Thus, the theories of utility, coherent risk measures, and deviation measures translate the uncertainty inherent in an r.v. into, although related but not the same, notions of utility, risk, and deviation.

The aim of this work is two-fold: (i) exploring axiomatic foundations of rational choice for the notions of risk and deviation, often used interchangeably in the finance literature, in connection with the underlined theories; and (ii) identifying a set of axioms for a preference relation, under which risk preferences can be characterized in terms of the mean and a deviation of an r.v. The main idea in (ii) is to relax the assumptions related to the dual independence axiom in the axiomatic foundation for coherent risk measures. Consequently, in the context of the theory of choice, the mean-deviation analysis generalizes both the mean-variance approach and the theory of coherent risk measures. The chart below illustrates these relationships.

$$\begin{array}{ccc} \textit{dual utility theory} & \Rightarrow & \textit{coherent risk measures} \\ & & \Downarrow \\ \textit{mean-variance analysis} & \Rightarrow & \textit{mean-deviation analysis} \end{array}$$

In particular, a mean-deviation functional induces a preference relation satisfying the monotonicity axiom (so-called *monotone preferences*) only if the deviation satisfies certain conditions. For example, those conditions do not hold for standard deviation, which agrees with the well-known fact that mean-variance analysis is inconsistent with monotone preferences.

The rest of the work is organized into four sections. Section 2 reviews classical axioms of rational choice: transitivity, monotonicity, continuity, and risk aversion. Section 3 discusses the axioms leading to the EUT, dual utility theory, and coherent risk measures, reviews paradoxes arising in these theories, and proposes a new set of less restrictive axioms. Section 4 proves the main result of this work: a preference relation satisfies the proposed set of axioms if and only if its numerical representation depends only on the expectation and some general deviation of an r.v. Section 5 applies the mean-deviation analysis to risk sharing and portfolio selection in the context of the theory of choice. Appendices 5.2–5.2 prove some auxiliary results.

2. THEORY OF CHOICE

The theory of choice under uncertainty studies an axiomatic foundation of rational preference relations for random outcomes that may represent, for example, returns of financial instruments, profits of ventures, prices of real estates, net values of construction projects under various risks and natural hazards, etc. In this work, the choice is contemplated among uncertain outcomes modeled as random variables.

Random variables (r.v.'s) are defined on $\Omega = (\Omega, \mathcal{M}, \mathbb{P})$, where Ω denotes the designated space of future states ω, $\mathcal{M}$ is a field of sets in Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{M})$. The probability space Ω is assumed to be *atomless*, i.e., there exists an r.v. with a continuous cumulative distribution function. This implies existence of r.v.'s on Ω with all possible distribution functions.[9]

We restrict our attention to r.v.'s from $\mathcal{L}^p(\Omega) = \mathcal{L}^p(\Omega, \mathcal{M}, P)$, $p \in [1, \infty]$. A relation between any two r.v.'s is understood to hold in the almost sure sense, e.g., $X = Y$ if $\mathbb{P}[X = Y] = 1$, and $X \geq Y$ if $\mathbb{P}[X \geq Y] = 1$. $E[X]$ will denote the expected value of an r.v. X.

On an arbitrary set of r.v.'s, let $\succeq$ be a weak preference relation, which means that if $X \succeq Y$ then either X is preferred over Y, i.e. $X \succ Y$, or the decision maker is indifferent in choosing between X and Y, i.e. $X \sim Y$ (indifference relation). The preference relation $\succeq$ is "*rational*," if it satisfies a certain set of *non contradictory* axioms.

We begin with reviewing basic axioms on $\succeq$, such as transitivity, monotonicity, and continuity, which are traditionally associated with rational choice.

The first axiom is that $\succeq$ defines a complete weak order.[5] It guarantees that the decision maker can always decide between two alternatives.

Axiom 1 (complete weak order): $\succeq$ is connected and transitive:

(i) $X \succeq Y$ or $Y \succeq X$ for all X and Y (completeness).
(ii) $X \succeq Y$ and $Y \succeq Z$ imply $X \succeq Z$ for all X, Y and Z (transitivity).

The next axiom is "the more the better."

Axiom 2 (strict monotonicity): If $X \geq Y$ then $X \succeq Y$. Moreover, for constants C_1 and C_2, $C_1 > C_2$ implies $C_1 \succ C_2$.

In contrast to the axiom of *non-strict monotonicity*,[5] frequently encountered in the literature and stated as "$X \geq Y \Rightarrow X \succeq Y$," the additional condition in A2 excludes the degenerate case of indifference between two unequal *sure* outcomes.

In real-life projects, probability distributions of natural hazards can only be estimated. If infinitely small variations in an r.v. do not significantly affect preferences, the preference relation $\succeq$ is called *continuous*.

Axiom 3 ($\mathcal{L}^p$-continuity): For every $X \in \mathcal{L}^p(\Omega)$, sets $\overline{\mathcal{B}}(X) = \{Y \in \mathcal{L}^p(\Omega) | Y \succeq X\}$ and $\underline{\mathcal{B}}(X) = \{Y \in \mathcal{L}^p(\Omega) | X \succeq Y\}$ are closed in $\mathcal{L}^p(\Omega)$.

Yaari[5] considered A3 for $p = 1$ only.

The next axiom precludes options such as "infinitely good" or "infinitely bad" in rational choice.

Axiom 4 (finiteness): For any r.v. $X \in \mathcal{L}^p(\Omega)$, there exist constants C_1 and C_2 such that $C_1 \succeq X \succeq C_2$.

If $p = \infty$ or the range of r.v.'s is bounded, A4 follows from A2.

Axioms A1–A4 imply that the preference relation $\succeq$ is *cardinal*, i.e., there exists a functional $U : \mathcal{L}^p(\Omega) \to \mathbb{R}$, called *numerical representation* of $\succeq$ or *utility functional*,[9] such that

(a) $X \succeq Y$ if and only if $U(X) \geq U(Y)$.

Any monotone transformation of $U(X)$ is a numerical representation of the same preference relation $\succeq$.

A1–A4 also imply that there exists a unique utility functional $U(X)$ such that

(b) $U(C) = C$ for any constant C.

If U satisfies both (a) and (b), then $X \sim U(X)$ for every r.v. X. In this case, $U(X)$ is called the *certainty equivalent* (17) of X.

In this work, special attention is paid to so-called law-invariant and risk-averse preference relations. A preference relation $\succeq$ is called *law invariant*[4] if $X \sim Y$ for any two r.v.'s X and Y with the same distribution. It is called *risk averse* if $X \succeq Y$ for any r.v. Y obtained from an r.v. X by *mean-preserving spread.* (19) In this case, $\succeq$ is also said to be *monotone* with respect to mean-preserving spread.

Definition 1: Let X and Z be r.v.'s such that $E[Z|X = x] = 0$ for all $x \in \mathbb{R}$. If Y and $X + Z$ have the same distribution, then we say that Y can be obtained from X by *mean-preserving spread.*[5]

Axiom 5 (risk aversion): If Y can be obtained from X by mean-preserving spread, then $X \succeq Y$. (5)

By Definition 1, X and Y, having the same distribution, can be obtained one from the other by mean-preserving spread, and consequently, in this case, $X \sim Y$, i.e., A5 implies law invariance.

It also follows from Definition 1 that every r.v. X can be obtained from its expected value $E[X]$ by mean-preserving spread, and thus, A5 implies $E[X] \succeq X$. In general, given a choice between two random outcomes with equal expected values, a risk-averse agent prefers the less "dispersive" one.

A preference relation $\succeq$ is *strictly risk averse* if in addition to axiom A5, $E[X] \succ X$ for every nonconstant X, i.e. the expectation of a nonconstant r.v. is strictly preferred to the r.v. itself.

Often risk-averse $\succeq$ is associated with the notion of second order stochastic dominance (SSD). (19)

An r.v. X dominates another r.v. Y by SSD, and we write $X \succcurlyeq_2 Y$, if and only if

$$\int_{-\infty}^{x} F_X(t)dt \le \int_{-\infty}^{x} F_Y(t)dt \quad \text{for all } x \in \mathbb{R}, \qquad (1)$$

where F_X and F_Y are the cumulative distribution functions (CDF's) of X and Y, respectively. Theorem 2 in Rothschild and Stiglitz (19) proves that Y can be obtained from X by mean-preserving spread if and only if $E[X] = E[Y]$ and $X \succcurlyeq_2 Y$. This fact results in the following proposition.

Proposition 1: Let $\succeq$ satisfy axiom A1. Then the following are equivalent:

(a) $\succeq$ satisfies axiom A5 and the non-strict monotonicity axiom.
(b) $\succeq$ is consistent with SSD, i.e. $X \succcurlyeq_2 Y$ implies $X \succeq Y$.

Proof. The proposition follows from Theorem 2 in Rothschild and Stiglitz (19) and the fact that, on an atomless probability space, $X \succcurlyeq_2 Y$ if and only if there exists an r.v. Z such that $X \ge Z$, $Z \succcurlyeq_2 Y$, and $E[Z] = E[Y]$ (see Theorems 2.2 and 2.4 in Bauerle and Muller (21)).

3. MODELS OF RATIONAL CHOICE

The five basic axioms, discussed in the previous section, are commonly agreed to represent main principles of rational behavior. In this section, we continue reviewing axioms on $\succeq$ in the context of the existing models of rational choice. However, in contrast to the previous five, next axioms are often inconsistent with empirical evidence, known as counter-examples (paradoxes). Remarkably, counter-examples of some axioms are resolved by other axioms and vice versa. We will discuss this phenomenon and suggest a new set of axioms, which resolve some counter-examples of the existing models.

Axioms on $\succeq$ are often stated for *lotteries.*

Definition 2: Let $F_X(z)$ and $F_Y(z)$ be CDF's for r.v.'s X and Y, respectively. For every $\lambda \in [0, 1]$, an r.v. Z with the CDF $F_Z(z) = \lambda F_X(z) + (1 - \lambda)F_Y(z)$ is called λ-*lottery* of X and Y. We write $Z = \lambda X \oplus (1 - \lambda)Y$.

The term "lottery" is justified by the fact that the total outcome of two alternative r.v.'s X and Y, realizing with probabilities λ and $1 - \lambda$, respectively, can be represented by the r.v. $Z = \lambda X \oplus (1 - \lambda)Y$.

The next axiom is due to von Neumann and Morgenstern. (2)

Axiom 6 (independence axiom): Let X, Y, and Z be arbitrary r.v.'s such that $X \succeq Y$. Then $\lambda X \oplus (1 - \lambda)Z \succeq \lambda Y \oplus (1 - \lambda)Z$ for any $\lambda \in [0, 1]$.

