

CDE  
November 2010

# **COMPARATIVE RISKINESS OF RANDOM VECTOR OUTCOMES**

**Sudhir A. Shah**  
Email: [sudhir@econdse.org](mailto:sudhir@econdse.org)  
**Delhi School of Economics**  
**University of Delhi**

**Working Paper No. 191**

**Centre for Development Economics**  
Department of Economics, Delhi School of Economics

# Comparative riskiness of random vector outcomes

Sudhir A. Shah\*

November 18, 2010

## Abstract

Consider lotteries  $\mu$  and  $\lambda$  with vector outcomes. Let  $\succeq_1$  be the relation that declares  $\mu$  to be riskier than  $\lambda$  if every risk averse decision-maker prefers  $\lambda$  to  $\mu$ . We introduce two other notions of the comparative riskiness of lotteries with vector outcomes. The first compares  $\mu$  and  $\lambda$  in terms of the expected utilities generated by a family of real-outcome lotteries derived from  $\mu$  and  $\lambda$  and a family of sufficiently smooth utilities defined over the real line. The second compares the riskiness of  $\mu$  and  $\lambda$  in terms of a generalized form of second-order stochastic dominance. Our main result is that these relations are identical to  $\succeq_1$  for every vector outcome space. This result complements classical characterizations of  $\succeq_1$  in terms of dilatations and martingales, thereby leading to a more complete theory of comparative vector riskiness. As our result admits outcomes in infinite dimensional vector spaces, it brings a host of new applications within the ambit of the theory. We illustrate this with applications to the theories of auctions, utility regulation, public goods, portfolio selection, multi-party moral hazard and teams.

JEL classification: C02, D01, D63, D81

Key words: vector outcome, comparative riskiness, generalized stochastic dominance

## 1 Introduction

Hadar and Russell [8], Hanoch and Levy [9] and Rothschild and Stiglitz [18] posed the question: When is one probability measure (henceforth, lottery) on the real line  $\mathfrak{R}$  riskier than another one? They provided three alternative formulations of “lottery  $\mu$  is riskier than lottery  $\lambda$ ”: (a) every risk averse expected utility maximizing decision-maker weakly prefers  $\lambda$  to  $\mu$ , (b)  $\mu$  and  $\lambda$  have the same mean and  $\lambda$  second order stochastically dominates  $\mu$ , and (c) there exist random variables  $f$  and  $g$  with distributions  $\lambda$  and  $\mu$  respectively, such that the expectation of  $g$  conditional on  $f$  is identical to  $f$

---

\*Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007, India. Email: sudhir@econ.dse.org

almost surely, i.e., the random process  $(f, g)$  is a martingale.<sup>1</sup> The principal result of the real outcomes theory is: if we consider lotteries with supports in a compact interval  $O$  in  $\mathfrak{R}$ , then all three criteria induce the same partial order on the admissible set of lotteries.

Given that economic problems with random vector outcomes are commonplace, it is natural to seek a generalization of the real outcomes theory to the situation where the set of outcomes  $O$  is a subset of a vector space  $X$ . Although it is tempting to finesse the problem by aggregating the vector outcomes into real outcomes and assuming that the resulting real risks are comparable, we shall argue in Section 3 that this strategy is unattractive.

### The vector riskiness literature

As real outcomes are not intrinsic to criteria (a) and (c), their definitions readily extend to the vector outcomes setting with only minor modifications. When  $O$  is a subset of a vector space  $X$ , it is known that for large classes of topological vector spaces  $X$  and outcome sets  $O$ , criterion (a) is equivalent to the existence of a suitable dilatation on  $O$ , and that criterion (a) is equivalent to criterion (c); see Theorems 4.2 and 4.5 in Appendix B for these results.<sup>2</sup>

On the other hand, as real outcomes are intrinsic to criterion (b), extension of this criterion to a vector space setting requires a non-trivial modification of its definition. The problem of extending criterion (b) to vector spaces and connecting it to criterion (a) has attracted some attention (Russell and Seo [20], Olson [13]).

Both these papers restrict  $X$  to be the Euclidean space  $\mathfrak{R}^n$  and  $O$  to be a convex compact subset of  $X$ . The Russell-Seo definition of the generalized second order stochastic dominance of a vector outcomes lottery  $\mu$  by another vector outcomes lottery  $\lambda$  amounts to checking that the distribution of utility  $\lambda \circ v^{-1}$  on  $\mathfrak{R}$  second order stochastically dominates the distribution of utility  $\mu \circ v^{-1}$  on  $\mathfrak{R}$  for every  $v \in \mathcal{V}_c$ , where  $\mathcal{V}_c$  is the set of functions  $v : O \rightarrow \mathfrak{R}$  that are increasing, concave and upper semicontinuous. The main Russell-Seo result is that criterion (a) is equivalent to their generalization of criterion (b) when  $X = \mathfrak{R}^n$  (see Remark 1 in Russell and Seo [20]). Olson [13] also contains an equivalence result, but compared to the Russell-Seo generalization, the Olson generalization of criterion (b) is opaque in terms of interpretation and difficult to apply because its verification intrinsically involves properties of certain non-denumerable families of sets.

---

<sup>1</sup>Second order stochastic dominance of  $\mu$  by  $\lambda$  means that, for every real number  $r$ , the area under the distribution function of  $\mu$  over the interval  $(-\infty, r]$  exceeds the area under the distribution function of  $\lambda$  over the same interval.

<sup>2</sup>Given its evolution, the characterization of criterion (a) in terms of dilatations may be referred to as the Hardy-Littlewood-Pólya-Blackwell-Stein-Sherman-Cartier-Fell-Meyer theorem. The equivalence of criteria (a) and (c) may be called the Blackwell-Stein-Sherman theorem. As is evident from the attributions, these results have a rich genealogy.

## Our contribution

We make two contributions to the literature described above.

First, we generalize the Russell-Seo equivalence result from the setting  $X = \mathfrak{R}^n$  to *all* vector spaces  $X$ . The only requirements on the set of outcomes are that  $O$  be a convex and measurable subset of  $X$ . This is the most general setting in which concave utilities can be specified and integrated to calculate their expectations. An implication of  $X$  not having topological or ordering structure is that, unlike in the Russell-Seo method, the notion of a utility  $v : O \rightarrow \mathfrak{R}$  being increasing and semicontinuous is not meaningful. This result exceeds, in scope and generality, the classical characterizations of criterion (a) in terms of dilatations and martingales and the Russell-Seo-Olson characterizations of criterion (a) in terms of generalizations of criterion (b), thereby substantially completing the extension of the real outcomes theory to general vector spaces.

Second, we show that the comparative vector riskiness theory enables the derivation of very general and economically meaningful predictions in a variety of applications. Although parts of the vector riskiness theory are quite classical, there is a notable absence of its applications in the literature. The applications studied in this paper begin to fill this gap.<sup>3</sup> As applications, we analyze and predict the effects of variations in vector riskiness in six settings: bidding in sealed bid auctions, the size of regulated utilities, the allocation of public goods, portfolio choice, multi-party moral hazard and team theory.

Our main theoretical results (Theorems 2.6 and 2.8) are: if  $O$  is a convex and measurable subset of a vector space  $X$ , then criterion (a) is equivalent to two other criteria for the comparison of vector-valued risks (Definitions 2.3 and 2.7). Of these other criteria, the first one is new, while the second is essentially the Russell-Seo generalization of criterion (b) that is adjusted to suit our more abstract setting; the modification involves expanding the set of utilities  $\mathcal{V}_c$  used in the Russell-Seo theory to  $\mathcal{V}$  which includes all function  $v : O \rightarrow \mathfrak{R}$  that are bounded, concave and measurable. We also derive some general implications of our results, such as the aggregation properties of the comparative vector riskiness relation (Corollary 2.11) and weak mean-preservation (Theorem 2.13).

## Choice of techniques

The Russell-Seo method of proving their result does not generalize to our setting. The reason is that the key step in their method is the construction of joint distribution functions on  $\mathfrak{R}^{n+1}$ , which intrinsically involves the usual

---

<sup>3</sup>However, an alternative interpretation of the comparative riskiness formalism has served as a starting point for the theory of comparative inequality of empirical distributions of multiple attributes (e.g., Kolm [10], Atkinson and Bourguignon [1], Tsui [22]).

ordering of  $\mathfrak{R}^n$  and its topological and measure-theoretic properties; the algebraic restriction in this method is already implicit in the fact that  $X = \mathfrak{R}^n$  is finite dimensional. Obviously, this method is unsuitable when  $\mathfrak{R}^n$  is replaced by an abstract vector space that does not possess the convenient ordering-algebraic-topological properties of Euclidean spaces.

However, we can work with general vector spaces because, unlike the Russell-Seo method, our arguments only involve distribution functions on  $\mathfrak{R}$ . Essentially, the utility functions in  $\mathcal{V}$  serve as a system of coordinates and the key observation is that all the necessary arguments to establish the equivalence result can be done coordinate-by-coordinate. So, instead of assuming and using the ordering-algebraic-topological properties of  $X$ , we transfer all arguments to  $\mathfrak{R}$  and feed off the ordering-algebraic-topological properties of  $\mathfrak{R}$ . This approach allows a result that is very general and technical arguments that are very elementary as they are confined to  $\mathfrak{R}$ .

The rest of this paper is organized as follows. Section 2 contains our general theoretical results. Section 3 contains applications of these results. We conclude in Section 4. Proofs of lemmas are collected in Appendix A. Appendix B contains statements of the classical equivalence results in the vector outcomes setting.

## 2 Main results

The following conventions and notation apply throughout this paper.  $\mathcal{N}$  is the set of natural numbers. The real line  $\mathfrak{R}$  is given the Euclidean metric topology and a topological space is given the Borel  $\sigma$ -algebra. Given topological spaces  $A$  and  $B$ , the set of continuous functions  $f : A \rightarrow B$  is denoted by  $\mathcal{C}(A, B)$ . A function  $f$  is said to be  $\mathcal{C}^k$  if it is  $k$ -times continuously differentiable for  $k \in \mathcal{N} \cup \{\infty\}$ . Consider  $u : A \rightarrow \mathfrak{R}$  with  $A \subset \mathfrak{R}$ .  $u$  is said to be increasing (resp. decreasing) if  $x, y \in A$  and  $x > y$  implies  $u(x) \geq u(y)$  (resp.  $u(x) \leq u(y)$ ).  $D^n u(x)$  (resp.  $u(x+)$ ,  $u(x-)$ ) denotes the  $n$ -th order derivative (resp. right-hand limit, left-hand limit) of  $u$  at  $x \in A$ . The symbol  $\neg P$  will denote the negation of a sentence  $P$ .

The following hypotheses hold throughout the paper.

**Assumption 2.1**  *$(O, \mathcal{O})$  is a measurable space, where  $O$  is a nonempty convex subset of a vector space  $X$  and  $\mathcal{O}$  is a  $\sigma$ -algebra on  $O$ .  $\mathcal{V}$  is the set of functions  $v : O \rightarrow \mathfrak{R}$  that are bounded, measurable and concave.*

Given  $v \in \mathcal{V}$ , let  $v_0 = \inf v(O)$  and  $v_1 = \sup v(O)$ . As  $v$  is bounded,  $v_0, v_1 \in \mathfrak{R}$ . Additional restrictions on  $O$  and  $X$  will be stated explicitly when required.  $O$  is the set of outcomes and  $\Delta(O)$  denotes the set of countably additive lotteries on  $(O, \mathcal{O})$ . Clearly, every  $v \in \mathcal{V}$  is integrable with respect to every  $\mu \in \Delta(O)$ .

## Equivalence results

The first relation on  $\Delta(O)$  declares a lottery  $\mu$  to be riskier than a lottery  $\lambda$ , denoted by  $\mu \succeq_1 \lambda$ , if every risk averse decision-maker prefers  $\lambda$  to  $\mu$ .<sup>4</sup>

**Definition 2.2** *Given  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_1 \lambda$  if  $\int_O \mu(dx) v(x) \leq \int_O \lambda(dx) v(x)$  for every  $v \in \mathcal{V}$ , i.e.,  $\lambda$  concave-dominates  $\mu$  over  $O$ .*

The hypothesized unanimity of risk averse preferences can be controlled by expanding or contracting the set of admissible functions  $\mathcal{V}$ . For instance, if  $\mathcal{V}$  is contracted by requiring greater regularity or monotonicity of its elements, then  $\succeq_1$  expands but the stronger requirements on  $v \in \mathcal{V}$  correspondingly contract the set of potential applications of the theory.

Relation  $\succeq_1$  is pivotal in the theory of comparative riskiness. One reason for this is that its definition requires no modification of the corresponding scalar definition other than the replacement of  $\mathfrak{R}$  by a vector space  $X$ . Second, it has obvious decision-theoretic and welfare-theoretic interpretations. Finally, well-known results such as Theorems 4.2 and 4.5 provide characterizations of  $\succeq_1$ .

In order to motivate the next definition, first consider the following easily verifiable characterization: given  $\mu, \lambda \in \Delta(O)$ ,  $\mu \succeq_1 \lambda$  if and only if  $v \in \mathcal{V}$  implies  $\int_O \mu(dx) u \circ v(x) \leq \int_O \lambda(dx) u \circ v(x)$  for every concave, increasing  $\mathcal{C}^2$  function  $u : [v_0, \infty) \rightarrow \mathfrak{R}$ . Using the change-of-variable formula (Billingsley [4], Theorem 16.13), we may re-write this characterization as: for  $\mu, \lambda \in \Delta(O)$ ,  $\mu \succeq_1 \lambda$  if and only if  $v \in \mathcal{V}$  implies

$$\int_{v(O)} \mu \circ v^{-1}(dy) u(y) \leq \int_{v(O)} \lambda \circ v^{-1}(dy) u(y) \quad (1)$$

for every concave, increasing  $\mathcal{C}^2$  function  $u : [v_0, \infty) \rightarrow \mathfrak{R}$ . Some interpretational remarks regarding this characterization are in order.

First, (1) suggests that comparing the riskiness of  $\mu$  and  $\lambda$  in  $\Delta(O)$  amounts to comparing, for every utility  $v \in \mathcal{V}$ , the riskiness of the corresponding utility distributions  $\mu \circ v^{-1}$  and  $\lambda \circ v^{-1}$  in terms of the class of preferences over real outcomes lotteries represented by the functions  $u$ . However, also note that, as each admissible  $u$  must be increasing and  $\mathcal{C}^2$ , the set of admissible  $u$  is a proper subset of the set of all bounded, measurable and concave functions on  $[v_0, \infty)$ . Therefore, (1) does not mean that  $\lambda \circ v^{-1}$  concave-dominates  $\mu \circ v^{-1}$  over  $[v_0, \infty)$ .

Alternatively, in this characterization and Definition 2.3,  $v$  may be interpreted as a statistic or aggregation function instead of a utility, while  $u$  may be interpreted as a utility; this interpretation is used in all the applications

---

<sup>4</sup>In the context of welfare economics,  $\mu \succeq_1 \lambda$  represents the notion that distribution  $\mu$  is more unequal than distribution  $\lambda$  because an egalitarian social planner, i.e., one with a concave welfare function, would prefer  $\lambda$  to  $\mu$ .

in Section 3. The next definition modifies the set of admissible  $u$  functions in this characterization.

Consider  $v \in \mathcal{V}$ . Let  $\mathcal{U}(v)$  be the set of concave, increasing  $\mathcal{C}^2$  functions  $u : [v_0, \infty) \rightarrow \mathfrak{R}$  such that  $Du(v_0) < \infty$  and  $Du(v_1) = 0$ . Let  $\overline{\mathcal{U}(v)}$  be the set of functions  $u : [v_0, \infty) \rightarrow \mathfrak{R}$ , where  $u$  is the uniform limit of a sequence of functions in  $\mathcal{U}(v)$ .

**Definition 2.3** *Given  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_2 \lambda$  if  $v \in \mathcal{V}$  implies  $\int_O \mu(dx) u \circ v(x) \leq \int_O \lambda(dx) u \circ v(x)$  for every  $u \in \overline{\mathcal{U}(v)}$ .*

$\mu \succeq_2 \lambda$  requires that (1) holds for every  $v \in \mathcal{V}$  and every  $u \in \overline{\mathcal{U}(v)}$ . Therefore, the interpretational remarks regarding (1) also apply to this definition. We prepare for our equivalence results with a pair of lemmas. The first one collects a number of facts regarding  $u \in \overline{\mathcal{U}(v)}$ .

**Lemma 2.4** *If  $v \in \mathcal{V}$  and  $u \in \overline{\mathcal{U}(v)}$ , then  $u$  is concave, increasing, continuous, measurable and bounded, with  $u(v(O)) \subset u([v_0, \infty)) \subset [u(v_0), u(v_1)]$  and  $u(x) = u(v_1)$  for every  $x \in (v_1, \infty)$ .*

The next lemma shows that mappings such as  $x \mapsto \min\{x, 0\}$  can be uniformly approximated by smooth functions with certain useful properties. The approximating functions are constructed by joining various smooth functions with care taken to match them smoothly at boundary points.

**Lemma 2.5** *Let  $v_0, b \in \mathfrak{R}$  and  $\epsilon > 0$  be such that  $v_0 < b - 2\epsilon$ . If  $u : [v_0, \infty) \rightarrow \mathfrak{R}$  is given by  $u(x) = \min\{x - b, 0\}$ , then there exists a concave, increasing  $\mathcal{C}^\infty$  function  $h : [v_0, \infty) \rightarrow \mathfrak{R}$  such that  $\sup\{|h(x) - u(x)| \mid x \in [v_0, \infty)\} \leq \epsilon$ ,  $h \geq u$  on  $[v_0, \infty)$ ,  $Dh = 1$  on  $[v_0, b - 2\epsilon]$ ,  $Dh \in (0, 1)$  on  $(b - 2\epsilon, b)$  and  $h = Dh = 0$  on  $[b, \infty)$ .*

Lemma 2.5 implies that the function  $x \mapsto \min\{x - b, 0\}$  belongs to  $\overline{\mathcal{U}(v)}$ . This fact is used to prove our equivalence results in this section and to analyze the size of a regulated utility in Section 3.

The following equivalence result transforms the problem of comparing the riskiness of lotteries on  $O \subset X$  to that of comparing specified families of their image measures on  $v(O) \subset \mathfrak{R}$ , i.e., it shifts the comparison problem from the vector domain to the real domain.

**Theorem 2.6** *Given Assumption 2.1,  $\succeq_1 = \succeq_2$ .*

*Proof.* Fix  $\mu, \lambda \in \Delta(O)$ . Suppose  $\neg\mu \succeq_2 \lambda$ . Then, there exists  $v \in \mathcal{V}$  and  $u \in \overline{\mathcal{U}(v)}$  such that  $\int_O \mu(dx) u \circ v(x) > \int_O \lambda(dx) u \circ v(x)$ . Using Lemma 2.4, it follows that  $u \circ v \in \mathcal{V}$ . Therefore,  $\neg\mu \succeq_1 \lambda$ .

Conversely, suppose  $\neg\mu \succeq_1 \lambda$ . Then, there exists  $v \in \mathcal{V}$  such that  $\int_O \mu(dx) v(x) > \int_O \lambda(dx) v(x)$ . Define  $u : [v_0, \infty) \rightarrow \mathfrak{R}$  by  $u(y) = \min\{y -$

$v_1, 0\}$ . For every  $n \in \mathcal{N}$ , Lemma 2.5 implies the existence of  $u_n \in \mathcal{U}(v)$  such that  $\sup\{|u_n(y) - u(y)| \mid y \in [v_0, \infty)\} \leq 1/n$ . As  $(u_n)$  converges uniformly to  $u$ , it follows that  $u \in \overline{\mathcal{U}(v)}$ . Since,  $v(x) \leq v_1$  for every  $x \in O$ ,  $u \circ v(x) = v(x) - v_1$  for every  $x \in O$ . Therefore,  $\int_O \mu(dx) u \circ v(x) = \int_O \mu(dx) v(x) - v_1 > \int_O \lambda(dx) v(x) - v_1 = \int_O \lambda(dx) u \circ v(x)$ . So,  $\neg\mu \succeq_2 \lambda$ . ■

Given  $\nu \in \Delta(\mathfrak{R})$ , define the distribution function  $F(\cdot, \nu) : \mathfrak{R} \rightarrow [0, 1]$  by  $F(x, \nu) = \nu(-\infty, x]$ . We now define the third relation on  $\Delta(O)$ .

**Definition 2.7** Given  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_3 \lambda$  if  $v \in \mathcal{V}$  implies  $\int_{(-\infty, y]} dx [F(x, \mu \circ v^{-1}) - F(x, \lambda \circ v^{-1})] \geq 0$  for every  $y \in \mathfrak{R}$ .

$\mu \succeq_3 \lambda$  is equivalent to the scalar second order stochastic dominance of  $\mu \circ v^{-1}$  by  $\lambda \circ v^{-1}$  for every  $v \in \mathcal{V}$ . As in Definition 2.3, the functions in  $\mathcal{V}$  may be interpreted as a system of coordinates that enable the comparative riskiness of  $\mu$  and  $\lambda$  to be evaluated by comparing the distributions of the real-valued coordinate functions.

**Theorem 2.8** Given Assumption 2.1,  $\succeq_2 = \succeq_3$ .

Proof. Consider  $\mu, \lambda \in \Delta(O)$  such that  $\mu \succeq_2 \lambda$ . Fix  $v \in \mathcal{V}$ . Given  $y \in \mathfrak{R}$ , define  $u_y : [v_0, \infty) \rightarrow \mathfrak{R}$  by  $u_y(x) = \min\{x - y, 0\}$ .

If  $y \leq v_0$ , then  $u_y(\cdot) = 0$ , and so  $u_y \in \mathcal{U}(v) \subset \overline{\mathcal{U}(v)}$ . Now consider  $y \in (v_0, v_1]$ . Then, there exists  $N \in \mathcal{N}$  such that  $y - 2/N > v_0$ . Consider  $n \in \mathcal{N}$  such that  $n \geq N$ . Then,  $y - 2/n \geq y - 2/N > v_0$ . By Lemma 2.5, there exists a concave, increasing  $\mathcal{C}^2$  function  $h_n : [v_0, \infty) \rightarrow \mathfrak{R}$  such that  $Dh_n(v_0) < \infty$ ,  $Dh_n(y) = 0$  and  $\sup\{|h_n(x) - u_y(x)| \mid x \in [v_0, \infty)\} \leq 1/n$ . As  $y \leq v_1$  and  $h_n$  is concave, increasing  $\mathcal{C}^2$ , it follows that  $Dh_n(v_1) = 0$ . So,  $h_n \in \mathcal{U}(v)$ . For  $n \in \mathcal{N}$  such that  $n < N$ , set  $h_n = h_N$ . By construction, the sequence  $(h_n)$  is in  $\mathcal{U}(v)$  and converges uniformly to  $u_y$ . So,  $u_y \in \overline{\mathcal{U}(v)}$ . Thus,  $u_y \in \overline{\mathcal{U}(v)}$  for every  $y \in (-\infty, v_1]$ .

Consider  $y \in (-\infty, v_1]$ . Then,  $\int_{(-\infty, y]} (x - y) dF(x, \lambda \circ v^{-1}) = \int_{\mathfrak{R}} (x - y) 1_{(-\infty, y]}(x) dF(x, \lambda \circ v^{-1}) = \int_{v(O)} \lambda \circ v^{-1}(dx) u_y(x)$ . Similarly,  $\int_{(-\infty, y]} (x - y) dF(x, \mu \circ v^{-1}) = \int_{v(O)} \mu \circ v^{-1}(dx) u_y(x)$ . As  $\mu \succeq_2 \lambda$  and  $u_y \in \overline{\mathcal{U}(v)}$ , (1) implies  $\int_{(-\infty, y]} (x - y) dF(x, \mu \circ v^{-1}) \leq \int_{(-\infty, y]} (x - y) dF(x, \lambda \circ v^{-1})$ . Integrating by parts, for every  $y \in (-\infty, v_1]$ ,

$$\int_{(-\infty, y]} dx F(x, \mu \circ v^{-1}) \geq \int_{(-\infty, y]} dx F(x, \lambda \circ v^{-1}) \quad (2)$$

Now consider  $y > v_1$ . Note that  $F(x, \mu \circ v^{-1}) = 1 = F(x, \lambda \circ v^{-1})$  for every  $x \in (v_1, y]$ . Therefore,  $\int_{(-\infty, y]} dx F(x, \mu \circ v^{-1}) = \int_{(-\infty, v_1]} dx F(x, \mu \circ v^{-1}) + y - v_1$  and  $\int_{(-\infty, y]} dx F(x, \lambda \circ v^{-1}) = \int_{(-\infty, v_1]} dx F(x, \lambda \circ v^{-1}) + y - v_1$ . Since (2) holds for  $y = v_1$ , we have  $\int_{(-\infty, y]} dx F(x, \mu \circ v^{-1}) \geq \int_{(-\infty, y]} dx F(x, \lambda \circ v^{-1})$ . Thus, (2) holds for every  $y > v_1$ . As (2) holds for every  $y \in \mathfrak{R}$ , we have  $\mu \succeq_3 \lambda$ .



Conversely, suppose  $\mu \succeq_3 \lambda$ . Fix  $v \in \mathcal{V}$  and consider  $u \in \mathcal{U}(v)$ . In order to simplify notation, define  $G : \mathfrak{R} \rightarrow \mathfrak{R}$  by  $G(\cdot) = F(\cdot, \lambda \circ v^{-1})$ . Using the integration-by-parts formula for Lebesgue-Stieltjes integrals (Rao [16], Exercise 11, pg 339), we have  $\int_{v(O)} \lambda \circ v^{-1}(dx) u(x) = \int_{v(O)} u(x) dG(x) = u(v_1+)G(v_1+) - u(v_0-)G(v_0-) - \int_{v(O)} dx Du(x)G(x)$ . As  $u$  is continuous,  $u(v_1+) = u(v_1)$ . As  $G$  is right-continuous and left-limited, we have  $G(v_1+) = G(v_1) = 1$  and  $G(v_0-) = 0$ . Therefore,

$$\int_{v(O)} \lambda \circ v^{-1}(dx) u(x) = u(v_1) - \int_{v(O)} dx Du(x)G(x) \quad (3)$$

Define  $H : \mathfrak{R} \rightarrow \mathfrak{R}$  by  $H(x) = \int_{(-\infty, x]} dy G(y)$ . By Lebesgue's differentiation theorem (Rao [16], Section 5.2, Theorem 7),  $H$  is differentiable Lebesgue almost everywhere, with  $DH = G$  at differentiability points. Therefore,  $\int_{v(O)} Du(x) dH(x) = \int_{v(O)} dx DH(x) Du(x) = \int_{v(O)} dx Du(x)G(x)$ . Integrating by parts,  $\int_{v(O)} Du(x) dH(x) = Du(v_1+)H(v_1+) - Du(v_0-)H(v_0-) - \int_{v(O)} dx D^2u(x)H(x)$ . As  $Du$  is continuous at  $v_1$  and  $u \in \mathcal{U}(v)$ , we have  $Du(v_1+) = Du(v_1) = 0$ . As  $G(y) = 0$  for every  $y < v_0$ , we have  $H(y) = 0$  for every  $y < v_0$ . Therefore,  $H(v_0-) = 0$ . As  $u \in \mathcal{U}(v)$ , we have  $Du(v_0-) < \infty$ . Consequently,

$$\int_{v(O)} dx Du(x)G(x) = - \int_{v(O)} dx D^2u(x)H(x) \quad (4)$$

Combining (3) and (4), we have

$$\int_{v(O)} \lambda \circ v^{-1}(dx) u(x) = u(v_1) + \int_{v(O)} dx D^2u(x) \int_{(-\infty, x]} dy F(y, \lambda \circ v^{-1})$$

Similarly,  $\int_{v(O)} \mu \circ v^{-1}(dx) u(x) = u(v_1) + \int_{v(O)} dx D^2u(x) \int_{(-\infty, x]} dy F(y, \mu \circ v^{-1})$ . As  $\mu \succeq_3 \lambda$  and  $D^2u \leq 0$ , we have  $\int_{v(O)} \lambda \circ v^{-1}(dx) u(x) \geq \int_{v(O)} \mu \circ v^{-1}(dx) u(x)$ .

Now consider  $u \in \overline{\mathcal{U}(v)}$ . Then, there exists a sequence  $(u_n) \subset \mathcal{U}(v)$  converging uniformly to  $u$ . By Lemma 2.4,  $u$  is bounded. As  $(u_n)$  converges uniformly to  $u$ , the  $u_n$  are uniformly bounded. As  $\int_{v(O)} \lambda \circ v^{-1}(dx) u_n(x) \geq \int_{v(O)} \mu \circ v^{-1}(dx) u_n(x)$  for every  $n$ , we have  $\int_{v(O)} \lambda \circ v^{-1}(dx) u(x) \geq \int_{v(O)} \mu \circ v^{-1}(dx) u(x)$  by applying the bounded convergence theorem (Billingsley [4], Theorem 16.5). Thus, by (1),  $\mu \succeq_2 \lambda$ .  $\blacksquare$

Combining Theorem 2 in Rothschild and Stiglitz [18] with Theorems 2.6 and 2.8, we have the following result.

**Corollary 2.9** *If Assumption 2.1 is specialized by setting  $X = \mathfrak{R}$  and assuming  $O$  is compact, then  $\succeq_I = \succeq_1 = \succeq_2 = \succeq_3$ , where  $\succeq_I$  is given by:  $\mu \succeq_I \lambda$  if and only if  $\int_{(-\infty, y]} dx [F(x, \mu) - F(x, \lambda)] \geq 0$  for every  $y \in \mathfrak{R}$  and  $\int_O \mu(dx) x = \int_O \lambda(dx) x$ .*

The compactness of  $O$  ensures that  $\mu$  and  $\lambda$  have finite means.

## Aggregation and mean-preservation

We address two inter-related questions in this section.

First, what is the relationship between the comparative riskiness of vector-outcome lotteries  $\mu, \lambda \in \Delta(O)$  and the comparative riskiness of real-outcome lotteries  $\mu \circ v^{-1}, \lambda \circ v^{-1} \in \Delta(v(O))$  for  $v \in \mathcal{V}$ ? The answer to this question will settle an issue raised in Section 1, namely: is it possible to aggregate vector risks into scalar risks without loss of generality?

Second, given the resemblance between  $\succeq_3$  and  $\succeq_I$ , do the relations comparing vector risks have a mean-preservation property analogous to that of the Rothschild-Stiglitz relations comparing scalar risks?

The next result helps to address the first question.

**Theorem 2.10** *Given Assumption 2.1, consider  $\mu, \lambda \in \Delta(O)$ .*

(A)  $\mu \circ v^{-1}$  and  $\lambda \circ v^{-1}$  have the same mean for every  $v \in \mathcal{V}$  if and only if  $\mu \succeq_1 \lambda$  and  $\lambda \succeq_1 \mu$ . This result also holds if  $\succeq_1$  is replaced by  $\succeq_2$  or  $\succeq_3$ .

(B)  $\mu \circ v^{-1} = \lambda \circ v^{-1}$  for every  $v \in \mathcal{V}$  if and only if  $\mu \succeq_3 \lambda$  and  $\lambda \succeq_3 \mu$ . This result also holds if  $\succeq_3$  is replaced by  $\succeq_1$  or  $\succeq_2$ .