[4] Such preferences have been intensively studied in the literature under different names. In particular, Machina and Schmeidler (18) called them "probabilistically sophisticated."
[5] This is equivalent to saying that X dominates Y in *concave order.* (20)

Axioms A1–A4 and A6 constitute the foundation of the *expected utility theory* (2) (EUT), i.e., for $\succeq$ satisfying A1–A4 and A6, there exists a *strictly increasing* function $u : \mathbb{R} \to \mathbb{R}$ such that $X \succeq Y$ if and only if $E[u(X)] \geq E[u(Y)]$ for every X and Y. Also, $\succeq$ satisfies risk aversion axiom A5 if and only if u is *concave.*

The EUT is one of the most well-known models of rational choice. However, Kahneman and Tversky (3) showed that the independence axiom is often violated in actual decision making, which was the reason for either omitting A6 or changing it to other axioms. In particular, Yaari (5) replaced the lottery operation $\oplus$ in A6 by the usual addition of *comonotone*[6] random variables and, thus, obtained the so-called *dual independence* axiom.

Axiom 7 (dual independence axiom): Let $X, Y, Z \in \mathcal{L}^1(\Omega)$ be pairwise comonotone r.v.'s such that $X \succeq Y$. Then $\lambda X + (1-\lambda)Z \succeq \lambda Y + (1-\lambda)Z$ for any $\lambda \in [0,1]$.

Axioms A1–A5 and A7 result in the *dual utility theory* (5) (DUT), i.e., for $\succeq$ satisfying A1–A5 (A3 for $p = 1$) and A7, there exists a nonnegative nonincreasing function $h(\alpha) \in \mathcal{L}^\infty[0,1]$ with $\int_0^1 h(\alpha)d\alpha = 1$ such that for every X and Y in $\mathcal{L}^1(\Omega)$, $X \succeq Y$ if and only if $\mathcal{U}(X) \geq \mathcal{U}(Y)$ with the dual utility functional $\mathcal{U}$ given by

$$\mathcal{U}(X) = \int_0^1 h(\alpha) \cdot q_X(\alpha) d\alpha, \tag{2}$$

where $q_X(\alpha) = \inf\{z | P[X \leq z] > \alpha\}$ is the α-quantile of an r.v. X.[7]

The DUT solves some well-known paradoxes of the EUT, e.g., Allais' paradox (certainty effect), common ratio effect, inconsistency with the Gini index,[8] etc. (5,23) However, according to Yaari, (5) "for each paradox of the EUT, one can usually construct a "dual paradox" of the EUT, by interchanging the roles of payments and probabilities."

Dual independence axiom A7 implies that the preference relation $\succeq$ is *uncertainty averse* and *constant risk averse.*

Axiom 8 (uncertainty aversion): Let X and Y be comonotone r.v.'s such that $X \sim Y$. Then $\lambda X + (1-\lambda)Y \succeq Y$ for any $\lambda \in [0,1]$.

Axiom 9 (constant risk aversion): Let X and Y be arbitrary r.v.'s. Then $X \succeq Y$ if and only if $\lambda X + C \succeq \lambda Y + C$ for any constant C and any $\lambda > 0$.

A8 and A9 are particular cases of A7. Indeed, A8 can be obtained from A7 by substituting $X \sim Y$ and $Y = Z$, whereas A9 follows from A7 by Proposition 2 in Yaari. (5)

A8 means that a linear combination of two equivalent comonotone r.v.'s is at least as good as each of them. Under A5, the comonotonicity assumption can be omitted. A preference relation $\succeq$ satisfying A1, A2, A5, and A8 is called "uncertainty averse." (24) A9 generalizes the notions of constant absolute risk aversion and constant relative risk aversion for an arbitrary preference relation. (5,25)

Replacing A7 in the DUT by A8–A9 results in the system of axioms A1–A5 and A8–A9 that corresponds to *coherent risk measures.* (8,9)

Definition 3 (coherent risk measures): A *coherent risk measure* is a functional $\mathcal{R} : \mathcal{L}^1(\Omega) \to (-\infty; \infty]$ satisfying

(R1) $\mathcal{R}(X+C) = \mathcal{R}(X) - C$ for all X and constants C,
(R2) $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for all X and all constant $\lambda > 0$,
(R3) $\mathcal{R}(X+Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$ for all X and Y,
(R4) $\mathcal{R}(X) \leq \mathcal{R}(Y)$ when $X \geq Y$.

A coherent risk measure $\mathcal{R}$ is called *law invariant* if $\mathcal{R}(X) = \mathcal{R}(Y)$ for any two r.v.'s X and Y with the same distribution. $\mathcal{R}$ is *continuous* if it is continuous in $\mathcal{L}^1$ norm; and $\mathcal{R}$ is *finite* if $\mathcal{R}(X) < \infty$ for all X. Theorem 4.1 in Dana (26) shows that every continuous and law-invariant coherent risk measure $\mathcal{R}$ is *SSD-consistent*, i.e. $\mathcal{R}(X) \leq \mathcal{R}(Y)$ for every two r.v.'s X and Y such that $X \succcurlyeq_2 Y$.

Proposition 2: The following are equivalent:

(a) A preference relation $\succeq$ on $\mathcal{L}^1(\Omega)$ satisfies axioms A1–A5 and A8–A9.
(b) There exists a law-invariant finite continuous *coherent* risk measure $\mathcal{R} : \mathcal{L}^1(\Omega) \to \mathbb{R}$ such that $X \succeq Y \Leftrightarrow \mathcal{R}(X) \leq \mathcal{R}(Y)$ for every $X, Y \in \mathcal{L}^1(\Omega)$.

[6] Two r.v.'s $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ are comonotone, if there exists a set $A \subseteq \Omega$ such that $P[A] = 1$ and $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0$ for all $\omega_1, \omega_2 \in A$.
[7] Yaari (5) proved this result for the case of bounded r.v.'s, whereas Guriev (22) proved it in a general case.
[8] The Gini index is defined by $1 - 2\int_0^1 L(x)dx$, where $L(x)$ is the Lorentz curve.

Proof. For nonnegative bounded r.v.'s, similar statements were proved by Alcantud and Bosi[17] and by Safra and Segal[25] (see Lemma 4). For a general case, the proof is given in Appendix 5.2.

The preference relation axioms corresponding to coherent risk measures are, however, also subjected to criticism. The following discussion addresses this issue.

Constant risk aversion A9 implies that profits and losses are weighed equally.

Example 1 (constant absolute risk aversion): If the sure payoff of \$4,500 is preferred to the 1/2-lottery of winning either \$0 or \$10,000, then A9 implies that the sure loss of \$5,500 should be preferred to the 1/2-lottery of loosing \$0 or \$10,000.

The observation that risk aversion becomes *much less* prevalent for negative payoffs is well-known and supported by empirical evidence. Moreover, Kahneman and Tversky[3] argued that for negative payoffs the majority of people even become *risk seeking*.[9]

Example 2 (constant relative risk aversion): If the 1/2-lottery of winning either \$0 or \$100 is preferred to the sure payoff of \$45, then A9 implies that the 1/2-lottery of winning \$0 or \$10,000 should be preferred to the sure payoff of \$4,500.

Holt and Laury[27] observed that "... risk aversion is *much more* prevalent when earnings are scaled up by a factor of twenty," which, in fact, does not support Example 2 and A9 in general.

In this work, we focus on a risk-averse preference relation and assume that risk aversion is also preserved for losses (though it may be much more prevalent for gains). In Examples 1 and 2, the lotteries have the expected gain greater than the corresponding sure payoff, which is the cause of the paradoxes associated with the axioms of constant absolute risk aversion and constant relative risk aversion. This and similar paradoxes do not arise if two random outcomes have the same expected payoff, in which case, only the *dispersion* of the outcomes is compared. This observation suggests relaxing axioms A8–A9 requiring them to hold only for r.v.'s with equal expectations.

[9] In terms of the EUT, this means that for gains, a utility function is concave and for losses, the utility function is convex.

Axiom 10: Let X and Y be comonotone r.v.'s such that $E[X] = E[Y]$ and $X \sim Y$. Then $\lambda X + (1-\lambda)Y \succeq Y$ for any constant $\lambda \in [0, 1]$.

Axiom 11: Let X and Y be arbitrary r.v.'s in $\mathcal{L}^p(\Omega)$ such that $E[X] = E[Y]$. Then $X \succeq Y$ if and only if $\lambda X + C \succeq \lambda Y + C$ for any constant C and $\lambda > 0$.

Example 3 (illustration of A11): If a random payoff with the uniform distribution between \$0 and \$200 is preferred to the lottery with 4/5 probability of winning \$125 (and \$0 otherwise), then A11 implies that the uniformly distributed payoff between \$0 and \$20,000 is preferred to the lottery with 4/5 probability of winning \$12,500 (and \$0 otherwise).

In Example 3, since both the alternatives have equal expected payoffs, a risk-averse agent compares only dispersion of the random outcomes, and if the uniform distribution is found to be less dispersive, then any linear transformation of the random outcomes should not change the original preference.[10]

We summarize the discussed models of rational choice:

- A1–A4 and A6: *expected utility theory*
- A1–A5 and A7: *dual utility theory*
- A1–A5 and A8–A9: *coherent risk measures*
- A1–A5 and A10–A11: *mean-deviation analysis*

The next section proves that a numerical representation of $\succeq$ satisfying A1–A5 and A10–A11 can be constructed in the mean-deviation framework, which generalizes the classical mean-variance approach.

4. MEAN-DEVIATION ANALYSIS IN THE THEORY OF CHOICE

In decision making, general deviation measures,[28,10] in contrast to expected utility, dual utility, and risk measures,[8,29] are usually used in the risk-reward setting, originated from Markowitz's mean-variance model.[30] In this case, there are two well-known approaches: (i) minimizing a deviation measure $\mathcal{D}$ of a random outcome X with constrained $E[X]$ (or maximizing $E[X]$ subject to a constraint on $\mathcal{D}(X)$), and (ii) maximizing a functional that

[10] Of course, this conclusion is based on the assumption that a measure of dispersion is positively homogeneous and invariant to constant shift.

combines $E[X]$ and $\mathcal{D}(X)$. An example of (i) is a Markowitz-type portfolio selection problem solved by Rockafellar et al.,[14] and an example of (ii) is the "mean-risk models" with standard lower semideviation and mean-absolute deviation analyzed by Ruszczynski and Ogryczak[31,32,33] in relation with SSD.

This section studies a mean-deviation model in the context of the theory of choice and shows that under certain conditions on a deviation measure, the preference relation $\succeq$ induced by the model satisfies axioms A1–A5 and A10–A11.