*Proof.* Fix  $\mu, \lambda \in \Delta(O)$ .

(A)  $\mu \succeq_1 \lambda$  and  $\lambda \succeq_1 \mu$  if and only if  $\int_{v(O)} \mu \circ v^{-1}(dx) x = \int_O \mu(dx) v(x) = \int_O \lambda(dx) v(x) = \int_{v(O)} \lambda \circ v^{-1}(dx) x$  for every  $v \in \mathcal{V}$ . Theorems 2.6 and 2.8 imply the other cases.

(B) If  $\mu \circ v^{-1} = \lambda \circ v^{-1}$  for every  $v \in \mathcal{V}$ , then they have the same mean for every  $v \in \mathcal{V}$ . The result follows from (A).

Conversely, suppose  $\mu \succeq_3 \lambda$ ,  $\lambda \succeq_3 \mu$  and  $v \in \mathcal{V}$ . For every  $y \in \mathfrak{R}$ ,  $c(y) \equiv \int_{(-\infty, y]} dx [F(x, \mu \circ v^{-1}) - F(x, \lambda \circ v^{-1})] = 0$ . Therefore,  $Dc(y) = 0$  for every  $y \in \mathfrak{R}$ . If  $y$  is a continuity point of  $F(\cdot, \mu \circ v^{-1}) - F(\cdot, \lambda \circ v^{-1})$ , then  $0 = Dc(y) = F(y, \mu \circ v^{-1}) - F(y, \lambda \circ v^{-1})$  (Bartle [2], Theorem 31.8). Consider  $y$  that is a discontinuity point of  $F(\cdot, \mu \circ v^{-1}) - F(\cdot, \lambda \circ v^{-1})$ . The set of such discontinuity points is countable (Billingsley [4], Theorem 10.2). Consequently, there exists a sequence of continuity points  $(y_n) \subset (y, \infty)$  converging to  $y$ . As  $F(y_n, \mu \circ v^{-1}) - F(y_n, \lambda \circ v^{-1}) = 0$  for every  $n$  and distribution functions are right-continuous, we have  $F(y, \mu \circ v^{-1}) - F(y, \lambda \circ v^{-1}) = 0$ . It follows that  $F(\cdot, \mu \circ v^{-1}) = F(\cdot, \lambda \circ v^{-1})$ .

Let  $\mathcal{P} = \{(-\infty, y] \mid y \in \mathfrak{R}\}$ .  $\mathcal{P}$  is a  $\pi$ -system (Billingsley [4], Page 41). Let  $\mathcal{L} = \{E \in \mathcal{B}(\mathfrak{R}) \mid \mu \circ v^{-1}(E) = \lambda \circ v^{-1}(E)\}$ , where  $\mathcal{B}(\mathfrak{R})$  is the Borel  $\sigma$ -algebra on  $\mathfrak{R}$ .  $\mathcal{L}$  is a  $\lambda$ -system (Billingsley [4], Page 41). By the above argument,  $\mathcal{P} \subset \mathcal{L}$ . By Dynkin's theorem (Billingsley [4], Theorem 3.2),  $\mathcal{B}(\mathfrak{R}) = \sigma(\mathcal{P}) \subset \mathcal{L}$ , i.e.,  $\mu \circ v^{-1}(E) = \lambda \circ v^{-1}(E)$  for every  $E \in \mathcal{B}(\mathfrak{R})$ . Theorems 2.6 and 2.8 imply the other cases. ■

So, the distributions  $\mu \circ v^{-1}$  and  $\lambda \circ v^{-1}$  have the same mean for every  $v \in \mathcal{V}$ , i.e., coordinate-wise mean-preservation holds, if and only if the riskiness of  $\mu$  and  $\lambda$  is indistinguishable. The following is a consequence of this result.

**Corollary 2.11** *Given Assumption 2.1, consider  $\mu, \lambda \in \Delta(O)$ . If, for every  $v \in \mathcal{V}$ , either  $\mu \circ v^{-1} \succeq_I \lambda \circ v^{-1}$  or  $\lambda \circ v^{-1} \succeq_I \mu \circ v^{-1}$ , then  $\mu \succeq_1 \lambda$ ,  $\lambda \succeq_1 \mu$  and  $\mu \circ v^{-1} = \lambda \circ v^{-1}$  for every  $v \in \mathcal{V}$ .*

Proof. Suppose  $\mu$  and  $\lambda$  are as hypothesized. By the definition of  $\succeq_I$ , we have  $\int_{v(O)} \mu \circ v^{-1}(dx) x = \int_{v(O)} \lambda \circ v^{-1}(dx) x$  for every  $v \in \mathcal{V}$ . Theorem 2.10(A) implies  $\mu \succeq_1 \lambda$  and  $\lambda \succeq_1 \mu$ . Theorem 2.10(B) implies  $\mu \circ v^{-1} = \lambda \circ v^{-1}$  for every  $v \in \mathcal{V}$ . ■

Suppose  $\mu$  and  $\lambda$  are lotteries with vector outcomes and  $v_0 \in \mathcal{V}$  is an aggregator function. The aggregation question is whether there is a relationship between the comparability of riskiness of  $\mu$  and  $\lambda$  on the one hand and of  $\mu \circ v_0^{-1}$  and  $\lambda \circ v_0^{-1}$  on the other hand?

Clearly, if  $\mu \circ v_0^{-1} \succeq_I \lambda \circ v_0^{-1}$  or  $\lambda \circ v_0^{-1} \succeq_I \mu \circ v_0^{-1}$ , then no unambiguous conclusion can be drawn regarding the comparative riskiness of  $\mu$  and  $\lambda$ . Conversely, if  $\mu \succeq_1 \lambda$  and  $\neg \lambda \succeq_1 \mu$ , then Corollary 2.11 implies the existence of some  $v \in \mathcal{V}$  such that  $\neg \mu \circ v^{-1} \succeq_I \lambda \circ v^{-1}$  and  $\neg \lambda \circ v^{-1} \succeq_I \mu \circ v^{-1}$ , i.e., the riskiness of the real-outcome lotteries  $\mu \circ v^{-1}$  and  $\lambda \circ v^{-1}$  is not comparable. As there is no *a priori* guarantee that  $v \neq v_0$ , nothing definite can be concluded about the comparative riskiness of the real-outcome lotteries  $\mu \circ v_0^{-1}$  and  $\lambda \circ v_0^{-1}$  and any assumption in this respect entails loss of generality. Thus, comparability of the riskiness of the aggregate real-outcome lotteries  $\mu \circ v_0^{-1}$  and  $\lambda \circ v_0^{-1}$  is neither necessary nor sufficient for the comparability of the riskiness of the vector-outcome lotteries  $\mu$  and  $\lambda$ . In other words, one cannot finesse the comparative vector risk problem *via* aggregation without loss of generality.

Now consider our second question regarding mean-preservation. Given  $\mu, \lambda \in \Delta(O)$ , Theorem 2.10 asserts that coordinate-wise mean-preservation holds if and only if  $\mu$  and  $\lambda$  are indistinguishable in terms of their riskiness. In order to consider another form of mean-preservation that holds more generally, we require  $X$  to be a topological vector space. Let  $X^*$  be the topological dual of  $X$ , i.e., the set of all continuous linear functionals  $h : X \rightarrow \mathfrak{R}$ . Then, the weak mean  $m_\mu \in X$  of  $\mu \in \Delta(O)$  is defined to be the Pettis integral (see Pettis [15]), i.e.,  $m_\mu$  solves the equation  $h(m_\mu) = \int_O \mu(dx) h(x)$  for every  $h \in X^*$ .<sup>5</sup> We now state conditions that ensure the existence and uniqueness of  $m_\mu$ .

**Lemma 2.12** *Given Assumption 2.1, if  $X$  is a locally convex topological vector space,  $O$  is compact and metrizable, and  $\mu \in \Delta(O)$ , then  $m_\mu$  exists, is unique and  $m_\mu \in O$ .*

Finally, we show that weak mean-preservation is generally necessary for the comparability of riskiness of lotteries with vector outcomes.

---

<sup>5</sup>For special classes of vector spaces, this notion coincides with other standard integrals of vector-valued functions such as the Bochner integral. In the case of  $\mathfrak{R}^n$ -valued functions, the Pettis integral is nothing but the familiar coordinate-by-coordinate integral.

**Theorem 2.13** *Given Assumption 2.1, if  $X$  is a locally convex topological vector space,  $O$  is compact and metrizable, and  $\mu \succeq_1 \lambda$  (resp.  $\mu \succeq_2 \lambda$ ,  $\mu \succeq_3 \lambda$ ) for  $\mu, \lambda \in \Delta(O)$ , then  $m_\mu = m_\lambda$  and  $m_\mu \in O$ .*

Proof. For every  $h \in X^*$ , the restrictions of  $h$  and  $-h$  to  $O$  belong to  $\mathcal{V}$ . So,  $\mu \succeq_1 \lambda$  implies  $\int_O \mu(dx) h(x) = \int_O \lambda(dx) h(x)$  for every  $h \in X^*$ . By Lemma 2.12, there exists a unique vector  $m_\mu \in O$  such that  $h(m_\mu) = \int_O \mu(dx) h(x)$  for every  $h \in X^*$  and a unique vector  $m_\lambda \in O$  such that  $h(m_\lambda) = \int_O \lambda(dx) h(x)$  for every  $h \in X^*$ . Thus,  $h(m_\mu) = \int_O \mu(dx) h(x) = \int_O \lambda(dx) h(x) = h(m_\lambda)$  for every  $h \in X^*$ . Consequently,  $h(m_\mu - m_\lambda) = h(m_\mu) - h(m_\lambda) = 0$  for every  $h \in X^*$ . Local convexity of  $X$  implies  $m_\mu = m_\lambda$  (Dunford and Schwartz [7], Corollary V.2.13). Theorems 2.6 and 2.8 imply the other cases. ■

### 3 Applications

In applications with random vector outcomes, one might attempt to finesse the problem of modeling comparative vector riskiness by aggregating vector outcomes into real outcomes and using the real outcomes theory to model the comparative riskiness of the aggregate outcomes. This tack is problematic for the following reasons.

First, at the conceptual level, an economically meaningful and fully adequate aggregator might not exist. For instance, consider the social planner's public good provision problem studied below. Any reduction of vector outcomes in this problem to scalar outcomes would involve a representative agent and a composite good, thereby suppressing substantive features of the stated problem, e.g., the manner in which an agent aggregates allocation vectors into utilities and the manner in which the planner aggregates individual utilities and allocations into a social utility.

Second, at the formal level, it is evident from Theorem 2.10 and Corollary 2.11 that the comparability of vector risks cannot be transformed into the comparability of aggregated scalar risks without loss of generality because the former does not imply the latter and *vice versa*.

Third, given the wide range of economic applications we study in this section, aggregation may simply be unnecessary because the vector outcomes theory appears to be as tractable as the real outcomes theory.

The common element of all the applications we study in this section is Theorem 3.1, which uses the results of Section 2 to generalize the analogous comparative statics result in Diamond and Stiglitz [5] from the real to the vector outcomes case. This result predicts the direction of variation in a decision-maker's optimal behavior in a situation with random vector outcomes when the riskiness of the outcomes varies.

Let  $X$  and  $O$  be as in Section 2 and let  $v \in \mathcal{V}$  with  $v_0 < v_1$ . Consider a family of lotteries  $\{\mu(r) \in \Delta(O) \mid r \in \mathfrak{R}\}$  where riskiness increasing with  $r \in$

$[0, 1]$ , i.e.,  $r_1, r_2 \in [0, 1]$  and  $r_2 > r_1$  implies  $\mu(r_2) \succeq_3 \mu(r_1)$ . Let  $H : \mathfrak{R}^2 \rightarrow \mathfrak{R}$  and  $h : \mathfrak{R}^2 \rightarrow \mathfrak{R}_+$  be such that  $F(x, \mu(r) \circ v^{-1}) = H(x, r) = \int_{(-\infty, x]} dy h(y, r)$  for  $(x, r) \in \mathfrak{R}^2$ , i.e.,  $H(\cdot, r)$  and  $h(\cdot, r)$  are the distribution function and density function, respectively, corresponding to lottery  $\mu(r) \circ v^{-1} \in \Delta(v(O))$ .

Let  $C = [a, b] \subset \mathfrak{R}$ , with  $a < b$ , be the decision-maker's choice space and let  $u : \mathfrak{R}^2 \rightarrow \mathfrak{R}$  be the decision-maker's utility function. If  $x \in O$  is the random outcome and the decision-maker chooses  $c \in C$ , then the resulting utility is  $u(v(x), c)$ . Formally define  $U : \mathfrak{R}^2 \rightarrow \mathfrak{R}$  by  $U(r, c) = \int_O \mu(r)(dx) u(v(x), c)$ . If  $\mu(r)$  is the lottery generating the random outcome and the decision-maker chooses  $c \in C$ , then  $U(r, c)$  is the expected utility.

**Theorem 3.1** *Suppose  $X, O, v, \{\mu(r) \in \Delta(O) \mid r \in \mathfrak{R}\}, H, h, C, u$  and  $U$  are as given above. If*

- (a)  $H$  is  $\mathcal{C}^1$ ,  $H(v_0, \cdot) = 0$  and  $h$  is  $\mathcal{C}^2$ ,
  - (b)  $u$  is  $\mathcal{C}^2$ ;  $D_{22}u < 0$  on  $v(O) \times (a, b)$ ;  $D_{21}u(\cdot, c)$  is  $\mathcal{C}^1$  for every  $c \in C$ ; and  $D_1u(v_1, \cdot) = 0$ , and
  - (c)  $c : [0, 1] \rightarrow (a, b)$  satisfies  $U(r, c(r)) \geq U(r, c)$  for every  $(r, c) \in [0, 1] \times C$ ,
- then  $c$  is differentiable and  $D_{211}u \geq 0$  (resp.  $\leq 0$ ) implies  $Dc \geq 0$  (resp.  $\leq 0$ ).

*Proof.* Define  $\phi : \mathfrak{R} \times \mathfrak{R}^2 \rightarrow \mathfrak{R}$  by  $\phi(x, (r, c)) = h(x, r)u(x, c)$ . Using Theorem XIII.7.1 in Lang [11], as  $h$  and  $u$  are  $\mathcal{C}^2$ , so is  $\phi$ . As  $v(O)$  is bounded,  $\phi(\cdot, (r, c))$  is bounded on  $v(O)$ . Consequently,  $\int_{v(O)} dx \phi(x, (r, c))$  exists and  $U(r, c) = \int_{v(O)} dx \phi(x, (r, c))$ . As  $\phi$  and  $D_2\phi = (uD_2h, hD_2u) : \mathfrak{R} \times \mathfrak{R}^2 \rightarrow \mathfrak{R}^2$  are continuous, Theorem XIII.8.1 in Lang [11] implies that  $U$  is differentiable and  $DU(r, c) = \int_{v(O)} dx D_2\phi(x, (r, c))$ . Analogous arguments applied to the component functions of  $DU$  imply that  $DU$  is differentiable and  $D^2U(r, c) = \int_{v(O)} dx D_{22}\phi(x, (r, c))$ , i.e.,  $U$  is twice differentiable. Since  $D_{22}\phi$  is continuous on the compact set  $v(O) \times [0, 1] \times C$ , it is bounded on this set. Consequently,  $D^2U$  is continuous on  $[0, 1] \times C$  (Lang [11], Exercise VI.10.13). Therefore,  $U$  is  $\mathcal{C}^2$  on  $[0, 1] \times C$ .