4.1 Mean-deviation model

Following Rockafellar et al.,[10] we define deviation measures on $\mathcal{L}^2(\Omega)$.

Definition 4 (general deviation measures): A *deviation measure* is a functional $\mathcal{D} : \mathcal{L}^2(\Omega) \to [0;\infty]$ satisfying

(D1) $\mathcal{D}(X + C) = \mathcal{D}(X)$ for all X and constants C,
(D2) $\mathcal{D}(0) = 0$, and $\mathcal{D}(\lambda X) = \lambda\mathcal{D}(X)$ for all X and all $\lambda > 0$,
(D3) $\mathcal{D}(X + Y) \le \mathcal{D}(X) + \mathcal{D}(Y)$ for all X and Y,
(D4) $\mathcal{D}(X) \ge 0$ for all X, with $\mathcal{D}(X) > 0$ for nonconstant X.

$\mathcal{D}$ is called a *lower range dominated* deviation measure[28,10] if it has the following property[11]

(D5) $\mathcal{D}(X) \le E[X] - \inf X$ for all X.

The most well-known examples of deviation measures are:

(a) deviation measures of $\mathcal{L}^p$ type $\mathcal{D}(X) = \|X - E[X]\|_p$, $p \in [1,\infty]$, e.g., standard deviation $\sigma(X) = \|X - E[X]\|_2 = \sqrt{E[X - E[X]]^2}$ and mean absolute deviation $\mathrm{MAD}(X) = \|X - E[X]\|_1 = E[|X - E[X]|]$, where $\|\cdot\|_p$ is the $\mathcal{L}^p$ norm;
(b) deviation measures of semi-$\mathcal{L}^p$ type $\mathcal{D}_\pm(X) = \|[X - E[X]]_\pm\|_p$, $p \in [1,\infty]$, e.g., *standard lower semideviation* $\sigma_-(X) = \|[X - E[X]]_-\|_2$ and *standard upper semideviation* $\sigma_+(X) = \|[X - E[X]]_+\|_2$, where $[X]_\pm = \max\{0, \pm X\}$;
(c) conditional value-at-risk (CVaR) deviation, defined for any $\alpha \in (0,1)$ by[12]

$$\mathrm{CVaR}_\alpha^\Delta(X) \equiv E[X] - \frac{1}{\alpha}\int_0^\alpha q_X(\beta)d\beta. \qquad (3)$$

For other examples, see Rockafellar et al.[10]

Deviation measures correspond one-to-one with *averse* measures of risk, $\mathcal{R}$, introduced by Rockafellar et al.,[28,10][13] through the relationships:

$$\begin{aligned} &(a)\ \ \mathcal{D}(X) = \mathcal{R}(X) + E[X], \\ &(b)\ \ \mathcal{R}(X) = \mathcal{D}(X) - E[X]. \end{aligned} \qquad (4)$$

The intersection of the classes of averse measures of risk and coherent risk measures[8] is a special class of so-called *coherent averse* risk measures, which in addition to R1–R4 satisfy

(R5) $\mathcal{R}(X) > E[-X]$ for all nonconstant X.

Through (4), coherent averse[14] risk measures correspond one-to-one with *lower range dominated* deviation measures.[28,10]

A deviation measure $\mathcal{D}$ is *continuous* if it is continuous in $\mathcal{L}^2$. Proposition 3.8 in Grechuk et al.[34] proves that $\mathcal{D}$ is continuous if and only if it is finite everywhere on $\mathcal{L}^2(\Omega)$.

On an atomless probability space Ω, a continuous law-invariant[15] deviation measure $\mathcal{D}$ is consistent with *concave ordering*, i.e., $E[X] = E[Y]$ and $X \succcurlyeq_2 Y$ imply $\mathcal{D}(X) \le \mathcal{D}(Y)$, see Theorem 4.1 in Dana.[26] Further, $\mathcal{D}$ is always assumed to be continuous and law invariant.

The expected value $E[X]$ and a deviation measure $\mathcal{D}(X)$ of a random outcome X can be combined in a single *mean-deviation functional*

$$U(X) = V(E[X], \mathcal{D}(X)) \qquad (5)$$

such that

(V1) $V(m,d)$ is strictly increasing as a function of m for every d.
(V2) $V(m,d)$ is strictly decreasing as a function of d for every m.
(V3) $V(m,0) = m$ for every m (normalization).

Condition V3 does not reduce generality, since any monotone transformation of $V(m,d)$ preserves $\succeq$.

[11] Here and further, inf is understood as the essential infimum.

[12] $\mathrm{CVaR}_1^\Delta(X) = -E[X]+E[X] = 0$ is not a deviation measure, since it vanishes for all r.v.'s (not only for constants).

[13] In Rockafellar et al.,[28,10] averse measures of risk were originally called *strictly expectation bounded* risk measures.

[14] For law-invariant risk measures, the relation $E[X] \succcurlyeq_2 X$ implies $\mathcal{R}(X) \ge \mathcal{R}(E[X]) = E[-X]$, and the novelty in R5 is the strict inequality.

[15] $\mathcal{D}$ depends only on the distribution of an r.v.

The mean-deviation functional (5), in which V satisfies V1–V3, will be called *mean-deviation model*. Rockafellar et al.[(15)] showed that there exists market equilibrium for agents whose risk-reward preferences are represented by the mean-deviation model (5) with different deviation measures.

4.2 Mean-deviation preference relation

A preference relation $\succeq$ consistent with the mean-deviation model (5), i.e., such that $X \succeq Y$ if and only if $V(E[X], \mathcal{D}(X)) \geq V(E[Y], \mathcal{D}(Y))$, will be called *mean-deviation preference relation*. The rest of the section explores conditions on $\mathcal{D}$ under which mean-deviation $\succeq$ satisfies axioms of rational choice.

PROPOSITION 3: If a mean-deviation preference relation satisfies axioms A1–A4, then $\mathcal{D}$ in (5) is such that

$$\sup_{X \not\equiv \text{const}} \frac{\mathcal{D}(X)}{E[X] - \inf X} < \infty. \tag{6}$$

Proof. By contradiction, suppose (6) is not true. Then $U(X) \geq U(Y)$ for every X and Y such that $E[X] > E[Y]$. Indeed, let $m_X = E[X]$, $d_X = \mathcal{D}(X)$, $m_Y = E[Y]$, $d_Y = \mathcal{D}(Y)$, and $e = m_X - m_Y$. Since by the assumption, the supremum in (6) is infinite, there exists an r.v. X_1 such that

$$\frac{\mathcal{D}(X_1)}{E[X_1] - \inf X_1} \geq \frac{d_X}{e}.$$

Then the r.v.

$$X_2 = e \cdot \frac{X_1 - \inf X_1}{E[X_1] - \inf X_1} + m_Y,$$

is such that $E[X_2] = e + m_Y = m_X$ and $\mathcal{D}(X_2) \geq d_X$, and V1 and V2 imply $V(m_X, d_X) \geq V(E[X_2], \mathcal{D}(X_2)) = U(X_2)$. On the other hand, observe that $X_2 \geq m_Y$. Consequently, by strict monotonicity axiom A2, $X_2 \succeq m_Y$, so that $U(X_2) \geq U(m_Y)$. By V3 and V2, $U(m_Y) = m_Y = V(m_Y, 0) \geq V(m_Y, d_Y)$. This reasoning can be summarized in a single inequality

$$\begin{aligned} V(m_X, d_X) &\geq V(E[X_2], \mathcal{D}(X_2)) = U(X_2) \\ &\geq m_Y = V(m_Y, 0) \geq V(m_Y, d_Y), \end{aligned}$$

which proves that $U(X) \geq U(Y)$ for every X and Y with $E[X] > E[Y]$.

Fix any m and any $d > 0$. Let X be an arbitrary r.v. with $m = E[X]$ and $d = \mathcal{D}(X)$. Then $m + \epsilon = U(m + \epsilon) \geq U(X) \geq U(m - \epsilon) = m - \epsilon$ for any $\epsilon > 0$, and continuity axiom A3 implies $U(X) = m$, or, equivalently, $V(m, d) \equiv m$. But this contradicts V2, and the proof is finished.

Proposition 3 restricts our attention to deviation measures with the following property.

DEFINITION 5: A deviation measure $\mathcal{D}$ is called *weakly lower-range dominated* if

$$\sup_{X \not\equiv \text{const}} \frac{\mathcal{D}(X)}{E[X] - \inf X} = K < \infty. \tag{7}$$

By Definition 5, every lower-range dominated deviation measure is weakly lower-range dominated with $K \leq 1$, e.g., standard lower semideviation and CVaR deviation (see Rockafellar et al.[(10)]). Then (7) implies that $\mathcal{D}$ is weakly lower range dominated with some $K > 0$ if and only if the deviation measure $\left(K^{-1}\mathcal{D}\right)(X) \equiv K^{-1} \cdot \mathcal{D}(X)$ is weakly lower-range dominated with $K = 1$.

EXAMPLE 4: MAD is weakly lower range dominated with $K = 2$.

Detail. $\mathcal{D}(X) = \frac{1}{2} \cdot E[|X - E[X]|]$ is lower range dominated (see Rockafellar et al.[(13)]), and thus, $\text{MAD}(X) \equiv E[|X - E[X]|]$ is weakly lower range dominated with $K \leq 2$. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence with $\mathbb{P}[X_n = 0] = 1 - 1/n$ and $\mathbb{P}[X_n = n] = 1/n$. Then $E[X_n] - \inf X_n = 1$, and $\lim_{n\to\infty} \text{MAD}(X_n) = 2$, so that $K = 2$. □

EXAMPLE 5: The standard deviation, σ, is not weakly lower range dominated.

Detail. By contradiction, suppose σ is weakly lower range dominated for some $K > 0$. Let X be the r.v. with $\mathbb{P}[X = 0] = 1-(K+1)^{-2}$ and $\mathbb{P}[X = (K+1)^2] = (K+1)^{-2}$. Then $E[X] - \inf X = 1$ and $\sigma(X) = \sqrt{K(K+2)} > K$, which contradicts (7). □

Next, we establish necessary and sufficient conditions for the function V, under which a mean-deviation preference relation satisfies axioms A1–A4.

PROPOSITION 4: Let $\mathcal{D}$ in the mean-deviation model (5) satisfy (7) for some K.[16] Then the following are equivalent:

(a) A mean-deviation preference relation satisfies axioms A1–A4 (A3 with $p = 2$).

[16] $\mathcal{D}$ is always assumed to be continuous and law-invariant.

(b) The function $V(m,d)$ in (5) is continuous and also satisfies

(V4) $V(m,d) \le V(m+a, d+a\cdot K)$ for every m, $d \ge 0$ and $a \ge 0$.

Proof. See Appendix 5.2.

Proposition 4 paves the way for constructing a mean-deviation model that induces the preference relation $\succeq$ satisfying A1–A4.

Example 6 (linear mean-deviation model): With the equality in V4, we obtain $V(m,d) = m - d/K$, which satisfies V1–V4. Then $U(X) = V(E[X], \mathcal{D}(X)) = E[X] - \mathcal{D}(X)/K$, and $\mathcal{R}(X) = -U(X)$ is a coherent averse risk measure. In this sense, coherent averse risk measures are a particular case of the mean-deviation model (5).

Example 7 (nonlinear mean-deviation model): For every $q \in (0,1)$, the functional (5) with $V(m,d) = m - ((1+d)^q - 1)/K$, satisfying V1–V4, induces a "nonlinear" preference relation that satisfies A1–A4.

Example 8 (piece-wise linear mean-deviation model): For every $K_1 > K_2 \ge K$ and $d_0 > 0$, the function

$$V(m,d) = \begin{cases} m - d/K_1 & d \le d_0, \\ m - d_0/K_1 - (d-d_0)/K_2 & d > d_0, \end{cases}$$

provides yet another example of the mean-deviation model (5), in which an agent prefers to have a deviation less than d_0.

Example 9 (model inconsistent with V4): There is no K, for which $V(m,d) = m - \lambda d^q$ with $\lambda > 0$ and $q > 1$, satisfies V4; see Example 1 in Rockafellar et al.[15] Thus, for any deviation measure, the corresponding mean-deviation preference relation fails to satisfy axioms A1–A4. For instance, let $V(m,d) = m - d^2$ with $\mathcal{D}(X) = \mathrm{MAD}(X)$, and let X be the 1/2-lottery between 0 and 2. Then $2X \ge X$ but $X \succ 2X$. Indeed, $V(E[X], \mathrm{MAD}(X)) = V(1,1) = 0$ and $V(E[2X], \mathrm{MAD}(2X)) = V(2,2) = -2 < 0$.

4.3 Main result

Our goal is to prove that a preference relation $\succeq$ on $\mathcal{L}^2(\Omega)$ satisfies A1–A5 and A10–A11 if and only if it is mean-deviation $\succeq$ for (5) with continuous V satisfying V1–V4 and with weakly lower-range dominated $\mathcal{D}$ for some $K > 0$. A3 will be considered for $p = 2$, and A5 will be understood in strict sense ($E[X] \succ X$ for every nonconstant X).

For brevity, let $\overline{X} = X - E[X]$. We begin with the following auxiliary results.

Proposition 5: Let $U(X)$ be the certainty equivalent of a preference relation $\succeq$ on $\mathcal{L}^2(\Omega)$ satisfying A1–A5 and A11, and let X_0 be a nonconstant r.v. with $E[X_0] = 0$. Then

(a) $\phi(\lambda) = -U(\lambda X_0)$ is a continuous strictly increasing function from $[0,\infty)$ to $[0,\infty)$ with $\phi(0) = 0$.

(b) $\mathcal{D}(X) \equiv \phi^{-1}(-U(X - E[X])) : \mathcal{L}^2(\Omega) \mapsto [0,\infty)$ is a well-defined continuous law-invariant functional.

Proof. See Appendix 5.2.

Proposition 6: Let a preference relation $\succeq$ on $\mathcal{L}^2(\Omega)$ satisfy A1–A5, A10, and A11. Then the functional in part (b) of Proposition 5 is a continuous law-invariant deviation measure $\mathcal{D}$.

Proof. D1 follows from the fact that $\overline{X+C} = \overline{X}$ for any constant C. Next we show D4. For any constant C, $\mathcal{D}(C) = \phi^{-1}(-U(0)) = 0$. For nonconstant X, strict risk aversion A5 implies $0 = E\left[\overline{X}\right] \succ \overline{X}$, and thus, by monotonicity of ϕ, $\mathcal{D}(X) = \phi^{-1}(-U(\overline{X})) > \phi^{-1}(-U(0)) = 0$, which proves D4.

Now let us show that the functional satisfies D2. We have $-U(\overline{X}) = \phi(\mathcal{D}(X)) = -U(\mathcal{D}(X)\cdot X_0)$ or

$$\overline{X} \sim \mathcal{D}(X)\cdot X_0 \quad \text{for all} \quad X \in \mathcal{L}^2(\Omega). \tag{8}$$

Since $E\left[\overline{X}\right] = 0 = E[\mathcal{D}(X)\cdot X_0]$, by A11, $\lambda\overline{X} \sim \lambda\mathcal{D}(X)\cdot X_0$ for every $\lambda > 0$. On the other hand, (8) implies $\lambda X - E[\lambda X] \sim \mathcal{D}(\lambda X)\cdot X_0$, and thus, $\lambda\mathcal{D}(X)\cdot X_0 \sim \mathcal{D}(\lambda X)\cdot X_0$ or $\phi(\lambda\mathcal{D}(X)) = \phi(\mathcal{D}(\lambda X))$. This fact and strict monotonicity of ϕ, shown in Proposition 5, imply that $\lambda\mathcal{D}(X) = \mathcal{D}(\lambda X)$, which is D2.

To prove D3, we start with the case of comonotone $X, Y \in \mathcal{L}^2(\Omega)$. A11 and (8) imply $\overline{X}/\mathcal{D}(X) \sim X_0$ and $\overline{Y}/\mathcal{D}(Y) \sim X_0$, so that $\overline{X}/\mathcal{D}(X) \sim \overline{Y}/\mathcal{D}(Y)$. Since $E\left[\overline{X}/\mathcal{D}(X)\right] = E\left[\overline{Y}/\mathcal{D}(Y)\right] = 0$, by A10,

$$\lambda\overline{X}/\mathcal{D}(X) + (1-\lambda)\overline{Y}/\mathcal{D}(Y) \succeq \overline{Y}/\mathcal{D}(Y) \sim X_0 \tag{9}$$

for all $\lambda \in [0,1]$. Substituting $\lambda = \mathcal{D}(X)/(\mathcal{D}(X)+\mathcal{D}(Y))$ into (9), we obtain $(\overline{X}+\overline{Y})/(\mathcal{D}(X)+\mathcal{D}(Y)) \succeq X_0$, and for all comonotone X and Y, A11 implies

$$\overline{X} + \overline{Y} \succeq (\mathcal{D}(X)+\mathcal{D}(Y))\cdot X_0. \tag{10}$$

Now with (8) and (10), $\mathcal{D}(X+Y)X_0 \sim \overline{X}+\overline{Y} \succeq (\mathcal{D}(X)+\mathcal{D}(Y))\cdot X_0$, and consequently, $\phi(\mathcal{D}(X)+\mathcal{D}(Y)) \geq \phi(\mathcal{D}(X+Y))$. Since ϕ is strictly increasing, D3 follows.

It is left to show that (10) holds true for noncomonotone r.v.'s. Lemma 4.2 in Dana [(26)] proves that for arbitrary r.v.'s X and Y, there is an r.v. X' comonotone with Y and having the same distribution as X. Then by Proposition 14 in Appendix 5.2, $X+Y \succcurlyeq_2 X'+Y$ or $\overline{X}+\overline{Y} \succcurlyeq_2 \overline{X}'+\overline{Y}$, and Proposition 1 and (10) imply

$$\begin{aligned}\overline{X}+\overline{Y} \succeq \overline{X}'+\overline{Y} &\succeq (\mathcal{D}(X')+\mathcal{D}(Y))\cdot X_0\\ &= (\mathcal{D}(X)+\mathcal{D}(Y))\cdot X_0,\end{aligned}$$

which finishes the proof of D3.

Now we state the main result of this work.

PROPOSITION 7: The following are equivalent

(a) a preference relation $\succeq$ on $\mathcal{L}^2(\Omega)$ satisfies axioms A1–A5 and A10–A11 (A3 for $p=2$ and A5 in strict sense).

(b) $X \succeq Y$ if and only if $V(E[X],\mathcal{D}(X)) \geq V(E[Y],\mathcal{D}(Y))$ for every $X,Y \in \mathcal{L}^2(\Omega)$, where

   (i) law-invariant continuous deviation measure $\mathcal{D}:\mathcal{L}^2(\Omega)\to[0,\infty)$ is weakly lower-range dominated with a constant $K>0$, and

   (ii) $V(m,d)$ is a continuous function satisfying V1–V3 and V4 with the constant K.

Proof. Let $\succeq$ on $\mathcal{L}^2(\Omega)$ satisfy (a). A1–A5 guarantee the existence of a certainty equivalent $U:\mathcal{L}^2(\Omega)\to\mathbb{R}$, and our goal is to show that U can be represented in the form (5), where the deviation measure $\mathcal{D}$ and function $V(m,d)$ satisfy the conditions in (b).

Let X_0 be an arbitrary *bounded* nonconstant random variable such that $E[X_0]=0$. As in Proposition 5, let $\phi(\lambda)=-U(\lambda X_0)$ and $\mathcal{D}(X)=\phi^{-1}\left(-U\left(\overline{X}\right)\right)$. Propositions 5b and 6 guarantee that $\mathcal{D}(X)$ is a well-defined continuous law-invariant deviation measure.

For every two r.v.'s X and Y, $E[X]=E[Y]$ and $\mathcal{D}(X)=\mathcal{D}(Y)$ imply $X\sim Y$. Indeed, (8) shows that $\overline{X}\sim\mathcal{D}(X)\cdot X_0$ and $\overline{Y}\sim\mathcal{D}(Y)\cdot X_0$. Since $\mathcal{D}(X)=\mathcal{D}(Y)$, we obtain $\overline{X}\sim\overline{Y}$, and thus, $X\sim Y$ by A11. This means that the certainty equivalent U can be represented in the form (5) for some function of two variables $V(m,d)$.

Next we show that $V(m,d)$ satisfies V1–V4. V3 follows from $U(C)=C$. To prove V2, consider arbitrary m and $d_1>d_2$. By Proposition 5, the function $\phi(\lambda)=-U(\lambda X_0)$ is strictly increasing, and therefore, $d_2X_0\succ d_1X_0$, so that $d_2X_0+m\succ d_1X_0+m$ by A11. Since $E[d_iX_0+m]=m$ and $\mathcal{D}(d_iX_0+m)=d_i$, $i=1,2$, this implies $V(m,d_2)>V(m,d_1)$.

To prove V1, consider arbitrary $m_1>m_2$ and $d\geq 0$. Since X_0 is assumed to be bounded, there exists $\delta>0$ such that $\delta X_0\geq m_2-m_1$. Then $m_1+(d+\delta)X_0\geq m_2+dX_0$, which with A2 implies $V(m_1,d+\delta)\geq V(m_2,d)$. At the same time, using V2, we obtain $V(m_1,d)>V(m_1,d+\delta)$, and consequently, $V(m_1,d)>V(m_2,d)$.