Define  $\Phi : v(O) \times \mathfrak{R} \rightarrow \mathfrak{R}$  by  $\Phi(x, r) = \int_{[v_0, x]} dy H(y, r)$ ; note that  $\Phi(x, r) = \int_{(-\infty, x]} dy H(y, r)$  as  $H(y, \cdot) = 0$  for  $y \leq v_0$ . Using Assumption (a) and Theorem XIII.8.1 in Lang [11],  $\Phi(x, \cdot)$  is differentiable and  $D_2\Phi(x, \cdot) = \int_{[v_0, x]} dy D_2H(y, \cdot)$  for every  $x \in v(O)$ . As  $r_1, r_2 \in [0, 1]$  and  $r_2 > r_1$  implies  $\mu(r_2) \succeq_3 \mu(r_1)$ ,  $\Phi(x, \cdot)$  is increasing on  $[0, 1]$  for every  $x \in v(O)$ . Thus,  $D_2\Phi \geq 0$  on  $v(O) \times [0, 1]$ .

Consider  $r \in [0, 1]$ . As  $U$  is  $\mathcal{C}^2$  on  $[0, 1] \times C$ , Assumptions (b) and (c) imply  $D_2U(r, c(r)) = 0$  and  $D_{22}U(r, c(r)) = \int_{v(O)} dx h(x, r)D_{22}u(x, c(r)) < 0$ . As  $D_2U$  is  $\mathcal{C}^1$ , the implicit function theorem implies that  $c$  is  $\mathcal{C}^1$  in a neighborhood of  $r$  and the chain rule yields  $D_{21}U(r, c(r)) + D_{22}U(r, c(r))Dc(r) = 0$ . As  $D_{22}U(r, c(r)) < 0$ ,  $Dc(r)$  and  $D_{21}U(r, c(r))$  have the same sign.

Consider  $(r, c) \in [0, 1] \times C$ . Then,  $D_2U(r, c) = \int_{v(O)} dx h(x, r) D_2u(x, c) = D_2u(v_1, c) - \int_{v(O)} dx D_{21}u(x, c) H(x, r)$ . It follows from Assumptions (a) and (b), and Theorem XIII.8.1 in Lang [11], that  $\int_{v(O)} dx D_{21}u(x, c) H(x, \cdot)$  is differentiable and its derivative is  $\int_{v(O)} dx D_{21}u(x, c) D_2H(x, \cdot)$ . Thus,  $D_{21}U(r, c) = - \int_{v(O)} dx D_{21}u(x, c) D_2H(x, r)$ . As Assumption (b) implies  $D_{21}u(v_1, \cdot) = D_{12}u(v_1, \cdot) = 0$ , integrating by parts yields

$$\begin{aligned} D_{21}U(r, c) &= \int_{v(O)} dx D_{211}u(x, c) \int_{[v_0, x]} dy D_2H(y, r) \\ &= \int_{v(O)} dx D_{211}u(x, c) D_2\Phi(x, r) \end{aligned}$$

As  $D_2\Phi \geq 0$  on  $v(O) \times [0, 1]$ ,  $D_{211}u \geq 0$  (resp.  $\leq 0$ ) implies  $D_{21}U(r, c(r)) \geq 0$  (resp.  $\leq 0$ ), and consequently,  $Dc(r) \geq 0$  (resp.  $\leq 0$ ). ■

While the above argument requires  $u$  to be  $\mathcal{C}^2$ , we note that in the following special case this requirement can be relaxed somewhat.

**Remark 3.2** Suppose  $D_2U(r, c) = \int_{v(O)} dx h(x, r) f(x) - De(c)$  for functions  $f$  and  $e$  such that  $f$  is  $\mathcal{C}^2$  and  $e$  is twice differentiable, with  $D^2f \leq 0$  and  $D^2e > 0$ . In order to simplify notation, set  $\alpha(r) \equiv \int_{v(O)} dx h(x, r) f(x)$  and  $\beta(c) \equiv De(c)$ . It is straightforward to check that  $D_2U$  is differentiable. Moreover,  $\beta$  has a function inverse  $\beta^{-1}$  and elementary arguments show that  $\beta^{-1}$  is differentiable, with  $D\beta^{-1} = [D\beta]^{-1} = -[D_{22}U]^{-1}$ . Consequently, the first order condition  $D_2U(r, c(r)) = 0$  implies  $c(r) = \beta^{-1} \circ \alpha(r)$  and  $Dc(r) = D\beta^{-1}(\alpha(r)) D\alpha(r) = -D_{21}U(r, c(r)) [D_{22}U(r, c(r))]^{-1}$ , as in the proof of Theorem 3.1. Note that  $D_{22}U(r, c) = -D^2e(c) < 0$  and  $D_{21}U(r, c) = \int_{v(O)} dx D_2h(x, r) f(x)$ . Copying the arguments in the proof of Theorem 3.1,  $D_{21}U(r, c) = \int_{v(O)} dx D^2f(x) D_2\Phi(x, r) \leq 0$ . Thus,  $Dc(r) \leq 0$ . Thus, if  $D_2U$  is additively separable as in this case, then the formula for  $Dc(r)$  is valid even if  $e$  is twice differentiable without being  $\mathcal{C}^2$ .

For some  $u$ , the smoothness requirements can be substantially relaxed.

**Remark 3.3** In special cases, such as the application involving utility regulation, the problem allows approximation of a non-differentiable  $u$  by smooth functions. As we show below, in such cases, it may be possible to analyze the approximate problems and take limits to solve the original problem.

Each of the following applications provides an economic interpretation of the formalism used in Theorem 3.1. We avoid needless repetition of hypotheses by stating a *portmanteau convention*: In every application, each hypothesis underlying Theorem 3.1 will hold. A given hypothesis will be either an explicit consequence of the relevant economic model, or if the model is silent with regard to that hypothesis, then it is assumed to hold implicitly at the level of generality of Theorem 3.1. Assumptions in addition to those of Theorem 3.1 will be stated explicitly.

## Sealed bid auctions with valuation risk

Consider the problem of bidding in a sealed bid auction when the bidder's valuation is determined by a vector-valued random variable whose outcome is unknown at the time of bidding.<sup>6</sup> In the cases of first-price and second-price auctions, we show that greater riskiness of the value-determining vector of variables leads to a lower bid.

We study the bidding strategy of bidder  $n + 1$  (henceforth, “the buyer”) in an auction of a single indivisible prize with  $n + 1$  bidders. All bidders simultaneously submit sealed bids. If  $(y_1, \dots, y_n)$  is the profile of bids by the other bidders, then the buyer wins the prize by bidding  $b$  if and only if  $y^* \equiv \max\{y_1, \dots, y_n\} \leq b$ ; for simplicity, we break ties in favor of the buyer. If the buyer wins, then he pays  $b$  in a first-price auction and  $y^*$  in a second price auction; otherwise, the buyer pays nothing in both auctions.

A random outcome  $x \in O$  may be a vector of opinions of appraisers and future buyers regarding the value of the prize. If the prize is a mining or oil exploration concession, a plot of land, a patent, or a financial asset (e.g., stock of a firm being privatized or taken over),  $x$  may be a vector of future revenues or prices that affects the current valuation of the prize. If the future evolution of the relevant variable is modeled by a random process, then  $x$  would be interpreted as a sample path of the process.

Suppose  $v \in \mathcal{V}$  is the statistic that aggregates the random vector  $x$  into the external valuation of the prize  $v(x)$ . Let  $\mu(r) \in \Delta(O)$ ,  $r \in [0, 1]$ , be the buyer's belief about  $x$ . Let  $f : v(O) \rightarrow \mathfrak{R}_+$  determine the buyer's value of the prize as a function of the external valuation. Assume that  $f$  is  $\mathcal{C}^2$ , increasing and concave, with  $Df(v_1) = 0$ . Let  $e(c)$  be the minimum expected payment by the buyer for ensuring the winning probability  $c \in C = [0, 1]$ . Define  $u : v(O) \times C \rightarrow \mathfrak{R}$  by  $u(v, c) = f(v)c - e(c)$ , which is the buyer's expected surplus when he chooses the winning probability  $c$  and the external valuation is  $v$ . The buyer chooses  $c \in C$  to maximize  $\int_O \mu(r)(dx) u(v(x), c)$ . We use “winning probability  $c$ ” as shorthand for “probability of winning that is at least  $c$ ”; this is sensible because  $f(v) \geq 0$  implies that, given an expected payment  $e$ , the buyer's payoff is increasing in  $c$ . While  $f$  is given exogenously,  $e$  is derived from the data as follows.

Since we are studying the buyer's decision problem, we take as given the distribution function  $G$  of  $y^*$ . For simplicity, let  $\text{supp } G = [0, 1]$ . Given either auction form, if the buyer bids  $b$ , then he wins with probability  $G(b)$ .<sup>7</sup>

<sup>6</sup>This set-up is inspired by the general model of Milgrom and Weber [12].

<sup>7</sup>A Bayesian game model of the auction generates  $G$  as follows. Let  $Q \in \Delta(O)$  be the common prior belief of the bidders. Let  $\sigma_i : O \rightarrow T_i$  generate bidder  $i$ 's private signal and let  $s_i : T_i \rightarrow \mathfrak{R}_+$  be bidder  $i$ 's bidding strategy. So, the maximum bid by the other bidders is generated by the random variable  $x \mapsto \max\{s_i \circ \sigma_i(x) \mid i = 1, \dots, n\}$ .  $G$  is the distribution function of this random variable. Given this set-up,  $\mu(r)$  is to be interpreted as the buyer's posterior belief conditional on his signal  $\sigma_{n+1}(x)$ .

Define  $b : C \rightarrow \mathfrak{R}$  by

$$b(c) = \begin{cases} 0, & \text{if } c = 0 \\ \inf G^{-1}([c, 1]), & \text{if } c \in (0, 1] \end{cases}$$

Consider  $c \in (0, 1]$ .  $G^{-1}([c, 1])$  is bounded below by 0. As  $G$  is right-continuous,  $G^{-1}([c, 1])$  is closed in  $\mathfrak{R}$ . Therefore,  $b(c) \in G^{-1}([c, 1])$ , i.e.,  $G(b(c)) \geq c$ . If  $b < b(c)$ , then  $G(b) < c$ . Thus,  $b(c)$  is the lowest bid by the buyer that yields the winning probability  $c$ . It is easy to verify that  $b$  is nonnegative, increasing and bounded.

Consider  $\alpha \in \mathfrak{R}$ . If  $c \in b^{-1}((\alpha, \infty))$ , then  $c \in C$  and  $G(\alpha) < c$  by the definition of  $b$ . Conversely, suppose  $c \in C$  and  $G(\alpha) < c$ . As  $G(b(c)) \geq c$ , we have  $G(\alpha) < G(b(c))$ . As  $G$  is increasing,  $\alpha < b(c)$ . Thus,  $b^{-1}((\alpha, \infty)) = (G(\alpha), \infty) \cap C$ . As this holds for every  $\alpha \in \mathfrak{R}$ ,  $b$  is left-continuous.

It is evident from inspection that a jump in the graph of  $b$  (resp.  $G$ ) corresponds to a flat portion of the graph of  $G$  (resp.  $b$ ). More precisely,  $G$  and  $b$  are dual in the following sense:

**Lemma 3.4** *If  $G$  and  $b$  are as defined above, then  $b$  (resp.  $G$ ) is continuous if and only if  $G$  (resp.  $b$ ) is strictly increasing on  $[0, 1]$ .*

Define  $g : C \rightarrow \mathfrak{R}$  by

$$g(c) = \begin{cases} b(c)G(b(c)), & \text{for the first-price auction} \\ \int_{[0, b(c)]} y^* dG(y^*), & \text{for the second price auction} \end{cases}$$

In both cases,  $g(c)$  is the buyer's expected payment if he bids efficiently to ensure the winning probability  $c$ . Clearly,  $g$  is nonnegative and bounded. As  $g$  is increasing, it is Borel measurable. Therefore, it is integrable. However, in general,  $g$  is not continuous and not a convex function. Therefore, given  $g$ , an optimal winning probability will generally be generated as the expectation of a random choice of winning probabilities. This possibility is formalized as follows.

For this application, let  $\Delta(C)$  denote the set of *finitely* additive lotteries on  $(C, \mathcal{C})$ , where  $\mathcal{C}$  is the Borel  $\sigma$ -algebra on  $C$ .<sup>8</sup> Define  $L : \Delta(C) \rightarrow \mathfrak{R}$  by  $L(\lambda) = \int_C \lambda(dx) g(x)$ . If the buyer chooses his winning probability randomly using  $\lambda \in \Delta(C)$ , then his expected payment is  $L(\lambda)$ . Given  $c \in C$ ,  $\Delta(C, c) = \{\lambda \in \Delta(C) \mid \int_C \lambda(dx) x \geq c\}$  is the set of probability measures on  $C$  that yield an expected winning probability greater than  $c$ . For  $c \in C$ , let

$$e(c) = \inf\{L(\lambda) \mid \lambda \in \Delta(C, c)\}$$

As  $g$  is nonnegative,  $e(c) \geq 0$ . If  $c_1, c_2 \in [0, 1]$  and  $c_1 \geq c_2$ , then  $\Delta(C, c_1) \subset \Delta(C, c_2)$ ; consequently,  $e$  is increasing. In order to show that  $e(\cdot)$  is real-valued and convex, we use the following result.

<sup>8</sup>The motivation for this expansion is Lemma 3.5. The more commonly used weak\* topology would yield the same consequences with countably additive lotteries if  $g$  is continuous; unfortunately it is not generally so.



**Lemma 3.5** *If  $\Delta(C)$  is given the  $B(C, C)$  topology, then  $L$  is continuous,  $\Delta(C, c)$  is compact, and for every  $c \in C$ , there exists  $\lambda \in \Delta(C, c)$  such that  $e(c) = L(\lambda)$ .*

Thus,  $e(\cdot)$  defines a function  $e : C \rightarrow \mathfrak{R}$ . Consider  $c_1, c_2 \in C$ ,  $t \in [0, 1]$  and  $c = tc_1 + (1-t)c_2$ . By Lemma 3.5, there exists  $\lambda_1 \in \Delta(C, c_1)$  and  $\lambda_2 \in \Delta(C, c_2)$  such that  $e(c_1) = L(\lambda_1)$  and  $e(c_2) = L(\lambda_2)$ . Set  $\lambda = t\lambda_1 + (1-t)\lambda_2$ . Since  $\int_C \lambda(dx) x = t \int_C \lambda_1(dx) x + (1-t) \int_C \lambda_2(dx) x \geq tc_1 + (1-t)c_2 = c$ , we have  $\lambda \in \Delta(C, c)$ . It follows that  $e(c) \leq L(\lambda) = tL(\lambda_1) + (1-t)L(\lambda_2) = te(c_1) + (1-t)e(c_2)$ . Thus,  $e$  is a convex function. This implies that  $e$  is continuous, indeed locally Lipschitz, on  $(0, 1)$ . Moreover, the Busemann-Feller-Alexandrov theorem implies that  $e$  is twice differentiable Lebesgue almost everywhere.

We now apply Theorem 3.1 to the buyer's problem. In addition to the *portmanteau* convention, suppose  $f$  is  $\mathcal{C}^2$  and  $e$  is twice differentiable, with  $D^2f \leq 0$  and  $D^2e > 0$ ; clearly, the regularity requirements on  $e$  are somewhat more than what can be deduced from the economic model itself. Then,  $D_2U(r, c) = \int_{v(O)} dx h(x, r) f(x) - De(c)$  and the requirements of Remark 3.2 are satisfied. It follows that  $Dc(r) \leq 0$ , i.e., greater riskiness of his valuation will induce the buyer to bid more conservatively.

### Size of regulated utility

Consider a firm that supplies electricity to a grid and is required by the regulator to match the random demand flow over the period  $T = [0, 1]$ . Failure to do so attracts a penalty. While the firm bears the fixed cost (e.g., capacity outlays), the firm's revenue from the sale of electricity and its variable cost (e.g., fuel) are monitored and fully internalized by the regulator *via* reimbursement of the net amount. In addition, the firm receives a lump-sum return for its participation in this arrangement. The firm chooses capacity to minimize the sum of the fixed cost and the expected penalty. We show that the chosen capacity is an increasing function of the riskiness of the demand process.