Thus, the utility functional $U(X)$ can be represented in the form (5) with a law-invariant continuous deviation measure $\mathcal{D}$ and a function $V(m,d)$ satisfying V1–V3. Proposition 3 implies that $\mathcal{D}$ is weakly lower range dominated with some constant K. Finally, by Proposition 4, $V(m,d)$ is continuous and satisfies V4 with the same constant K. The proof of (b) is finished.

Now let (b) hold. We should show that $\succeq$ satisfies axioms in (a). A1–A4 hold by Proposition 4. V2 and the fact that every law-invariant continuous deviation measure on an atomless probability space is consistent with concave ordering (see Theorem 4.1 in Dana [(26)]) imply that risk aversion A5 holds for $\succeq$. The risk aversion is then strict by D4.

To show A10, consider comonotone X and Y such that $E[X]=E[Y]$ and $X\sim Y$. Then $V(E[X],\mathcal{D}(X))=V(E[Y],\mathcal{D}(Y))$, which implies $\mathcal{D}(X)=\mathcal{D}(Y)$ by V2. It follows from D2 and D3 that $\mathcal{D}(Z)\leq\lambda\mathcal{D}(X)+(1-\lambda)\mathcal{D}(Y)=\mathcal{D}(Y)$ for every $Z=\lambda X+(1-\lambda)Y$ with $\lambda\in[0,1]$. Since $E[Z]=E[Y]$, by V2, we obtain $V(E[Z],\mathcal{D}(Z))\geq V(E[Y],\mathcal{D}(Y))$ or $Z\succeq Y$.

It is left to show that $\succeq$ satisfies A11. Let X and Y be r.v.'s such that $E[X]=E[Y]$ and $X\succeq Y$. Then $V(E[X],\mathcal{D}(X))\geq V(E[Y],\mathcal{D}(Y))$, whence $\mathcal{D}(X)\leq\mathcal{D}(Y)$ by V2. Now $E[\lambda X+C]=E[\lambda Y+C]$ and $\mathcal{D}(\lambda X+C)\leq\mathcal{D}(\lambda Y+C)$ for any $\lambda>0$ and constant C, and V2 implies $V(E[\lambda X+C],\mathcal{D}(\lambda X+C))\geq V(E[\lambda Y+C],\mathcal{D}(\lambda Y+C))$ or $\lambda X+C\succeq\lambda Y+C$, which is A11.

4.4 Discussion

The main contribution of this work is the set of axioms for a preference relation $\succeq$ and the conditions on V and $\mathcal{D}$ in the mean-deviation model

(5) under which (5) represents $\succeq$. Namely, $\succeq$ that satisfies basic axioms A1–A5 and relaxed axioms of uncertainty aversion A8 and constant risk aversion A9 (required to hold only for random outcomes with equal expected values) can be represented by (5) with V satisfying V1–V3. Example 9 shows, however, that, in general, the converse does not hold: the model (5) with $V(m,d) = m - \lambda d^q$, $\lambda > 0$, $q > 1$, satisfying V1–V3, induces nonmonotone preferences. Propositions 3 and 4 state that (5) is consistent with monotonicity axiom A2 only if $\mathcal{D}$ is weakly lower-range dominated and if in addition to V1–V3, V satisfies V4. This result implies that the mean-deviation model (5) is

(a) consistent with monotone preferences if $\mathcal{D} = \sigma_-$ or $\mathcal{D} = \mathrm{CVaR}_\alpha^\Delta$ and if V satisfies V4 for $K = 1$.
(b) consistent with monotone preferences if $\mathcal{D} =$ MAD and V satisfies V4 for $K = 2$ (Example 4).
(c) *inconsistent* with monotone preferences if $\mathcal{D} = \sigma$ and V satisfies V1–V3 (Example 5).

Key examples of function V satisfying V1–V4 include

(a) Linear mean-deviation model $V(m,d) = m - d/K$ with weakly lower range dominated $\mathcal{D}$ with $K > 0$ (Example 6).
(b) Linear mean-deviation model $V(m,d) = m - d$ with lower range dominated $\mathcal{D}$ (i.e. $K \leq 1$) corresponding to coherent averse risk measures.
(c) Non-linear mean-deviation model $V(m,d) = m - ((1+d)^q - 1)/K$, $q \in (0,1)$, with weakly lower range dominated $\mathcal{D}$ with $K > 0$ (Example 7).

The mean-deviation analysis preserves all the main properties of the mean-variance approach and provides flexibility in expressing individual risk preferences, particularly those that are nonsymmetric with respect to ups and downs of an r.v. In this sense, preference relations satisfying A1–A5 and A10–A11 can be viewed as a monotone approximation for the preference relation $\succeq_\sigma$ induced by the mean-variance approach. Maccheroni et al. [35] found a monotone preference relation "closest" to $\succeq_\sigma$ but failing constant risk aversion A11, whereas a monotone approximation of $\succeq_\sigma$ consistent with A1–A5 and A10–A11 was suggested by Filipovic and Kupper. [36]

Since A8–A9 are less restrictive than dual independence axiom A7, and A10–A11 are less restrictive than A8–A9, both the DUT (A1–A5 and A7) and the theory of coherent averse risk measures (A1–A5 and A8–A9) are particular cases of the mean-deviation model (5) (A1–A5 and A10–A11). Consequently, the mean-deviation analysis resolves those paradoxes of the EUT that the DUT [5] does, and also resolves the paradoxes of the DUT and the theory of coherent averse risk measures related to the axioms of constant risk aversion; see Examples 1 and 2.

5. APPLICATIONS OF THE MEAN-DEVIATION ANALYSIS

This section applies mean-deviation analysis to risk sharing and portfolio selection in the context of theory of choice. In several applications including portfolio optimization, [13,14,15] risk modeling, [31] estimation of return distribution through the maximum entropy principle, [37] etc., risk-reward preferences of an agent are often assumed to depend only on the expectation and some deviation of a random outcome. Proposition 7 justifies this assumption from the perspective of the theory of choice and also indicates what restrictions should be imposed on agent's utility function in each case to guarantee that the corresponding system of axioms for a preference relation is "rational."

5.1 Risk Sharing

Risk sharing is one of the central problems in the actuarial science (insurance) and finance, but it also arises in any project involving cooperation. For example, in the dam proposal selection problem, discussed in the introduction, several agencies with different risk attitudes may contribute to the construction and maintenance of a selected dam and are also assumed to share potential losses due to an earthquake. Given that the risk is shared, how do the agencies make a cooperative choice, and what are conditions that guarantee that the made choice is rational? Other examples of risk sharing include cooperative investment projects in the construction and agricultural sectors under the risks of various natural hazards and disasters, e.g. floods, fires, tornados, hurricanes, earthquakes, volcanic eruptions, etc.

Suppose a random outcome X is divided among n agencies such that agency i receives its share Y_i. A division $Y = (Y_1, \ldots, Y_n)$ is called X-feasible if $\sum_{i=1}^n Y_i = X$. The risk preferences of agency i are represented by a finite utility functional U_i, which induces a connected and transitive preference relation $\succeq_i$ by $X_1 \succeq_i X_2 \Leftrightarrow U_i(X_1) \geq U_i(X_2)$. A utility functional is called *monotone* or *(strictly) risk averse*, if the corresponding preference relation satisfies A2 or A5 (in strict form), respectively. A division $Z = (Z_1, \ldots, Z_n)$ dominates Y if $Z_i \succeq_i Y_i$ for all i and strictly dominates Y if at least one preference is strict, and it ϵ-dominates Y if $U_i(Z_i) > U_i(Y_i) - \epsilon$ for all i.

A preference relation $\succeq_c$ for a coalition of n agencies can be introduced in several ways. The most conservative way is to say that X_1 is preferred over X_2, i.e. $X_1 \succeq_c X_2$, if for every X_2-feasible division Y and every $\epsilon > 0$, there exists an X_1-feasible division Z that ϵ-dominates Y. If $\succeq_i$ is continuous for all $i = 1, \ldots, n$, then $\succeq_c$ is also continuous and independent of the choice of the continuous utility functionals U_i, $i = 1, \ldots, n$. The next proposition proves this fact.

Proposition 8: Let $\succeq_i$ satisfy A3 for all i, and let functionals U_i and $\widetilde{U}_i$, $i = 1, \ldots, n$, be continuous and such that $X_1 \succeq_i X_2$ if and only if $U_i(X_1) \geq U_i(X_2) \Leftrightarrow \widetilde{U}_i(X_1) \geq \widetilde{U}_i(X_2)$. Then the coalition of the agencies with either U_i or $\widetilde{U}_i$, $i = 1, \ldots, n$, has same $\succeq_c$.

Proof. Since $U_i(X_1) \geq U_i(X_2) \Leftrightarrow \widetilde{U}_i(X_1) \geq \widetilde{U}_i(X_2)$ for any X_1 and X_2, there exist continuous increasing functions f_i such that $f_i(U_i(X)) = \widetilde{U}_i(X)$, $i = 1, \ldots, n$. Now let us fix X_1 and X_2 and assume that for every X_2-feasible division $Y = (Y_1, \ldots, Y_n)$ and every $\epsilon > 0$, there exists an X_1-feasible division $Z(\epsilon) = (Z_1(\epsilon), \ldots, Z_n(\epsilon))$ such that $U_i(Z_i(\epsilon)) > U_i(Y_i) - \epsilon$, $i = 1, \ldots, n$. Then for any $\delta > 0$, there exists $\epsilon_i > 0$ such that $f_i(z) > f_i(y) - \delta$ whenever $z > y - \epsilon_i$. Consequently, $\widetilde{U}_i(Z_i(\epsilon)) > \widetilde{U}_i(Y_i) - \delta$, $i = 1, \ldots, n$, for $\epsilon = \min_i \epsilon_i$.

In general, the preference relation $\succeq_c$ fails to satisfy A1, namely, there may exist "incomparable" r.v.'s X_1 and X_2 such that neither $X_1 \succeq_c X_2$ nor $X_2 \succeq_c X_1$. However, the following result holds.