We formalize the firm's problem as follows. Let  $X = \mathcal{C}(T, \mathfrak{R})$ , which is given the compact-open topology. As  $T$  is compact, this topology coincides with the topology of uniform convergence generated by the supremum norm (Dugundji [6], Theorem XII.8.2). Let  $O$  be a convex subset of  $\mathcal{C}(T, [0, \gamma]) \subset X$  for some  $\gamma > 0$ , such that  $x \equiv 0 \in O$ . The firm's demand flow is generated by the random process  $\delta : O \times T \rightarrow \mathfrak{R}$  where  $\delta(x, t) = x(t)$  is the demand at time  $t$  generated by the sample path  $x$ . Applying Theorem XII.2.4 in Dugundji [6],  $\delta$  is continuous. Let  $\mu(r) \in \Delta(O)$  be the firm's belief about the demand trajectory, with the interpretation that the perceived riskiness of the demand path increases with parameter  $r \in [0, 1]$ .

Given a demand trajectory  $x \in O$ , let  $w(x) = \sup x(T)$  be the largest instantaneous demand over  $T$ . Clearly,  $w : O \rightarrow [0, \gamma]$  is convex and bounded. For every  $\alpha \in \mathfrak{R}$ ,  $w^{-1}((\alpha, \infty)) = \cup_{t \in T} \{x \in O \mid x(t) > \alpha\} = \cup_{t \in T} \{x \in O \mid \delta(x, t) > \alpha\} = \cup_{t \in T} \delta(\cdot, t)^{-1}((\alpha, \infty))$  is open in  $O$  as  $\delta$  is continuous. Thus,  $w$  is lower semicontinuous.<sup>9</sup> Therefore,  $w$  is Borel measurable.

Suppose the firm has to pay a penalty based on the largest deficit over  $T$ . Let  $C = [a, b]$  be the set of capacity choices, with  $0 < a < b$  and  $a < \gamma$ . Given a demand trajectory  $x \in O$  and capacity  $c \in C$ , the largest deficit over  $T$  is  $\sup_{t \in T} \max\{x(t) - c, 0\} = \max\{\sup x(T) - c, 0\} = \max\{w(x) - c, 0\} \in [0, \gamma]$  and the resulting penalty is  $P(\max\{w(x) - c, 0\})$ . Let  $S(c)$  be the cost of setting-up capacity  $c \in C$ . We assume that  $S$  is  $\mathcal{C}^3$ , with  $D^2S > 0$ , and the penalty function  $P : [-b, \gamma] \rightarrow \mathfrak{R}$  is  $\mathcal{C}^3$ , with  $P(x) = 0$  for  $x \leq 0$  and  $DP, D^2P, D^3P \geq 0$ . While excess supply is not rewarded, penalties increase in severity with the maximum supply shortfall.

Suppose, given  $r \in [0, 1]$ , capacity  $c = c(r) \in (a, b)$  minimizes over  $C$  the firm's decision criterion, namely, the expected penalty plus the capacity building cost:  $\int_O \mu(r)(dx) P(\max\{w(x) - c, 0\}) + S(c)$ . We show that, if  $r_1, r_2 \in [0, 1]$  and  $r_1 > r_2$ , then  $c(r_1) \geq c(r_2)$ , i.e., greater riskiness of the demand trajectory induces the firm to choose a larger capacity.

We begin by transforming the firm's problem to the format of the problem in Theorem 3.1. It follows from the properties of  $w$  that  $v = -w \in \mathcal{V}$  and  $v(O) \subset [-\gamma, 0]$ . Define  $g : [-b, \gamma] \rightarrow [-b, 0]$  by  $g(x) = \min\{x, 0\}$ . Then, for  $r \in [0, 1]$ ,  $c = c(r) \in (a, b)$  maximizes

$$U(r, c) = \int_O \mu(r)(dx) u(v(x), c)$$

over  $C$ , where  $u : [-\gamma, 0] \times C \rightarrow \mathfrak{R}$  is given by  $u(v, c) = -S(c) - P \circ (I - g)(-v - c)$  and  $I$  is the identity mapping. Although this problem formally resembles the problem of Theorem 3.1, we cannot use this theorem to analyze  $c(\cdot)$  directly as  $u$  is not sufficiently regular. Before plunging into the details of our indirect method of analyzing  $c(\cdot)$ , we outline the broad steps.

First, we uniformly approximate  $g$  by a sequence  $(g_n)$  of smooth functions. Each  $g_n$  generates a uniform approximation  $u_n$  of  $u$ , which generates an approximate version of the firm's problem and a solution  $c(\cdot; n)$  of this approximate problem.

Second, applying Theorem 3.1, we find that  $Dc(r; n) \geq 0$  for all  $r \in [0, 1]$  and  $n \in \mathcal{N}$ . So, for  $r_1 > r_2$ , we have  $c(r_1; n) \geq c(r_2; n)$  for every  $n \in \mathcal{N}$ .

Finally, we generate cluster points  $c^*(r_1)$  and  $c^*(r_2)$  of the sequences  $(c(r_1; n))$  and  $(c(r_2; n))$  respectively, and show that  $c(r_1) = c^*(r_1) \geq c^*(r_2) = c(r_2)$ , thus answering our basic question.

We now implement this plan.

**Step 1.** Fix  $N \in \mathcal{N}$  such that  $2/N < a$  and consider  $n \geq N$ . Lemma 2.5 implies that there exists a concave, increasing  $\mathcal{C}^\infty$  function  $g_n : [-b, \gamma] \rightarrow$

<sup>9</sup>In this non-Euclidean setting, convexity of  $w$  is not a sufficient condition for continuity.

$[-b, 0]$  such that  $\sup\{|g_n(x) - g(x)| \mid x \in [-b, \gamma]\} \leq 1/n$ ,  $g_n \geq g$  on  $[-b, \gamma]$ ,  $Dg_n = 1$  on  $[-b, -2/n]$ ,  $Dg_n \in (0, 1)$  on  $(-2/n, 0)$  and  $g_n = Dg_n = 0$  on  $[0, \gamma]$ . Set  $g_1 = \dots = g_{N-1} = g_N$ . Given  $n \in \mathcal{N}$ , define  $u_n : [-\gamma, 0] \times C \rightarrow \mathfrak{R}$  by  $u_n(v, c) = -S(c) - P \circ (I - g_n)(-v - c)$ . Clearly,  $u_n$  is  $\mathcal{C}^3$ .

**Lemma 3.6** *The sequence  $(u_n)$  converges uniformly to  $u$ .*

Let  $c = c(r; n) \in (a, b)$  solve the approximate problem of maximizing

$$U(r, c; n) = \int_O \mu(r)(dx) u_n(v(x), c)$$

**Step 2.** We now analyze  $c(\cdot; n)$ . The *portmanteau* convention and the above structural model ensure that the conditions of Theorem 3.1 are met by the approximate version of the firm's problem.

Consider  $n \geq N$ . It is easily confirmed that  $D_{22}u_n < 0$ . As  $2/n < a$  and  $Dg_n(x) = 1$  for every  $x \in [-b, -2/n]$ , we have  $Dg_n(x) = 1$  for every  $x \in [-b, -a]$ . Since  $0 \in O$ ,  $v_1 = \sup v(O) = 0$ . Consequently, for every  $c \in C$ ,  $D_1u_n(v_1, c) = D_1u_n(0, c) = DP((I - g_n)(-c))[1 - Dg_n(-c)] = 0$ . It is straightforward to calculate that  $D_{211}u_n = -DPD^3g_n - 3D^2P(1 - Dg_n)D^2g_n + (1 - Dg_n)^3D^3P$ , where the derivatives of  $P$  are evaluated at  $(I - g_n)(w - c)$  and the derivatives of  $g_n$  are evaluated at  $w - c$ . If  $w - c < 0$ , then  $(I - g_n)(w - c) = w - c - g_n(w - c) < w - c - \min\{w - c, 0\} = 0$ . Consequently,  $DP \circ (I - g_n)(w - c) = 0$  as  $P(x) = 0$  for  $x \leq 0$ . On the other hand, if  $w - c \geq 0$ , then  $D^3g_n(w - c) = 0$  as  $g_n(x) = 0$  for  $x \geq 0$ . Given our hypotheses regarding  $P$  and the facts that  $g_n$  is concave and  $Dg_n \leq 1$ , we have  $D_{211}u_n = -3D^2P(1 - Dg_n)D^2g_n + (1 - Dg_n)^3D^3P \geq 0$ . By Theorem 3.1,  $Dc(r; n) \geq 0$  for  $n \geq N$ .

So, for  $n \geq N$  and  $r_1, r_2 \in \mathfrak{R}$  with  $r_1 > r_2$ , we have  $c(r_1; n) \geq c(r_2; n)$ .

**Step 3.** As  $C$  is compact, the Bolzano-Weierstrass property implies the existence of a subsequence  $(c(r_1, k))_{k \in \mathcal{N}}$  of  $(c(r_1, n))_{n \in \mathcal{N}}$ , converging to some  $c^*(r_1) \in C$ . For the same reason, there exists a subsequence  $(c(r_2, j))_{j \in \mathcal{N}}$  of  $(c(r_2, k))_{k \in \mathcal{N}}$ , converging to some  $c^*(r_2) \in C$ . Since every subsequence of a convergent sequence converges to the limit point of the sequence, the subsequence  $(c(r_1, j))_{j \in \mathcal{N}}$  of  $(c(r_1, k))_{k \in \mathcal{N}}$  also converges to  $c^*(r_1)$ .

By Step 2,  $c(r_1; j) \geq c(r_2; j)$  for all but a finite set of  $j \in \mathcal{N}$ . Taking limits as  $j \uparrow \infty$ , we have  $c^*(r_1) \geq c^*(r_2)$ .

**Lemma 3.7** *For  $r \in \{r_1, r_2\}$ ,  $c^*(r)$  maximizes  $U(r, \cdot)$  over  $C$ .*

As  $U(r, \cdot)$  is strictly concave, its maximum over  $C$  is unique. Therefore, Lemma 3.7 implies  $c(r_1) = c^*(r_1) \geq c^*(r_2) = c(r_2)$ , as claimed.

## Allocation of a public good

In this application we study the planner's problem of using a given resource to provide public goods for two periods, say 1 and 2. These goods are assumed to be excludable across periods but non-excludable and non-rival within a period. Alternatively, jurisdictions may be defined by a characteristic other than time, say by location or some socio-economic criterion.

Let  $I$  be a finite set of agents. An agent's allocation in a period consists of the public good provided by the planner in that period and a vector of other goods, which may be private or public. The planner's resource constraint is modeled by interpreting  $\phi(c)$  as the maximum amount of public good that the planner can provide in period 1 if  $c$  is to be supplied in period 2. While the allocation of other goods in period 1 is known to the planner, the allocation of other goods in period 2 depends on a vector-valued random variable. We consider various specifications of this scenario and analyze how the provision of the planner's public good varies with the riskiness of the random outcome in period 2.

**Example 3.8** *Let agent  $i$ 's period 2 allocation  $(x_i, c) \in \mathbb{R}_+^n \times \mathbb{R}_+$  consist of the planner's public good  $c$  and a vector  $x_i$  of private goods. Agent  $i$ 's utility from this allocation is  $g_i(f_i(x_i), c)$ , where  $f_i : \mathbb{R}_+^n \rightarrow \mathbb{R}$  is continuous and concave. Suppose the allocation of private goods in period 2 is random. This is formalized by setting  $X = (\mathbb{R}_+^n)^{|I|}$  and  $O \subset (\mathbb{R}_+^n)^{|I|}$ . Given  $x = (x_i)_{i \in I} \in O$ , let  $v(x) = \min\{f_i(x_i) \mid i \in I\}$ . The public good  $c$  in period 2 is chosen by the planner from  $C \subset \mathbb{R}_+$ . Given an allocation  $(z, c) \in O \times C$  in some period, let the social welfare in that period be  $w(v(z), c)$ . If  $y \in O$  is the known allocation of private goods in period 1, then the aggregate two-period social welfare is  $w(v(y), \phi(c)) + w(v(x), c)$ . Since  $y$  is known and plays no role in the analysis, we suppress it and interpret  $\psi(c) = -w(v(y), \phi(c))$  as the social cost of supplying  $c$  in period 2. So, the planner's objective function is  $u(v(x), c) = w(v(x), c) - \psi(c)$ , given  $c$  and the random allocation  $x$ .*

In this example,  $v$  generates the value of a private goods allocation using the Rawlsian criterion. Given the product topology on  $(\mathbb{R}_+^n)^{|I|}$ , the  $i$ -th projection mapping  $\pi_i : (\mathbb{R}_+^n)^{|I|} \rightarrow \mathbb{R}_+^n$  is continuous. Consequently,  $f_i \circ \pi_i$  is continuous. Thus,  $v$  is continuous (Berge [3], Theorems IV.8.3 and IV.8.4).

**Example 3.9** *Suppose the aggregate endowment of private goods in period 2 is random. Once the random endowment is realized, the planner chooses its allocation. Accordingly, let  $X = \mathbb{R}_+^n$  and  $O \subset \mathbb{R}_+^n$ . Given an endowment of private goods  $x \in O$ , the set of possible allocations is  $S(x) = \{(x_i)_{i \in I} \in (\mathbb{R}_+^n)^{|I|} \mid \sum_{i \in I} x_i = x\}$ . Set  $v(x) = \sup\{\min\{f_i(x_i) \mid i \in I\} \mid (x_i)_{i \in I} \in S(x)\}$  and define  $w$ ,  $\phi$ ,  $\psi$  and  $u$  as in Example 3.8.*

In this example,  $v$  generates the value of a social endowment of private goods from the optimal allocation of that endowment using the Rawlsian

criterion. We have shown above that the mapping  $(x_i)_{i \in I} \mapsto \min\{f_i(x_i) \mid i \in I\}$  is continuous. It is easily shown that the set-valued mapping  $S$  is continuous. Therefore,  $v$  is continuous (Berge [3], Theorem VI.3).

In the next example, private goods are excluded for the sake of simplicity as we consider various situations in which agent  $i$ 's allocation  $(x, c) \in \mathfrak{R}_+^n \times \mathfrak{R}_+$  consists only of the planner's public good  $c$  and a vector  $x$  of other public goods. Unlike in Examples 3.8 and 3.9,  $f_i$  is now interpreted as agent  $i$ 's utility over the other public goods.

**Example 3.10** *Suppose the allocation of other public goods is random. Accordingly, let  $X = \mathfrak{R}^n$ ,  $O \subset \mathfrak{R}_+^n$  and  $\Delta = \{\lambda \in \mathfrak{R}_+^{|I|} \mid \sum_{i \in I} \lambda_i = 1\}$ . Given an allocation of other public goods  $x \in O$ , define  $v : O \rightarrow \mathfrak{R}$  by any of the following formulae: (a)  $v(x) = \sum_{i \in I} \lambda_i f_i(x)$  for some  $\lambda \in \Delta$ , (b)  $v(x) = \sup\{\sum_{i \in I} \lambda_i f_i(x) \mid \lambda \in \Delta\}$ , (c)  $v(x) = \min\{\lambda_i f_i(x) \mid i \in I\}$  for some  $\lambda \in \Delta$ , or (d)  $v(x) = \sup\{\min\{\lambda_i f_i(x) \mid i \in I\} \mid \lambda \in \Delta\}$ . Define  $w$ ,  $\phi$ ,  $\psi$  and  $u$  as in Example 3.8.*

In case (a) (resp. (c)),  $v$  generates the value of an allocation of public goods using the weighted Benthamite (resp. Rawlsian) criterion. In case (b) (resp. (d)),  $v$  generates the value of an allocation of public goods using the valuation of agents who value it the most (resp. least). Continuity of  $v$  in case (a) follows from Theorem IV.8.5 in Berge [3]. Continuity of  $v$  in case (c) follows from Theorems IV.8.3 and IV.8.4 in Berge [3]. Continuity of  $v$  in cases (b) and (d) follows from Theorem VI.3 in Berge [3].

In all the above examples, the concavity of  $v$  follows from Theorem 5.8 in Rockafellar [17]. In every case, the continuity of  $v$  implies its Borel measurability. Assume that  $O$  is compact and convex. Then,  $v$  is bounded in every case. Consequently,  $v \in \mathcal{V}$  in each case. In addition to the *portmanteau* convention, suppose  $w$  and  $\phi$  are  $\mathcal{C}^2$ ,  $D_2 w > 0$ ,  $D^2 \phi < 0$ ,  $D_{22} w < 0$ ,  $D_1 w(v_1, \cdot) = 0$  and  $D_{21} w(\cdot, c)$  is  $\mathcal{C}^1$  for every  $c \in C$ . Then,  $u$  satisfies the requirements of Theorem 3.1.

If  $D_{211} w \geq 0$ , then Theorem 3.1 implies that, in all the above examples, the planner's optimal response to an increase in riskiness of the random outcome  $x \in O$  is to increase the second period public good allocation  $c$ .

## Portfolio choice

This application extends the analysis of the static portfolio problem from the scalar risk (e.g., in Rothschild and Stiglitz [19]) to the vector risk case.

Consider an investor who has to choose a portfolio of two financial assets, a stock and a bond. The returns *per share* and *per bond* are defined over the time horizon  $T \subset \mathfrak{R}_+$ ; e.g.,  $T = \mathcal{N}$  or  $T = \mathfrak{R}_+$ . The dividend *per share* fluctuates randomly over  $T$ . Suppose the random process generating the stock dividends has sample paths that belong to  $O \subset X$ , where  $X$  is a space

of functions defined on  $T$ . Suppose  $O$  and  $X$  are as in Section 2; see Dunford and Schwartz [7] for numerous examples of appropriate function spaces  $X$ . Let  $d : O \times T \rightarrow \mathfrak{R}$  generate the dividends *per* share, with  $d(x, t) \equiv x(t)$  as the dividend at time  $t \in T$ , contingent on sample path  $x \in O$ . Consider a family of stock dividend processes represented by an indexed family of probability measures  $\{\mu(r) \in \Delta(O) \mid r \in \mathfrak{R}\}$ , such that the riskiness of these measures increases with  $r$ . Unlike the stock, the stream of returns *per* bond, say  $y \in O$ , is known when the portfolio is chosen.

Consider  $v \in \mathcal{V}$  and a function  $f : v(O) \rightarrow \mathfrak{R}$ . If the stream of returns from a unit of an asset is  $x \in O$ , then the investor's valuation of that unit is  $v(x)$  and the resulting utility is  $f \circ v(x)$ . We assume that  $f$  is  $\mathcal{C}^2$ , with  $Df > 0$ ,  $D^2f < 0$  and  $Df(v_1) = 0$ .

Provided  $T$  and  $O$  are appropriately specified, linear examples of  $v$  include the projection mapping  $v(x) = x(\tau)$  for  $\tau \in T$  and the discounted sum-of-returns functionals  $v(x) = \int_T dt e^{-\delta t} x(t)$  and  $v(x) = \sum_{t \in T} \delta^t x(t)$ . Non-linear examples include the supremum and infimum mappings  $v(x) = -\sup x(T)$  and  $v(x) = \inf x(T)$ , the norm mapping  $v(x) = -\|x\|$  where  $\|\cdot\|$  is a norm on  $X$  and discounted non-linear functionals  $v(x) = \int_T dt e^{-\delta t} \psi \circ x(t)$  and  $v(x) = \sum_{t \in T} \delta^t \psi \circ x(t)$  where  $\psi$  is concave and continuous.

Normalize the share price to unity and suppose the investor has no initial wealth. Consequently, the investor's portfolio  $(c, Q(c))$  satisfies the identity  $c + Q(c)P(c) = 0$ , where  $c \in C = [0, 1]$  is the share holding,  $Q(c)$  is the bond holding and  $P(c)$  is the bond price. As the available stock is normalized to 1,  $c$  may be interpreted as a proportion of the available stock. In order to study the investor's choice of  $c$ , we need to know the behavior of  $Q : C \rightarrow \mathfrak{R}$  and  $P : C \rightarrow \mathfrak{R}$  separately because bond returns are specified *per* bond in  $Q(c)$  rather than *per* dollar in  $Q(c)P(c)$ . We do this as follows.

For some  $P_0, P_1 > 0$  with  $P_0 < P_1$ , let  $G : (P_0, P_1) \rightarrow \mathfrak{R}$  be a  $\mathcal{C}^2$  function such that  $G > 0$ ,  $DG < 0$ ,  $D^2G < 0$ ,  $G(P_0) + P_0 DG(P_0) < 0$ , and for every  $c \in C$ , there exists  $p \in (P_0, P_1)$  such that  $c = G(p)p$ . We interpret  $G(p)$  as the demand for bonds when the investor announces a bond price  $p$ . Then, there exists a unique  $\mathcal{C}^2$  function  $P : C \rightarrow (P_0, P_1)$  such that  $c = G(P(c))P(c)$  for every  $c \in C$ . It is easily verified that  $P > 0$ ,  $DP < 0$  and  $D^2P < 0$ . Moreover,  $DP(c) < 0 < P(c)/c$  for every  $c \in C$ . Since  $Q(c) = -c/P(c)$ , it is easily verified that  $Q$  is  $\mathcal{C}^2$ ,  $Q < 0$ ,  $DQ < 0$  and  $D^2Q < 0$ .  $Q < 0$  means that the investor sells bonds in order to finance his stock holding.

Define  $u : v(O) \times C \rightarrow \mathfrak{R}$  by  $u(v, c) = cf(v) + Q(c)\alpha$ , where  $\alpha > 0$ . Suppose  $x \in O$  is the stream of dividends *per* share,  $y \in O$  is the stream of returns *per* bond and  $\alpha \equiv f \circ v(y) > 0$ . Then, the investor's utility from portfolio  $(c, Q(c))$  is  $cf \circ v(x) + Q(c)f \circ v(y) = cf \circ v(x) + Q(c)\alpha = u(v(x), c)$ . We have already assumed that  $X, O, v$  and  $C$  are as *per* the requirements of Theorem 3.1. Suppose  $\{\mu(r) \in \Delta(O) \mid r \in \mathfrak{R}\}$ ,  $H, h$  and  $c(\cdot)$  also satisfy the requirements of this theorem. Since  $D_{211}u = D^2f < 0$ , Theorem 3.1 implies

that  $Dc(r) \leq 0$ , i.e., an increase in the riskiness of the stock dividend process implies a smaller stock holding in the portfolio.

### Multi-party moral hazard

Consider a decision-maker who might suffer a loss  $a > 0$  on account of a transmittable disease. The probability of loss is  $Q(v(x), c)$ , where  $c \in C = [a, b]$  is the care taken by the decision-maker,  $x = (x_1, \dots, x_n)$  is the profile of care taken by other persons and  $v(x) = \min\{x_1, \dots, x_n\}$  is the effective care exercised by them. Suppose care  $c$  implies cost  $c$  to the decision-maker. Thus, the decision-maker selects  $c \in C$  to minimize  $c + a \int_O \mu(r)(dx) Q(v(x), c)$ , where  $\mu(r)$  is the decision-maker's belief about the care levels chosen by other persons. Equivalently,  $c$  is chosen to maximize  $\int_O \mu(r)(dx) u(v(x), c)$ , where  $u(v, c) = -c - aQ(v, c)$ .

Suppose  $Q : v(O) \times C \rightarrow [0, 1]$  is  $\mathcal{C}^2$  and decreasing,  $D_{22}Q > 0$  on  $v(O) \times (a, b)$ ,  $D_{21}Q(., c)$  is  $\mathcal{C}^1$  for every  $c \in C$  and  $D_1Q(v_1, .) = 0$ . Set  $X = \mathfrak{R}^n$  and let  $O \subset X$  be convex and compact. If  $c : [0, 1] \rightarrow (a, b)$  is as in Theorem 3.1, then the sign of  $Dc(r)$  is the opposite of the sign of  $D_{211}Q$ .

### Team theory

A slight variation on the last application is as follows. Consider a team consisting of a principal and  $n$  agents. Let  $O \subset \mathfrak{R}_+^n$  be the set of possible effort levels of the agents. Define  $v \in \mathcal{V}$  by  $v(x) = \min\{x_1, \dots, x_n\}$ . Suppose the team's output is generated by the function  $u : v(O) \times C \rightarrow \mathfrak{R}$ . If  $x \in O$  is the profile of agents' efforts and the principal's management contribution is  $c \in C$ , then the effective effort input by the agents is  $v(x)$  and the resulting team output is  $u(v(x), c)$ . Suppose the principal does not know  $x$  when choosing  $c$ . Let  $\mu(r) \in \Delta(O)$  be the principal's belief about  $x$ . Suppose  $O$ ,  $C$  and  $u$  satisfy the assumptions of Theorem 3.1 and  $c = c(r)$  maximizes expected team output  $\int_O \mu(r)(dx) u(v(x), c)$ . Using Theorem 3.1, if an increase in the principal's contribution  $c$  reduces the concavity of  $u(., c)$ , then greater uncertainty regarding the agents' effort levels will induce the principal to increase his own contribution.

## 4 Conclusions

Given an outcome set  $O$  that is a nonempty, convex and measurable subset of a vector space  $X$ , we have defined relations  $\succeq_1$ ,  $\succeq_2$  and  $\succeq_3$  over the set of probability measures  $\Delta(O)$ .  $\succeq_1$  is defined in terms of concave-dominance over  $O$ ,  $\succeq_2$  is defined in terms of a family of concave-dominance relations over various subsets of  $\mathfrak{R}$  and  $\succeq_3$  is defined in terms of a generalized form of second-order stochastic dominance. Our main results, Theorems 2.6 and 2.8, show that these relations are identical.

Our result showing the identity of  $\succeq_1$  and  $\succeq_3$  significantly generalizes the corresponding Russell-Seo result, which was established in the Euclidean setting. In the special case when  $X = \mathfrak{R}$ , these relations coincide with the Rothschild-Stiglitz relation based on mean-preserving second order stochastic dominance. In Appendix B we recall, for the sake of completeness, two classical equivalence results that characterize comparative riskiness in terms of dilatations and martingale dominance.

In addition, we show two other results. First, Theorem 2.13 shows that if two measures are ordered by any of the above-mentioned relations, then their weak means must coincide. Second, we show *via* Theorem 2.10 and Corollary 2.11 that the comparison of riskiness of vector-valued random variables cannot be reduced to the comparison of riskiness of real-valued random variables without loss of generality.

Our main results are used to study six economic applications. We identify conditions under which

(1) a bidder in a sealed-bid (first-price or second-price) auction will bid more conservatively with increased riskiness of the vector-valued random variable determining his valuation of the prize;

(2) a regulated firm (e.g., an electricity generator) that is required to meet a random demand over a period will choose a higher capacity in response to greater uncertainty regarding the random demand process;

(3) a social planner would increase the supply of a public good in response to greater uncertainty regarding the consumption allocations of the citizens;

(4) a risk averse investor will increase the riskless component of her portfolio in response to increased riskiness of the stream of returns generated by the risky asset;

(5) the effort level chosen by a decision-maker in a multi-party moral hazard situation will increase or decrease as uncertainty regarding the effort levels chosen by the other decision-makers increases; and

(6) the principal in a team would increase her contribution of managerial input in response to increased uncertainty regarding the efforts contributed by the agents in the team.

These applications amply demonstrate the usefulness of the comparative vector riskiness theory. More specifically, our generalizations allow the tractable modeling of comparative riskiness even when the risk is embodied in an infinity of random variables, e.g., in the evolution of a continuous time random process.

## Appendix A: Proofs

**Proof of Lemma 2.4** As  $u \in \overline{\mathcal{U}(v)}$ , there is a sequence  $(u_n) \subset \mathcal{U}(v)$  converging uniformly to  $u$ . Since every  $u_n$  is concave and increasing, so is  $u$ . As each  $u_n$  is continuous, so is  $u$  (Bartle [2], Theorem 24.1). Therefore,  $u$  is



measurable. As  $v$  is bounded and  $u$  is increasing and continuous,  $u(v(O)) \subset [u(v_0), u(v_1)]$ . As  $u$  is increasing,  $u(v_1) \leq u(x)$  for every  $x \in (v_1, \infty)$ . We show that  $u(v_1) \geq u(x)$  for every  $x \in (v_1, \infty)$ . So,  $u([v_0, \infty)) \subset [u(v_0), u(v_1)]$ .

Suppose  $u(x) - u(v_1) = \epsilon > 0$  for some  $x \in (v_1, \infty)$ . As each  $u_n$  is concave, it attains a global maximum at  $v_1$ . Since  $u_n$  is increasing,  $u_n(x) = u_n(v_1)$  for every  $n \in \mathcal{N}$ . As  $(u_n)$  converges uniformly to  $u$ , there exists  $N \in \mathcal{N}$  such that  $n \geq N$  implies  $|u(x) - u_n(x)| < \epsilon/2$  and  $|u(v_1) - u_n(v_1)| < \epsilon/2$ . Thus, for  $n \geq N$ ,  $|u(v_1) - u(x)| \leq |u(v_1) - u_n(x)| + |u(x) - u_n(x)| = |u(v_1) - u_n(v_1)| + |u(x) - u_n(x)| < \epsilon$ , a contradiction.  $\blacksquare$

**Proof of Lemma 2.5** The proof is by construction. Let  $a = b - 2\epsilon$ .

Define  $\xi : \mathfrak{R} \rightarrow \mathfrak{R}$  by  $\xi(x) = (2x - b - a)/(b - a)$ . Then,  $\xi(a) = -1$  and  $\xi(b) = 1$ . Define  $\psi : \mathfrak{R} \rightarrow \mathfrak{R}$  by

$$\psi(x) = \begin{cases} \exp\{-(1 - x^2)^{-1}\}, & \text{if } |x| < 1 \\ 0, & \text{if } |x| \geq 1 \end{cases}$$

Define  $f : [v_0, \infty) \rightarrow \mathfrak{R}$  by  $f(x) = -\psi \circ \xi(x) / \int_{\mathfrak{R}} dx \psi \circ \xi(x)$ . Clearly,  $f \leq 0$ ,  $f = 0$  on  $\mathfrak{R} - (a, b)$ ,  $f$  is symmetric around  $(a + b)/2$  and  $\int_{\mathfrak{R}} dx f(x) = \int_{[a,b]} dx f(x) = -1$ . As  $-f$  is a probability density function on  $\mathfrak{R}$  that is symmetric around  $(a + b)/2$ , it follows that  $\int_{\mathfrak{R}} dx x f(x) = \int_{[a,b]} dx x f(x) = -(a + b)/2$ . As we shall verify that  $\psi$  is  $\mathcal{C}^\infty$ , so is  $f$ .

Define  $g : [v_0, \infty) \rightarrow \mathfrak{R}$  by  $g(x) = 1 + \int_{[v_0,x]} dy f(y)$ . It follows that  $g \geq 0$ ,  $g$  is  $\mathcal{C}^\infty$  and  $Dg = f$ .