Proposition 9: Let U_i, $i = 1, \ldots, n$, be continuous and constant translation invariant: $U_i(X + C) = U_i(X) + C$ for all X and constants C. Then $\succeq_c$ satisfies A1 and is represented by the utility functional U_c of the coalition given by

$$U_c(X) = \sup_Y \sum_{i=1}^n U_i(Y_i) \quad \text{subject to} \quad \sum_{i=1}^n Y_i = X. \tag{11}$$

Proof. Let $U_c(X_1) \geq U_c(X_2)$, and let Y be an X_2-feasible division. Then (11) implies that for any $\epsilon > 0$, there exists an X_1-feasible division Z such that $\delta/n < \epsilon$, where $\delta = \sum_{i=1}^n U_i(Y_i) - \sum_{i=1}^n U_i(Z_i)$. Then the division $Z^* = (Z_1^*, \ldots, Z_n^*)$ given by $Z_i^* = Z_i - U_i(Z_i) + U_i(Y_i) - \delta/n$ is an X_1-feasible division that ϵ-dominates Y. Conversely, if $U_c(X_1) < U_c(X_2)$, then there exists an X_2-feasible division Y such that $\delta \equiv \sum_{i=1}^n U_i(Y_i) - U_c(X_1) > 0$, so that no X_1-feasible division ϵ-dominates Y for $\epsilon = \delta/n$.

Proposition 9 shows that if U_i, $i = 1, \ldots, n$, are constant translation invariant then it is enough to maximize the sum of these functionals. In other words, in this case, the utility functional (11) represents integral risk preferences of the coalition. If U_i, $i = 1, \ldots, n$, are not constant translation invariant, the coalition preference relation $\succeq_c$ can still be represented by (11), but in this case, $U_c(X_1) \geq U_c(X_2)$ does not imply that every X_2-feasible division is ϵ-dominated by some X_1-feasible division for any $\epsilon > 0$.

Next we establish conditions that guarantee that $\succeq_c$ represented by (11) conforms to the axioms of rational choice. The next proposition states that, surprisingly, if $\succeq_i$ of just some i is monotone, then so is $\succeq_c$.

Proposition 10: Let all U_i, $i = 1, \ldots, n$, be constant translation invariant, and let at least one U_i be monotone. Then U_c given by (11) is monotone.

Proof. Let $X_1 \geq X_2$. For any $C < U_c(X_2)$, there exists an X_2-feasible division Y such that $C < \sum_{i=1}^n U_i(Y_i)$. Let U_j be monotone for some j. Then for X_1-feasible division $Z = (Z_1, \ldots, Z_n)$ given by $Z_i = Y_i$ for $i \neq j$ and $Z_j = Y_j + (X_1 - X_2)$, we have $C < \sum_{i=1}^n U_i(Y_i) \leq \sum_{i=1}^n U_i(Z_i) \leq U_c(X_1)$, so that $U_c(X_2) \leq U_c(X_1)$.

Under V1–V3, the utility functionals $U_i(X) = V_i(E[X], \mathcal{D}_i(X))$, $i = 1, \ldots, n$, are constant translation invariant if and only if V_i, $i = 1, \ldots, n$, are linear: $V_i(m, d) = m - \rho_i \cdot d$ for some constants ρ_i. We recall that all deviation measures are assumed to be law-invariant and continuous (therefore finite), and consequently, U_i, $i = 1, \ldots, n$, are also continuous and finite. The next proposition establishes necessary and sufficient conditions for $\succeq_c$ to be strictly risk averse.

Proposition 11: Let $\succeq_i$ be represented by $U_i(X) = E[X] - \rho_i \, \mathcal{D}_i(X)$, $i = 1, \ldots, n$. Then $\succeq_c$

represented by (11) is strictly risk averse if and only if $\rho_i > 0$, $i = 1, \ldots, n$.

Proof. By contradiction, assume that $\succeq_c$ is strictly risk averse and that $\rho_1 \le 0$. Let X be a noncostant r.v. with $E[X] = 0$, so that $U_c(X) < U_c(0)$. For any 0-feasible division Y, a division $Z = (Z_1, \ldots, Z_n)$ given by $Z_1 = Y_1 + X$, $Z_i = Y_i$, $i \ge 2$, is X-feasible, and $\sum_{i=1}^n U_i(Y_i) - \sum_{i=1}^n U_i(Z_i) = -\rho_1(\mathcal{D}(Y_1) - \mathcal{D}(Y_1 + X)) \le -\rho_1\mathcal{D}(-X)$, so that $U_c(0) - U_c(X) \le -\rho_1\mathcal{D}(-X)$. If $\rho_1 = 0$ or $U_c(0) = +\infty$, then $U_c(0) \le U_c(X)$, which is a contradiction. Now assume that $\rho_1 < 0$ and $U_c(0) < +\infty$. For any $C > 0$, a division $Z = (Z_1, \ldots, Z_n)$ given by $Z_1 = C\,X$, $Z_i = 0$, $i \ge 2$, is $(C\,X)$-feasible, and consequently, $U_c(C\,X) \ge \sum_{i=1}^n U_i(Z_i) = -\rho_1 C\,\mathcal{D}(X)$, which for $C = U_c(0)/(-\rho_1\mathcal{D}(X))$ implies $U_c(C\,X) \ge U_c(0)$, which is a contradiction.

Next, let $\rho_i > 0$, $i = 1, \ldots, n$. In this case, $U_c(C) = C$ for constants C, and for any X-feasible division $Y = (Y_1, \ldots, Y_n)$,

$$U_c(E[X]) - \sum\nolimits_{i=1}^n U_i(Y_i) = \sum\nolimits_{i=1}^n \rho_i\mathcal{D}_i(Y_i) \ge \min_i \rho_i \cdot \left(\inf_Y \max_i \mathcal{D}_i(Y_i)\right).$$

For a nonconstant r.v. X, the right-hand side in the last expression is positive by Proposition 3.3 in Grechuk et al.,[34] and consequently, $U_c(E[X]) > U_c(X)$. Also, this implies that U_c is finite. Then, with $\rho_i > 0$, $i = 1, \ldots, n$, all U_i, $i = 1, \ldots, n$, are concave, continuous and law-invariant, and so is their sup-convolution U_c. Thus, by Theorem 4.1 in Dana,[26] $U_c(X_1) \ge U_c(X_2)$ provided that $E[X_1] = E[X_2]$ and $X_1 \succcurlyeq_2 X_2$. Hence, $\succeq_c$ is strictly risk averse.

Proposition 10 implies that $\succeq_c$ satisfies monotonicity axiom A2 if U_i is monotone for some i. This condition holds if the corresponding deviation measure $\mathcal{D}_i(X)$ is weakly lower range dominated with the constant $1/\rho_i$.

Proposition 12: If $\succeq_i$ is represented by $U_i(X) = E[X] - \rho_i\,\mathcal{D}_i(X)$, $\rho_i > 0$, $i = 1, \ldots, n$, then $\succeq_c$ represented by (11) satisfies A1, A3, A4, and A5 in strict form. If, in addition, for some j, $\mathcal{D}_j(X)$ is weakly lower range dominated with the constant $1/\rho_j$, then $\succeq_c$ also satisfies monotonicity axiom A2.

Proof. A1 obviously holds. A3 follows from the continuity of $\mathcal{D}_i$, $i = 1, \ldots, n$, whereas A4 and A5 in strict form follow from Proposition 11. If, in addition, for some j, $\mathcal{D}_j(X)$ is weakly lower range dominated with the constant $1/\rho_j$, corresponding U_j is monotone by Proposition 7, and so is U_c by Proposition 10.

Remarkably, $\succeq_c$ is strictly risk averse only if $\succeq_i$ is strictly risk averse for *all* i, whereas $\succeq_c$ is guaranteed to be monotone if $\succeq_i$ of at least *one* i is monotone.

Example 10 (dam proposal selection): Suppose that two agencies participate in making a choice between two dam proposals A and B, discussed in the introduction, and share construction and maintenance costs of the selected dam as well as any potential losses due to an earthquake. If $U_1(\cdot) = E[\cdot] - \rho_1\,\mathrm{MAD}(\cdot)$ and $U_2(\cdot) = E[\cdot] - \rho_2\,\sigma_-(\cdot)$ are the utility functionals of the agencies, then a cooperative choice is guaranteed to be strictly risk averse if $\rho_i > 0$, $i = 1, 2$. In this case, $\succeq_c$ satisfies A1, A3, and A4. In addition, it satisfies A2 if either $\rho_1 \le 1/2$ or $\rho_2 \le 1$.

If all the agencies have the utility functions $U_i(\cdot) = E[\cdot] - \rho_i\,\mathcal{D}(\cdot)$, $i = 1, \ldots, n$, with same deviation measure $\mathcal{D}$, the cooperative utility functional U_c in (11) considerably simplifies.

Proposition 13: If $\succeq_i$ is represented by $U_i(X) = E[X] - \rho_i\,\mathcal{D}(X)$, $\rho_i > 0$, $i = 1, \ldots, n$, then with $\rho_{\min} = \min_i \rho_i$, (11) reduces to

$$U_c(X) = E[X] - \rho_{\min}\,\mathcal{D}(X). \tag{12}$$

If, in addition, $\mathcal{D}$ is weakly lower range dominated with the constant $1/\rho_{\min}$, then $\succeq_c$ is monotone.

Proof. For any X-feasible division $Y = (Y_1, \ldots, Y_n)$,

$$\begin{aligned}\sum\nolimits_{i=1}^n U_i(Y_i) &= E[X] - \sum\nolimits_{i=1}^n \rho_i\mathcal{D}_i(Y_i)\\ &\le E[X] - \rho_{\min}\sum\nolimits_{i=1}^n \mathcal{D}_i(Y_i)\\ &\le E[X] - \rho_{\min}\,\mathcal{D}(X).\end{aligned}$$

Let $\rho_j = \min_i \rho_i$. For the division $Y = (Y_1, \ldots, Y_n)$ given by $Y_i = 0$ for $i \ne j$ and $Y_j = X$, the last inequality reduces to the equality.

Example 11 (dam proposal selection): Suppose that n agencies participate in making a choice between two dam proposals A and B, discussed in the introduction: dam A has a constant net value P_A, whereas dam B is risky with the uncertain net value P_B. Suppose that the agencies share construction and maintenance costs of the selected dam as well as any potential losses due to an earthquake. Let agency i have the utility functional

$U_i(\cdot) = E[\cdot] - \rho_i\, \sigma_-(\cdot)$ with $\rho_i > 0$, $i = 1, \ldots, n$, and let $\rho_{\min} = \min_i \rho_i$, then

(i) Proposition 13 implies that the coalition chooses the risky proposal B if and only if $E[P_B] - \rho_{\min}\, \sigma_-(P_B) > P_A$ or, equivalently, $\max_i (E[P_B] - \rho_i\, \sigma_-(P_B)) > P_A$, which means that if at least one participating agency prefers dam B then so does the coalition. At first glance, this conclusion seems to be counterintuitive, since we would think that the coalition chooses dam B only if the majority of the participating agencies prefer dam B. However, it has the following explanation: if just one agency j prefers dam B then it agrees to cover the possible loss $E[P_B]-P_B$, provided that all other agencies pay it a total fixed amount $P > \rho_j\, \sigma_-(P_B)$. If, in addition, $P < E[P_B] - P_A$, then every agency is better off with the choice of dam B.