Define  $h : [v_0, \infty) \rightarrow \mathfrak{R}$  by  $h(x) = v_0 - b + \epsilon + \int_{[v_0,x]} dy g(y)$ . Then,  $h$  is  $\mathcal{C}^\infty$ ,  $Dh = g \geq 0$  and  $D^2h = f \leq 0$ . Thus,  $h$  is increasing and concave. The claims regarding the values of  $Dh$  are easily verified using the properties of  $g$  and  $f$ .

We now compare  $h$  and  $u$ .

Consider  $x \in [v_0, a)$ . Clearly,  $f = 0$  and  $g = 1$  on  $[v_0, x]$ . Therefore,  $h(x) = v_0 - b + \epsilon + \int_{[v_0,x]} dy = x - b + \epsilon = u(x) + \epsilon$ .

Consider  $x \in (b, \infty)$ . Since  $g = 0$  on  $(b, \infty)$ , we have  $h(x) = v_0 - b + \epsilon + \int_{[v_0,b]} dy g(y)$ . Now,  $\int_{[v_0,b]} dy g(y) = \int_{[v_0,a]} dy g(y) + \int_{[a,b]} dy g(y) = a - v_0 + \int_{[a,b]} dy [1 + \int_{[a,y]} dz f(z)] = b - v_0 + \int_{[a,b]} dy \int_{[a,y]} dz f(z)$ . Integrating by parts,  $\int_{[a,b]} dy \int_{[a,y]} dz f(z) = b \int_{[a,b]} dz f(z) - \int_{[a,b]} dy y f(y) = -b + (a + b)/2$ . Combining these calculations, we have  $h(x) = v_0 - b + \epsilon + b - v_0 + (a - b)/2 = 0 = u(x)$  for  $x \in (b, \infty)$ .

Consider  $x \in [a, b]$ . Observe that  $\int_{[v_0,x]} dy g(y) = a - v_0 + \int_{[a,x]} dy [1 + \int_{[a,y]} dz f(z)] = x - v_0 + \int_{[a,x]} dy \int_{[a,y]} dz f(z)$ . Therefore,  $h(x) = v_0 - b + \epsilon + x - v_0 + \int_{[a,x]} dy \int_{[a,y]} dz f(z) = x - b + \epsilon + \int_{[a,x]} dy \int_{[a,y]} dz f(z) = u(x) + \epsilon + \int_{[a,x]} dy \int_{[a,y]} dz f(z)$ . Integrating by parts,  $0 \geq \int_{[a,x]} dy \int_{[a,y]} dz f(z) = \int_{[a,x]} dz (x - z) f(z) \geq \int_{[a,b]} dz (b - z) f(z) = (a + b)/2 - b = (a - b)/2 = -\epsilon$ . Since  $\int_{[a,x]} dy \int_{[a,y]} dz f(z) \in [-\epsilon, 0]$ , we have  $h(x) - u(x) \in [0, \epsilon]$ .

Combining the above arguments, it follows that  $h \geq u$  on  $[v_0, \infty)$  and  $\sup\{|h(x) - u(x)| \mid x \in [v_0, \infty)\} \leq \epsilon$ .

It only remains to confirm that  $\psi$  is  $\mathcal{C}^\infty$ . Observe that  $\psi = \phi \circ \eta$ , where  $\phi : \mathfrak{R} \rightarrow \mathfrak{R}$  is given by

$$\phi(y) = \begin{cases} e^{-1/y}, & \text{if } y > 0 \\ 0, & \text{if } y \leq 0 \end{cases} \quad (5)$$

and  $\eta$  is given by  $\eta(x) = 1 - x^2$ . Using the chain-rule, as  $\eta$  is  $\mathcal{C}^\infty$ , it suffices to show that  $\phi$  is  $\mathcal{C}^\infty$ . As preparation, we note two facts.

First, for  $n \in \mathcal{N}$ , it is easily verified that

$$D^n \phi(y) = \begin{cases} p_n(y) \phi(y) y^{-2n}, & \text{if } y > 0 \\ 0, & \text{if } y < 0 \end{cases} \quad (6)$$

where  $p_n(y)$  is the polynomial of degree  $n - 1$  generated by the recursive rule:  $p_1(y) = 1$  and  $p_{k+1}(y) = y^2 D p_k(y) - (2ky - 1)p_k(y)$  for  $k \in \mathcal{N}$ .

Second, consider  $y > 0$  and  $n \in \mathcal{N} \cup \{0\}$ . Using the power series representation of  $e^{1/y}$ , we have the estimate  $0 \leq y^{-n} = y y^{-(n+1)} \leq y(n+1)! e^{1/y}$ . Thus,  $0 \leq \lim_{y \downarrow 0} e^{-1/y} y^{-n} \leq \lim_{y \downarrow 0} y(n+1)! = 0$ , i.e.,  $\lim_{y \downarrow 0} \phi(y) y^{-n} = 0$ .

Setting  $n = 0$  implies that  $\lim_{y \downarrow 0} \phi(y) = 0$ . Using (5),  $\phi$  is continuous.

Using (5), the left-hand derivative of  $\phi$  at 0 is  $D_-^1 \phi(0) = \lim_{y \uparrow 0} [\phi(y) - \phi(0)]/y = 0$  and the right-hand derivative of  $\phi$  at 0 is  $D_+^1 \phi(0) = \lim_{y \downarrow 0} [\phi(y) - \phi(0)]/y = \lim_{y \downarrow 0} \phi(y) y^{-1} = 0$ . Thus,  $D^1 \phi(0) = 0$ . Using (6),  $D^1 \phi$  exists on  $\mathfrak{R}$  and is continuous.

Now we do the inductive step. Suppose  $D^n \phi$  exists on  $\mathfrak{R}$  for  $n \in \mathcal{N}$  and  $D^n \phi(0) = 0$ . Using (6),  $D_-^{n+1} \phi(0) = D_-^1 D^n \phi(0) = \lim_{y \uparrow 0} [D^n \phi(y) - D^n \phi(0)]/y = 0$  and  $D_+^{n+1} \phi(0) = D_+^1 D^n \phi(0) = \lim_{y \downarrow 0} [D^n \phi(y) - D^n \phi(0)]/y = \lim_{y \downarrow 0} p_n(y) \phi(y) y^{-(2n+1)} = p_n(0) \lim_{y \downarrow 0} \phi(y) y^{-(2n+1)} = 0$ . It follows that  $D^{n+1} \phi(0) = 0$ . Using (6),  $D^{n+1} \phi$  exists on  $\mathfrak{R}$  and is continuous.

It follows that  $\phi$  is  $\mathcal{C}^\infty$ . ■

**Proof of Lemma 2.12** Let  $X^*$  be the set of all continuous linear functionals  $h : X \rightarrow \mathfrak{R}$ . Local convexity of  $X$  ensures that, if  $x \in X$  is such that  $h(x) = 0$  for every  $h \in X^*$ , then  $x = 0$  (Dunford and Schwartz [7], Corollary V.2.13). Define  $H : X \rightarrow \mathfrak{R}^{X^*}$  by  $H(x) = (h(x))_{h \in X^*}$ . Give  $\mathfrak{R}^{X^*}$  the product topology. Consequently,  $H$  is continuous as every component function  $H_h = h$  is continuous. Moreover,  $H$  is injective; if  $H(x) = H(y)$  for some  $x, y \in X$ , then  $h(x - y) = h(x) - h(y) = 0$  for every  $h \in X^*$ , which implies  $x - y = 0$ . As  $O$  is compact and  $\mathfrak{R}^{X^*}$  is Hausdorff,  $H$  imbeds  $O$  in  $\mathfrak{R}^{X^*}$ . This implies  $H(O)$  is closed in  $\mathfrak{R}^{X^*}$  and metrizable.

Consider  $\mu \in \Delta(O)$  with  $|\text{supp } \mu| < \infty$ . The linearity of  $h \in X^*$  implies

$$\int_O \mu(dz) h(z) = \sum_{z \in \text{supp } \mu} \mu(\{z\}) h(z) = h \left( \sum_{z \in \text{supp } \mu} \mu(\{z\}) z \right) \quad (A.1)$$

Setting  $m_\mu = \sum_{z \in \text{supp } \mu} \mu(\{z\})z$ , we have  $m_\mu \in O$  as  $O$  is convex and  $\text{supp } \mu \subset O$ . Thus,  $H(m_\mu) \in H(O)$  for every  $\mu \in \Delta(O)$  with  $|\text{supp } \mu| < \infty$ .

Consider  $\mu \in \Delta(O)$ . As  $O$  is compact and metric, it is separable. Consequently, there exists a sequence  $(\mu_n) \subset \Delta(O)$  converging to  $\mu$  in the  $\mathcal{C}(O, \mathfrak{R})$  topology such that  $|\text{supp } \mu_n| < \infty$  for every  $n \in \mathcal{N}$  (Parthasarathy [14], Theorem II.6.3). By the above argument,  $m_{\mu_n}$  exists,  $m_{\mu_n} \in O$  and  $H(m_{\mu_n}) \in H(O)$  for every  $n \in \mathcal{N}$ . Using (A.1) and the definition of convergence in the  $\mathcal{C}(O, \mathfrak{R})$  topology, we have

$$\lim_{n \uparrow \infty} h(m_{\mu_n}) = \lim_{n \uparrow \infty} \int_O \mu_n(dz) h(z) = \int_O \mu(dz) h(z)$$

for every  $h \in X^*$ . Thus,  $\lim_{n \uparrow \infty} H(m_{\mu_n}) = (\int_O \mu(dz) h(z))_{h \in X^*}$ . As the sequence  $(H(m_{\mu_n})) \subset H(O)$  and  $H(O)$  is closed in  $\mathfrak{R}^{X^*}$  and metrizable, we have  $(\int_O \mu(dz) h(z))_{h \in X^*} \in H(O)$ . As  $H$  imbeds  $O$  in  $\mathfrak{R}^{X^*}$ , there exists a unique  $x \in O$  such that  $H(x) = (\int_O \mu(dz) h(z))_{h \in X^*}$ . By the definition of  $H$ , we have  $h(x) = \int_O \mu(dz) h(z)$  for every  $h \in X^*$ . Set  $m_\mu = x$ . ■

**Proof of Lemma 3.4** Suppose  $G$  is not strictly increasing on  $[0, 1]$ . Then, there exist  $b_1, b_2 \in [0, 1]$  such that  $b_1 < b_2$  and  $G(b_1) = G(b_2) = c$ . As  $\text{supp } G = [0, 1]$ , we have  $c \in (0, 1)$ . So, there exists  $N \in \mathcal{N}$  such that  $c + 1/N < 1$ . For every  $n \geq N$ , we have  $b(c + 1/n) > b_2 > b_1 \geq b(c)$ . Thus,  $b(c+) \geq b_2 > b(c)$ , i.e.,  $b$  is not continuous.

Conversely, suppose  $G$  is strictly increasing on  $[0, 1]$ . We already know that  $b$  is left-continuous. To show that  $b$  is right-continuous, fix  $\alpha \in \mathfrak{R}$ . If  $\alpha > 1$ , then  $b^{-1}((-\infty, \alpha)) = C$ , which is open in  $C$ . If  $\alpha \leq 0$ , then  $b^{-1}((-\infty, \alpha)) = \emptyset$ , which is open in  $C$ .

Finally, let  $\alpha \in (0, 1]$ . Consider  $c \in b^{-1}((-\infty, \alpha))$ . Then,  $c \in C$  and  $0 \leq b(c) < \alpha \leq 1$ . It follows that  $c \leq G(b(c)) < G(\alpha-)$  as  $G$  is strictly increasing on  $[0, 1]$ . So,  $b^{-1}((-\infty, \alpha)) \subset (-\infty, G(\alpha-)) \cap C$ . Conversely, suppose  $c \in C$  and  $c < G(\alpha-)$ . As  $\alpha > 0$ , there exists  $\epsilon > 0$  such that  $\alpha - \epsilon \in [0, 1]$  and  $c < G(\alpha - \epsilon)$ . It follows that  $b(c) \leq \alpha - \epsilon < \alpha$ . Consequently,  $b^{-1}((-\infty, \alpha)) \supset (-\infty, G(\alpha-)) \cap C$ . Thus,  $b^{-1}((-\infty, \alpha)) = (-\infty, G(\alpha-)) \cap C$ , which is open in  $C$ . We conclude that  $b$  is right-continuous.

Suppose  $b$  is not strictly increasing. Then, there exist  $c_1, c_2 \in C$  such that  $c_1 < c_2$  and  $b(c_1) = b(c_2)$ . For every  $n \in \mathcal{N}$ ,  $G(b(c_1)) = G(b(c_2)) \geq c_2 > c_1 > G(b(c_1) - 1/n)$ . Therefore,  $G(b(c_1)) > c_1 \geq G(b(c_1)-)$ , i.e.,  $G$  is not continuous.

Conversely, suppose  $G$  is not continuous. As  $G$  is right-continuous and increasing, there exists  $\alpha \in \mathfrak{R}$  such that  $G(\alpha) > G(\alpha-)$ . Let  $c_1, c_2 \in (G(\alpha-), G(\alpha))$  such that  $c_1 < c_2$ . As  $G(\alpha) > c_1$  and  $G$  is increasing,  $[\alpha, \infty) \subset G^{-1}([c_1, \infty))$ . If  $\beta < \alpha$ , then  $G(\beta) \leq G(\alpha-) < c_1$ , which means  $\beta \notin G^{-1}([c_1, \infty))$ . So,  $[\alpha, \infty) \supset G^{-1}([c_1, \infty))$ . Thus,  $G^{-1}([c_1, \infty)) = [\alpha, \infty)$ . Similarly,  $G^{-1}([c_2, \infty)) = [\alpha, \infty)$ . Therefore,  $b(c_1) = b(c_2)$ , i.e.,  $b$  is not strictly increasing. ■

**Proof of Lemma 3.5** Let  $\text{ba}(C, \mathcal{C})$  be the set of bounded, finitely additive real-valued functions defined on  $\mathcal{C}$ . With the total variation norm,  $\text{ba}(C, \mathcal{C})$  is a Banach space (Dunford and Schwartz [7], Sections III.7 and IV.9). Let  $B(C, \mathcal{C})$  be the set of real-valued functions on  $C$  that are uniform limits of simple measurable functions on  $(C, \mathcal{C})$ , i.e., functions that are finite linear combinations of characteristic functions of sets in  $\mathcal{C}$ . With the supremum norm,  $B(C, \mathcal{C})$  is a Banach space (Dunford and Schwartz [7], Section IV.5). By a Riesz-type representation theorem (Dunford and Schwartz [7], Theorem IV.5.1),  $\text{ba}(C, \mathcal{C})$  may be identified with  $B(C, \mathcal{C})^*$ , i.e., the conjugate of  $B(C, \mathcal{C})$ . The  $B(C, \mathcal{C})$  topology is the weakest topology on  $\text{ba}(C, \mathcal{C})$  that makes the linear functional  $\mu \mapsto \int_C \mu(dx) h(x)$  continuous for every  $h \in B(C, \mathcal{C})$ . As  $g$  is bounded and measurable,  $g \in B(C, \mathcal{C})$ . Thus,  $L$  is continuous.