(ii) Proposition 12 implies that the mean-deviation model $E[\cdot] - \rho_{\min}\, \sigma_-(\cdot)$ conforms to all the axioms of the rational choice except for possibly monotonicity. If, in addition, $\rho_{\min} \le 1$, then the cooperative preference relation is monotone: the coalition never chooses dam A if $P_B \ge P_A$ with probability 1.

5.2 Portfolio Selection

Rockafellar et al.[13,14,15] solved and analyzed a Markowitz-type portfolio selection problem: minimize a deviation measure $\mathcal{D}$ of the portfolio return X subject to a constraint on the expected return $E[X]$, i.e.

$$\min_{X\in\mathcal{F}} \mathcal{D}(X) \quad \text{subject to} \quad E[X] \ge \pi, \tag{13}$$

where $\mathcal{F}$ is a set of all feasible portfolio returns, and π is a desirable expected gain.

The formulation (13) is general enough and readily extends to not purely financial applications. For example, a portfolio may consist of investment projects in the construction and agricultural sectors under the risk of natural hazards and disasters, and a portfolio selection problem is to allocate an initial capital among different construction and agricultural investment projects, e.g. building construction, industrial construction, land settlement, agricultural extension, irrigation, soil conservation, etc.[1]

The problem (13) can be reconstituted in the form

$$\max_{X\in\mathcal{F}} V(E[X], \mathcal{D}(X)), \tag{14}$$

where V is given by $V(m,d) = -d$ for $m \ge \pi$ and $V(m,d) = -\infty$ otherwise. The function V is not continuous and does not satisfy V4 in Proposition 7. As a consequence, axioms A3 and A2 are violated. Indeed, the constraint in (13) implies that an investor with the desired expected gain π rejects a project which guarantees *sure* profit $\pi - \epsilon$ for any $\epsilon > 0$. This is not a rational choice. Also, the model (13) fails to correctly order certain nonnegative r.v.'s. For example, if $X_1 = X$ and $X_2 = 2X$ with $X \ge 0$ and $E[X] = \pi$, then $X_2 \ge X_1$, but X_1 is preferred over X_2 because it has smaller deviation.

All these issues, however, can be easily overcome, if instead of choosing fixed π, the problem is to find the best trade off between the deviation and the expected value of an optimal return X. For any V satisfying V1–V3, if X^* is an optimal portfolio return in (14), then, obviously, X^* is also an optimal solution to (13) with $\pi = E[X^*]$. This suggests the following two-step portfolio selection procedure:

(i) Find an optimal portfolio return $X(\pi)$ in (13) for every $\pi > 0$.

(ii) Choose a continuous function V satisfying V1–V3 and find π^* maximizing $f(\pi) \equiv V(E[X(\pi)], \mathcal{D}(X(\pi)))$.

The one-parameter family of optimal solutions $X(\pi)$ is determined by Rockafellar et al.,[14] and step (ii) reduces to a one-parameter optimization problem. Proposition 7 implies that if $\mathcal{D}$ is weakly lower-range dominated with a constant K, and V satisfies V4 with the same constant K, then the induced preference relation satisfies A1–A5. This means that $X(\pi^*)$ is a rational choice. In particular, A2 implies that there is no portfolio return $Y \in \mathcal{F}$ such that $Y > X(\pi^*)$, whereas A5 implies that there is no $Y \in \mathcal{F}$ such that $Y \succ_2 X(\pi^*)$.

ACKNOWLEDGMENT

We are grateful to the Managing Editor, Dr. Karen Lowrie, for the valuable suggestions, and to the Area Editor, Dr. Tony Cox, for the encouraging comments.

REFERENCES

1. Department of Regional Development and Environment Executive Secretariat for Economic and Social Affairs Organization of American States. Primer on natural hazard management in integrated regional development planning, Washington, D.C., 1991.
2. von Neumann J, Morgenstern O. Theory of Games and Economic Behavior, 3rd ed., Princeton, NJ: Princeton University Press, 1953.
3. Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica, 1979; 47:263–291.
4. Machina M. Choice under uncertainty: problems solved and unsolved. Journal of Economic Perspectives, 1987; 1:121–154.
5. Yaari ME. The dual theory of choice under risk. Econometrica, 1987; 55:95–115.
6. Quiggin J. A theory of anticipated utility. Journal of Economic Behavior and Organization, 1982; 3:323–343.
7. Quiggin J. Generalized Expected Utility Theory – The Rank-Dependent Expected Utility Model, Dordrecht: Kluwer Academic Publishers, 1993.
8. Artzner P, Delbaen F, Eber JM, Heath D. Coherent measures of risk. Mathematical Finance, 1999; 9:203–227.
9. Föllmer H, Schied A. Stochastic finance, 2nd ed., Berlin New York: de Gruyter, 2004.
10. Rockafellar RT, Uryasev S, Zabarankin M. Generalized deviations in risk analysis. Finance and Stochastics, 2006; 10(1):51–74.
11. Rockafellar RT, Uryasev S. Conditional value-at-risk for general loss distributions. Journal of Banking & Finance, 2002; 26(7):1443–1471.
12. Rockafellar RT, Uryasev S. Optimization of conditional value-at-risk. Journal of Risk, 2000; 2:21–41.
13. Rockafellar RT, Uryasev S, Zabarankin M. Optimality conditions in portfolio analysis with general deviation measures. Mathematical Programming, 2006; 108(2-3):515–540.
14. Rockafellar RT, Uryasev S, Zabarankin M. Master funds in portfolio analysis with general deviation measures. The Journal of Banking and Finance, 2006; 30(2):743–778.
15. Rockafellar RT, Uryasev S, Zabarankin M. Equilibrium with investors using a diversity of deviation measures. The Journal of Banking and Finance, 2007; 31(11):3251–3268.
16. Rockafellar RT, Uryasev S, Zabarankin M. Risk tuning with generalized linear regression. Mathematics of Operations Research, 2008; 33(3):712–729.
17. Alcantud J, Bosi G. On the existence of certainty equivalents of various relevant types. Journal of Applied Mathematics, 2003; 9:447–458.
18. Machina M, Schmeidler D. A more robust definition of subjective probability. Econometrica, 1992; 60(4):745–780.
19. Rothschild M, Stiglitz J. Increasing risk I: A definition. Journal of Economic Theory, 1970; 2(3):225–243.
20. Levy H. Stochastic Dominance, Boston-Dodrecht-London: Kluwer Academic Publishers, 1998.
21. Bauerle N, Muller A. Stochastic orders and risk measures: Consistency and bounds. Insurance: Mathematics and Economics, 2006; 38:132–148.
22. Guriev S. On microfoundations of the dual theory of choice. The Geneva papers on risk and insurance theory, 2001; 26(2):117–137.
23. Roell A. Risk aversion in Quiggin and Yaari's rank–order model of choice under uncertainty. The Economic Journal, 1987; 97:143–159.
24. Cerreia-Vioglio S, Maccheroni F, Marinacci M, Montrucchio L. Uncertainty averse preferences, 2009; preprint, Bocconi University.
25. Safra Z, Segal U. Constant risk aversion. Journal of Economic Theory, 1998; 83(1):19–82.
26. Dana RA. A representation result for concave Schur-concave functions. Mathematical Finance, 2005; 15(4):613–634.
27. Holt CA, Laury SK. Risk aversion and incentive effects. American Economic Review, 2002; 92(5):1644–1655.
28. Rockafellar RT, Uryasev S, Zabarankin M. Deviation measures in risk analysis and optimization, Tech. Rep. 2002-7, Industrial and Systems Engineering Department, University of Florida, 2002.
29. Krokhmal P. Higher moment risk measures. Quantitative Finance, 2007; 7:373–387.
30. Markowitz HM. Portfolio selection. The Journal of Finance, 1952; 7(1):77–91.
31. Ogryczak W, Ruszczyński A. Dual stochastic dominance and related mean-risk models. SIAM Journal on Optimization, 2002; 13(1):60–78.
32. Ogryczak W, Ruszczyński A. From stochastic dominance to mean-risk models: Semideviations as risk measures. European Journal of Operational Research, 1999; 116(1):33–50.
33. Ogryczak W, Ruszczyński A. On consistency of stochastic dominance and mean-semideviation models. Mathematical Programming, 2001; 89:217–232.
34. Grechuk B, Molyboha A, Zabarankin M. Cooperative games with general deviation measures. Mathematical Finance, 2011; to appear.
35. Maccheroni F, Marinacci M, Rustichini A, Taboga M. Portfolio selection with monotone mean-variance preferences. Mathematical Finance, 2009; 19(3):487–521.
36. Filipovic D, Kupper M. Monotone and cash-invariant convex functions and hulls. Insurance: Mathematics and Economics, 2007; 32(2):1–16.
37. Grechuk B, Molyboha A, Zabarankin M. Maximum entropy principle with general deviation measures. Mathematics of Operations Research, 2009; 34(2):445–467.

APPENDIX

Auxiliary Proposition

PROPOSITION 14: Let $X, Y \in \mathcal{L}^1(\Omega)$ be arbitrary r.v.'s, and let X' and Y' be comonotone r.v.'s, whose distributions coincide with those of X and Y, respectively. Then $X + Y \succcurlyeq_2 X' + Y'$.

Proof. Let $q_X(\alpha)$ be the quantile function of X and X', and let $q_Y(\alpha)$ be the quantile function of Y and Y'. Then, since X' and Y' are comonotone, the quantile function of $X' + Y'$ is $q_X(\alpha) + q_Y(\alpha)$; see Lemma 4.84 in Föllmer and Schied.[9] Rockafellar et al.[10] showed that the *conditional value-at-risk* (CVaR), defined by

$$\mathrm{CVaR}_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q_X(\beta)d\beta \qquad \text{(A.1)}$$

for $\alpha \in (0, 1]$, is a coherent risk measure for every

$\alpha \in (0,1]$. Thus, by R3,

$$\int_0^\beta q_{X+Y}(\alpha)d\alpha \geq \int_0^\beta q_X(\alpha)d\alpha + \int_0^\beta q_Y(\alpha)d\alpha = \int_0^\beta q_{X'+Y'}(\alpha)d\alpha,$$

and $X+Y \succcurlyeq_2 X'+Y'$ by Theorem 2.58 in Föllmer and Schied.[9]

This result is used to prove Propositions 2 and 6.

Proof of Proposition 2

Suppose that (a) holds. A1–A4 imply that there exists a continuous certainty equivalent $U(X)$. Then $\mathcal{R}(X) = -U(X)$ is a coherent risk measure specified in (b). Indeed, (i) $\mathcal{R}(0) = 0$, (ii) $X \succeq Y$ if and only if $\mathcal{R}(X) \leq \mathcal{R}(Y)$, and (iii) $X \sim -\mathcal{R}(X)$ for all X. This fact along with A4 implies finiteness of $\mathcal{R}$, and also, along with A5, proves law-invariance of $\mathcal{R}$. Then it follows from $X \sim -\mathcal{R}(X)$ and A9 that $X + C \sim -\mathcal{R}(X) + C$, so that $-\mathcal{R}(X+C) = -\mathcal{R}(X) + C$, which is R1. Also, $X \sim -\mathcal{R}(X)$ and A9 imply $\lambda X \sim -\lambda\mathcal{R}(X)$ for any $\lambda > 0$. This is equivalent to $\mathcal{R}(\lambda X) = \mathcal{R}(-\lambda\mathcal{R}(X)) = \lambda\mathcal{R}(X)$, which proves R2. R4 holds by A2.

To prove R3, we first assume that r.v.'s X and Y are comonotone. With A9, $X \sim -\mathcal{R}(X)$ and $Y \sim -\mathcal{R}(Y)$ imply $X + \mathcal{R}(X) \sim 0$ and $Y + \mathcal{R}(Y) \sim 0$, respectively. Then $X + \mathcal{R}(X) \sim Y + \mathcal{R}(Y)$, and by A8, $\lambda(X + \mathcal{R}(X)) + (1-\lambda)(Y + \mathcal{R}(Y)) \succeq 0$ for any $\lambda \in [0,1]$, so that $\mathcal{R}\big(\lambda X + (1-\lambda)Y + \lambda\mathcal{R}(X) + (1-\lambda)\mathcal{R}(Y)\big) \leq \mathcal{R}(0) = 0$ or, equivalently, $\mathcal{R}(\lambda X + (1-\lambda)Y) \leq \lambda\mathcal{R}(X) + (1-\lambda)\mathcal{R}(Y)$, which for $\lambda = 1/2$ reduces to R3.

If X and Y are not comonotone, then there exists an r.v. X' comonotone with Y and having the distribution of X (see Lemma 4.2 in Dana[26]). In this case, Proposition 14 implies $X + Y \succcurlyeq_2 X' + Y$, and therefore $X + Y \succeq X' + Y$ by Proposition 1. Consequently, $\mathcal{R}(X+Y) \leq \mathcal{R}(X'+Y) \leq \mathcal{R}(X') + \mathcal{R}(Y) = \mathcal{R}(X) + \mathcal{R}(Y)$. This finishes the proof of R3, and $\mathcal{R}(X)$ is a law-invariant continuous coherent risk measure.

Now we assume that (b) holds. A1 is obvious. A2 follows from R4 and R1, whereas A3 and A4 follow from the continuity and finiteness of $\mathcal{R}(X)$, respectively. Since every law-invariant continuous coherent risk measure is SSD-consistent, A5 holds by Proposition 1. A8 follows from R2 and R3. Finally, for $X \succeq Y$, the fact $\mathcal{R}(\lambda X) = \lambda\mathcal{R}(X) \leq \lambda\mathcal{R}(Y) = \mathcal{R}(\lambda Y)$ implies $\lambda X \succeq \lambda Y$, which together with R1 proves A9.

Proof of Proposition 4

Suppose that (b) holds. Then obviously, a mean-deviation preference relation satisfies A1 and A4. A3 on $\mathcal{L}^2(\Omega)$ follows from the continuity of $\mathcal{D}$ and $V(m,d)$. It is left to show that A2 holds as well.

Let $X \geq Y$ and $X \neq Y$, then $E[X] = E[Y] + a$ with $a > 0$. Since $\inf(X - Y) \geq 0$, the condition (7) implies $\mathcal{D}(X-Y) \leq K \cdot (E[X-Y] - \inf(X-Y)) \leq K \cdot E[X-Y] = K \cdot a$. Consequently, $\mathcal{D}(X) \leq \mathcal{D}(Y) + \mathcal{D}(X-Y) \leq \mathcal{D}(Y) + K\cdot a$, and with V4 and V2, we obtain

$$V(E[Y],\mathcal{D}(Y)) \leq V(E[Y]+a, \mathcal{D}(Y) + K\cdot a) \leq V(E[X],\mathcal{D}(X)),$$

so that the mean-deviation preference relation satisfies A2.

Now suppose (a) holds. To show continuity of V at some point (m,d), $d \geq 0$, consider a sequence (m_n, d_n) with $d_n \geq 0$ such that $m_n \to m$ and $d_n \to d$. For an r.v. X with $E[X] = 0$ and $\mathcal{D}(X) = 1$, the sequence $Y_n = d_n X + m_n$ converges to $Y = dX + m$ in $\mathcal{L}^2(\Omega)$, and thus, $U(Y_n) \to U(Y)$. On the other hand, (5) implies that $U(Y) = V(E[Y],\mathcal{D}(Y)) = V(m,d)$ and $U(Y_n) = V(E[Y_n],\mathcal{D}(Y_n)) = V(m_n,d_n)$, and consequently, the continuity of V follows.

Next we show that V4 holds. The condition (7) implies that for any $\epsilon > 0$, there exists an r.v. $X_1 \in \mathcal{L}^2(\Omega)$ such that

$$K \geq \frac{\mathcal{D}(X_1)}{E[X_1] - \inf X_1} > K - \epsilon.$$

We can assume that $K - \epsilon > 0$, and thus, $\inf X_1 > -\infty$. Let m, $d \geq 0$, and $a \geq 0$ be given. Then the r.v.'s

$$X_2 = a\frac{X_1 - \inf X_1}{E[X_1] - \inf X_1}, \qquad X_3 = \frac{d}{K\cdot a}X_2 + m - \frac{d}{K},$$

are such that $X_2 \geq 0$, $E[X_2] = a$, $a \cdot K \geq \mathcal{D}(X_2) > a(K-\epsilon)$, $E[X_3] = m$, and $\mathcal{D}(X_3) \leq d$. Also, $E[X_2 + X_3] = m + a$ and $\mathcal{D}(X_2 + X_3) = (d/(K\cdot a) + 1)\mathcal{D}(X_2) > d + aK - \epsilon(d/K + a)$.

Since $X_2 + X_3 \geq X_3$, A2 implies $U(X_2 + X_3) \geq U(X_3)$, and consequently, $V(E[X_3],\mathcal{D}(X_3)) \leq V(E[X_2+X_3],\mathcal{D}(X_2+X_3))$. By V2,

$$\begin{aligned} V(m,d) &\leq V(E[X_3],\mathcal{D}(X_3)) \\ &\leq V(E[X_2+X_3],\mathcal{D}(X_2+X_3)) \\ &< V\left(m+a, d+aK-\epsilon\left(d/K+a\right)\right). \end{aligned}$$

Finally, passing $\epsilon \to 0$ and using the continuity of V, we obtain $V(m,d) \leq V(m+a, d+a\cdot K)$.

Proof of Proposition 5

First, we prove (a). Continuity of $\phi(\lambda)$ follows from the continuity of U, and obviously, $\phi(0) = 0$ holds. It is left to prove that $\phi(\lambda)$ is strictly increasing. A5 in strict sense implies $0 = E[\lambda X_0] \succ \lambda X_0$ for any $\lambda > 0$, so that $U(\lambda X_0) < U(0) = 0$ or $\phi(\lambda) > 0$. Then, by Theorem 2.58 in Föllmer and Schied,[9] $\lambda_2 X_0 \succcurlyeq_2 \lambda_1 X_0$ for any $\lambda_1 > \lambda_2 > 0$, and by Proposition 1, $\lambda_2 X_0 \succeq \lambda_1 X_0$, so that $\phi(\lambda_1) \geq \phi(\lambda_2)$. To prove that the inequality is strict, we assume the contrary, i.e., $\phi(\lambda_1) = \phi(\lambda_2) > 0$. In this case, $U(\lambda_1 X_0) = U(\lambda_2 X_0)$ or $\lambda_1 X_0 \sim \lambda_2 X_0$. Since $E[\lambda_1 X_0] = E[\lambda_2 X_0] = 0$, A11 implies $\alpha\lambda_1 X_0 \sim \alpha\lambda_2 X_0$ for every $\alpha > 0$. Choosing $\alpha_0 = \lambda_2/\lambda_1 < 1$, we obtain $\lambda_2 X_0 = \alpha_0\lambda_1 X_0 \sim \alpha_0\lambda_2 X_0 = \alpha_0^2\lambda_1 X_0 \sim \alpha_0^2\lambda_2 X_0 = \ldots$, and by induction, $\lambda_2 X_0 \sim \alpha_0^n\lambda_2 X_0$ for all $n \in N$. But since $\alpha_0 < 1$, $\alpha_0^n\lambda_2 X_0 \to 0$ as $n \to \infty$, and by A3, $\lambda_2 X_0 \sim 0$, which contradicts A5 in strict sense.

Now we prove (b). The properties of ϕ in (a) guarantee existence of an inverse function $\phi^{-1} : [0, M) \to [0, \infty)$ with $M = \sup_{\lambda\geq 0}(\phi(\lambda))$. Thus, to prove that $\mathcal{D}(X)$ is defined for every X, we should show that $-U\left(\overline{X}\right) \in [0, M)$. The inequality $-U\left(\overline{X}\right) \geq 0$ follows from A5 and the fact that $E[X] \succcurlyeq_2 X$. Next we show that $-U\left(\overline{X}\right) < M$. The continuity of U and $U(0) = 0$ imply that for every X, there exists $\alpha_0 > 0$ such that $-U(\alpha_0\overline{X}) < M$. Thus, there exists $\lambda_0 > 0$ such that $-U(\alpha_0\overline{X}) = \phi(\lambda_0) = -U(\lambda_0 X_0)$ or, equivalently, $\alpha_0\overline{X} \sim \lambda_0 X_0$. The equality $E[\alpha_0\overline{X}] = E[\lambda_0 X_0] = 0$ and A11 result in $\overline{X} \sim (\lambda_0/\alpha_0)X_0$ or $-U\left(\overline{X}\right) = -U((\lambda_0/\alpha_0)X_0) = \phi(\lambda_0/\alpha_0) < M$. This proves that $\mathcal{D}(X) = \phi^{-1}(-U\left(\overline{X}\right))$ is well-defined.

The law invariance of $\mathcal{D}$ follows from A5, whereas the continuity of ϕ and U imply that $\mathcal{D}$ is continuous.