By Alaoglu's theorem (Dunford and Schwartz [7], Theorem V.4.2), the closed unit sphere of  $\text{ba}(C, \mathcal{C})$  is compact. As  $\Delta(C)$  is a subset of the closed unit sphere of  $\text{ba}(C, \mathcal{C})$ ,  $\Delta(C)$  is compact if it is closed. Consider  $\lambda \in \text{ba}(C, \mathcal{C})$  that is an accumulation point of  $\Delta(C)$ . Then, there exists a net  $(\lambda_n) \subset \Delta(C)$  converging to  $\lambda$ . By the definition of the  $B(C, \mathcal{C})$  topology,  $\lim_n \lambda_n(E) = \lim_n \int_C \lambda_n(dx) 1_E(x) = \int_C \lambda(dx) 1_E(x) = \lambda(E)$  for every  $E \in \mathcal{C}$ . Consequently,  $\lambda(\emptyset) = 0$ ,  $\lambda(C) = 1$  and  $\lambda \geq 0$ . Consider sets  $E_1, \dots, E_k \in \mathcal{C}$  that are pairwise disjoint and  $E = \cup_{i=1}^k E_i$ . Then,  $E \in \mathcal{C}$ . As each  $\lambda_n$  is finitely additive, we have  $\lambda(E) = \lim_n \lambda_n(E) = \lim_n \sum_{i=1}^k \lambda_n(E_i) = \sum_{i=1}^k \lim_n \lambda_n(E_i) = \sum_{i=1}^k \lambda(E_i)$ , i.e.,  $\lambda$  is finitely additive. Thus,  $\lambda \in \Delta(C)$ , and consequently,  $\Delta(C)$  is closed in the  $B(C, \mathcal{C})$  topology.

Consider a net  $(\lambda_n) \subset \Delta(C, c)$  converging to  $\lambda$  in the  $B(C, \mathcal{C})$  topology. As  $\Delta(C, c) \subset \Delta(C)$  and  $\Delta(C)$  is closed,  $\lambda \in \Delta(C)$ . The identity map  $I$  on  $C$  is bounded and measurable. Therefore,  $I \in B(C, \mathcal{C})$ . Consequently,  $c \leq \lim_n \int_C \lambda_n(dx) x = \int_C \lambda(dx) x$ . Thus,  $\lambda \in \Delta(C, c)$ . As  $\Delta(C, c)$  is a closed subset of  $\Delta(C)$ ,  $\Delta(C, c)$  is compact. ■

**Proof of Lemma 3.6** Note that  $I - g_n, I - g \in \mathcal{C}([-b, \gamma], [-b, \gamma])$  and  $P \in \mathcal{C}([-b, \gamma], \mathfrak{R})$ . Consequently,  $P \circ (I - g_n), P \circ (I - g) \in \mathcal{C}([-b, \gamma], \mathfrak{R})$ . Endow  $\mathcal{C}([-b, \gamma], [-b, \gamma])$  and  $\mathcal{C}([-b, \gamma], \mathfrak{R})$  with their compact-open topologies. Given  $P$ , the composition mapping  $f \mapsto P \circ f$ , from  $\mathcal{C}([-b, \gamma], [-b, \gamma])$  to  $\mathcal{C}([-b, \gamma], \mathfrak{R})$ , is continuous (Dugundji [6], Theorem XII.2.1).

Since  $(g_n)$  converges uniformly to  $g$ ,  $(I - g_n)$  converges uniformly to  $I - g$ . By Theorem XII.8.2 in Dugundji [6], the compact-open topologies on  $\mathcal{C}([-b, \gamma], [-b, \gamma])$  and  $\mathcal{C}([-b, \gamma], \mathfrak{R})$  coincide with the respective topologies of uniform convergence. Thus,  $(I - g_n)$  converges to  $I - g$  in the compact-open topology of  $\mathcal{C}([-b, \gamma], [-b, \gamma])$ . It follows from the continuity of the composition mapping that  $P \circ (I - g_n)$  converges to  $P \circ (I - g)$  in the compact-open topology of  $\mathcal{C}([-b, \gamma], \mathfrak{R})$ . Therefore,  $P \circ (I - g_n)$  converges uniformly to  $P \circ (I - g)$ . It follows that  $(u_n)$  converges uniformly to  $u$ . ■

**Proof of Lemma 3.7** Fix  $r \in \{r_1, r_2\}$  and the subsequence  $(c(r, n_j))$  of  $(c(r, n))$ , converging to  $c^*(r)$ . For  $m, j \in \mathcal{N}$ , let  $F(m, j) = U(r, c(r, n_m); n_j)$ ,  $f(m) = U(r, c(r, n_m))$  and  $\phi(j) = U(r, c^*(r); n_j)$ . For  $(m_1, j_1), (m_2, j_2) \in \mathcal{N}^2$ , we say that  $(m_1, j_1) \succeq (m_2, j_2)$  if  $m_1 \geq m_2$  and  $j_1 \geq j_2$ . Then,  $(\mathcal{N}^2, \succeq)$  is a directed set and  $\{F(m, j) \mid (m, j) \in \mathcal{N}^2; \succeq\}$  is a net in  $\mathfrak{R}$ .

Fix  $\epsilon > 0$ . As  $(u_n)$  converges uniformly to  $u$  by Lemma 3.6, the subsequence  $(u_{n_j})$  converges uniformly to  $u$ , i.e., there exists  $J \in \mathcal{N}$  such that  $j \geq J$  implies  $\|u_{n_j} - u\| < \epsilon/8$ , where  $\|\cdot\|$  is the supremum norm. So,  $j \geq J$  implies the inequality  $|U(r, c; n_j) - U(r, c)| \leq \int_O \mu(r)(dx) |u_{n_j}(v(x), c) - u(v(x), c)| \leq \|u_{n_j} - u\| < \epsilon/8$  for every  $c \in C$ , i.e.,  $\lim_{j \uparrow \infty} U(r, c; n_j) = U(r, c)$  for every  $c \in C$ .

If we set  $c = c^*(r)$  in the above inequality, then  $j \geq J$  implies  $|\phi(j) - U(r, c^*(r))| = |U(r, c^*(r); n_j) - U(r, c^*(r))| < \epsilon/8$ . Thus,  $\lim_{j \uparrow \infty} \phi(j) = U(r, c^*(r))$ .

If we set  $c = c(r, n_m)$  in the above inequality, then  $m \in \mathcal{N}$  and  $j \geq J$  implies  $|F(m, j) - f(m)| = |U(r, c(r, n_m); n_j) - U(r, c(r, n_m))| < \epsilon/8$ . Consequently,  $m \in \mathcal{N}$  and  $j \geq J$  implies  $|F(m, j) - F(m, J)| \leq |F(m, j) - f(m)| + |F(m, J) - f(m)| < \epsilon/4$ .

The continuity of  $u_{n_j}$  and the bounded convergence theorem (Billingsley [4], Theorem 16.5) imply that  $\phi(j) = \lim_{m \uparrow \infty} F(m, j)$  for every  $j \in \mathcal{N}$ . So, given  $J$ , there exists  $M \in \mathcal{N}$  such that  $m \geq M$  implies  $|F(m, J) - \phi(J)| < \epsilon/8$ . Consequently,  $m \geq M$  implies  $|F(m, J) - F(M, J)| \leq |F(m, J) - \phi(J)| + |F(M, J) - \phi(J)| < \epsilon/4$ .

It follows that  $j \geq J$  and  $m \geq M$  implies  $|F(m, j) - F(M, J)| \leq |F(m, j) - F(m, J)| + |F(m, J) - F(M, J)| < \epsilon/2$ . Therefore, if  $(m_1, j_1) \succeq (M, J)$  and  $(m_2, j_2) \succeq (M, J)$ , then  $|F(m_1, j_1) - F(m_2, j_2)| \leq |F(m_1, j_1) - F(M, J)| + |F(m_2, j_2) - F(M, J)| < \epsilon$ . Thus,  $\{F(m, j) \mid (m, j) \in \mathcal{N}^2; \succeq\}$  is a Cauchy net with respect to the Euclidean space  $\mathfrak{R}$ . As the Euclidean space  $\mathfrak{R}$  is complete, this net converges to some  $p \in \mathfrak{R}$  (Dugundji [6], Theorem XIV.3.2). Consequently, there exists  $(M, J) \in \mathcal{N}^2$  such that  $(m, j) \succeq (M, J)$  implies  $|F(m, j) - p| < \epsilon$ .

Without loss of generality, let  $J \geq M$ . So,  $j \geq J$  implies  $(j, j) \succeq (J, J) \succeq (M, J)$ , which implies  $|F(j, j) - p| < \epsilon$ . So,  $\lim_{j \uparrow \infty} F(j, j) = p$ .

Moreover,  $|p - \phi(j)| = |p - \lim_{m \uparrow \infty} F(m, j)| = \lim_{m \uparrow \infty} |p - F(m, j)| \leq \epsilon$  for every  $j \geq J$ . Thus,  $\lim_{j \uparrow \infty} \phi(j) = p$ .

By definition,  $U(r, c(r, n_j); n_j) \geq U(r, c; n_j)$  for all  $j \in \mathcal{N}$  and  $c \in C$ . Therefore, we have  $U(r, c^*(r)) = \lim_{j \uparrow \infty} \phi(j) = p = \lim_{j \uparrow \infty} F(j, j) = \lim_{j \uparrow \infty} U(r, c(r, n_j); n_j) \geq \lim_{j \uparrow \infty} U(r, c; n_j) = U(r, c)$  for every  $c \in C$ . ■

## Appendix B: Classical results

For the sake of completeness, we recall two classical results that complement our results by showing the equivalence between a variation of  $\succeq_1$ , denoted

$\succeq_1^*$ , and relations  $\succeq_4$  and  $\succeq_5$  based on the notions of dilatations and martingale dominance respectively.

Specialize Assumption 2.1 by requiring  $X$  to be a topological vector space and  $O$  to be compact. For  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_1^* \lambda$  if  $\int_O \mu(dx) v(x) \leq \int_O \lambda(dx) v(x)$  for every continuous  $v \in \mathcal{V}$ . Clearly,  $\succeq_1^*$  is an extension of  $\succeq_1$ , i.e.,  $\succeq_1 \subset \succeq_1^*$ , since  $\mu \succeq_1 \lambda$  implies  $\mu \succeq_1^* \lambda$ .

A Markov kernel on  $(O, \mathcal{O})$  is a function  $P : O \times \mathcal{O} \rightarrow \mathfrak{R}$  such that  $P(x, \cdot) \in \Delta(O)$  for every  $x \in O$  and  $P(\cdot, E)$  is measurable for every  $E \in \mathcal{O}$ . A dilatation on  $(O, \mathcal{O})$  is a Markov kernel  $P$  on  $(O, \mathcal{O})$  such that  $f(\cdot) = \int_O P(\cdot, dx) f(x)$  for every continuous affine function  $f : O \rightarrow \mathfrak{R}$ .

**Definition 4.1** *Given  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_4 \lambda$  if there is a dilatation  $P$  on  $(O, \mathcal{O})$  such that  $\mu(\cdot) = \int_O \lambda(dx) P(x, \cdot)$ .*

The following result is Theorem 2 in Strassen [21].

**Theorem 4.2** *Given Assumption 2.1, if  $X$  is a locally convex topological vector space and  $O$  is compact and metrizable, then  $\succeq_1^* = \succeq_4$ .*

In order to state the other well-known equivalence result, we first define a martingale.

**Definition 4.3** *Let  $X$  be a separable Banach space. An ordered pair  $(f, g)$  is an  $X$ -valued martingale if there is a probability space  $(\Omega, \mathcal{F}, Q)$  such that*

- (a)  $f, g \in L_1(\Omega, \mathcal{F}, Q; X)$
- (b)  $\sigma(f) \equiv \{f^{-1}(B) \mid B \in \mathcal{B}(X)\} \subset \{g^{-1}(B) \mid B \in \mathcal{B}(X)\} \equiv \sigma(g)$ , and
- (c)  $\int_E Q(dx) f(x) = \int_E Q(dx) g(x)$  for every  $E \in \sigma(f)$ .

Condition (a) means that  $f$  and  $g$  are Bochner integrable random variables with values in  $X$ . Condition (b) means that  $(\sigma(f), \sigma(g))$  is a filtration, i.e., the  $\sigma$ -algebra  $\sigma(g)$  contains finer information about the state than the  $\sigma$ -algebra  $\sigma(f)$ . By definition,  $(f, g)$  is adapted to this filtration, i.e.,  $f$  is  $\sigma(f)/\mathcal{B}(X)$  measurable and  $g$  is  $\sigma(g)/\mathcal{B}(X)$  measurable. Condition (c) means that  $f$  is (a version of) the expectation of  $g$  conditional on  $f$ , i.e.,  $f = E_{\sigma(f)} g$ . Setting  $h = g - f$ , we have  $E_{\sigma(f)} h = E_{\sigma(f)} g - f = 0$ . Therefore, we have the representation  $g = f + h$ , where  $E_{\sigma(f)} h = 0$ , i.e.,  $g$  is  $f$  plus noise. This interpretation motivates the following definition.

**Definition 4.4** *Let  $X$  be a separable Banach space. Given  $\mu, \lambda \in \Delta(O)$ , we say that  $\mu \succeq_5 \lambda$  if there exists an  $X$ -valued martingale  $(f, g)$  with  $Q \circ f^{-1} = \lambda$  and  $Q \circ g^{-1} = \mu$ .*

The following result is a consequence of Theorem 8 in Strassen [21].

**Theorem 4.5** *In addition to the assumptions of Theorem 4.2, if  $X$  is a separable Banach space, then  $\succeq_1^* = \succeq_5$ .*

This result is useful because it characterizes a convenient source of lotteries that are ordered by the various riskiness relations: a set of lotteries on  $O$  is ordered by  $\succeq_1^*$  if and only if the lotteries are distributions of random variables that constitute an  $X$ -valued martingale.

## References

- [1] A. Atkinson, F. Bourguignon: The comparison of multi-dimensional distributions of economic status, *The Review of Economic Studies* 49, 183-201 (1982)
- [2] R. Bartle: *The Elements of Real Analysis*, Wiley, New York (1976)
- [3] C. Berge: *Topological Spaces*, Macmillan, New York (1963)
- [4] P. Billingsley: *Probability and Measure, 3rd ed*, Wiley-Interscience, New York (1995)
- [5] P. Diamond, J. Stiglitz: Increases in risk and in risk aversion, *Journal of Economic Theory* 8, 337-360 (1974)
- [6] J. Dugundji: *Topology*, Wm. C. Brown Publishers, Dubuque (1989)
- [7] N. Dunford, J. Schwartz: *Linear Operators, Part I: General Theory*, Wiley Interscience, New York (1988)
- [8] J. Hadar, W. Russell: Rules for ordering uncertain prospects, *American Economic Review* 59, 25-34 (1969)
- [9] G. Hanoeh, H. Levy: The efficiency analysis of choices involving risk, *The Review of Economic Studies* 36, 335-346 (1969)
- [10] S.-C. Kolm: Multidimensional egalitarianisms, *Quarterly Journal of Economics* 91, 1-13 (1977)
- [11] S. Lang: *Real and Functional Analysis*, Springer-Verlag, New York 1993
- [12] P. Milgrom, R. Weber: A theory of auctions and competitive bidding, *Econometrica* 50, 1089-1122 (1982)
- [13] L. Olson: Multivariate decision-making under risk aversion, *Journal of Economic Theory* 50, 193-203 (1990)
- [14] K. Parthasarathy: *Probability Measures on Metric Spaces*, Academic Press, New York (1967)
- [15] B. Pettis: On integration in vector spaces, *Transactions of the American Mathematical Society* 44, 277-304 (1938)

- [16] M. Rao: *Measure Theory and Integration*, Wiley Interscience, New York (1987)
- [17] R. Rockafellar: *Convex Analysis*, Princeton University Press, Princeton (1970)
- [18] M. Rothschild, J. Stiglitz: Increasing risk: I. A definition, *Journal of Economic Theory* 2, 225-243 (1970)
- [19] M. Rothschild, J. Stiglitz: Increasing risk II: Its economic consequences, *Journal of Economic Theory* 3, 66-84 (1971)
- [20] W. Russell, T. Seo: Ordering uncertain prospects: the multivariate utility functions case, *The Review of Economic Studies* 45, 605-610 (1978)
- [21] V. Strassen: The existence of probability measures with given marginals, *The Annals of Mathematical Statistics* 36, 423-439 (1965)
- [22] K.-Y. Tsui: Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach, *Journal of Economic Theory* 67, 251-265 (1995)