CDE
September 2006

COMPARATIVE RISK AVERSION WHEN THE OUTCOMES ARE VECTORS

Sudhir A. Shah
Email: sudhir@econdse.org

Delhi School of Economics
University of Delhi

Working Paper No. 149

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

Comparative risk aversion when the outcomes are vectors

Sudhir A. Shah*
Delhi School of Economics
University of Delhi

Pratt (1964) and Yaari (1969) contain the classical results pertaining to the equivalence of various notions of comparative risk aversion of von Neumann-Morgenstern utilities in the setting with real-valued outcomes. Some of these results have been extended to the setting with outcomes in $\Re^n$. We obtain analogues of the classical results in the setting with outcomes in ordered topological vector spaces when differentiability is not required, and in the setting with outcomes in ordered Hilbert spaces when differentiability is required, as is the case when we work with a vector-valued generalized notion of an Arrow-Pratt coefficient.

JEL classification: C02, D01, D81

Key words and phrases: Comparative risk aversion, vector space of outcomes, acceptance set, vector-valued risk premia, vector-valued Arrow-Pratt coefficient, Pettis integral, ordered topological vector spaces, ordered Hilbert spaces

---

* I am grateful to the Centre for Development Economics at the Delhi School of Economics for institutional support. All correspondence should be addressed to: Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007, India. Telephone: (+91)(11) 2766-7005. Fax: (+91)(11) 2766-7159. E-mail: sudhir@econdse.org

1. Introduction

Consider a von Neumann-Morgenstern utility function $u : O \to \Re$, where $O \subset \Re$ is the sets of outcomes. Let $\Delta(O)$ be the set of all probability measures (also called lotteries) on O. u is said to be risk averse (Pratt 1964, Arrow 1971) if it always rejects every non-degenerate fair lottery, i.e.,

$$u\left(\int_O \mu(dy)\, y\right) > \int_O \mu(dy)\, u(y) \tag{1.1}$$

for every non-degenerate lottery $\mu \in \Delta(O)$. By Jensen's inequality, this property is equivalent to u being strictly concave.

The salient formalizations of the notion that "a risk averse utility function u is more risk averse than a risk averse utility function v" are:

(a) every lottery accepted by u is accepted by v (Yaari 1969),

(b) the lottery dependent risk premia associated with u are weakly larger than the corresponding risk premia associated with v (Pratt 1964),

(c) $u = f \circ v$, where f is concave (Pratt 1964), and

(d) the Arrow-Pratt coefficient of absolute risk aversion is everywhere weakly larger for u than for v (Pratt 1964).

It is well-known (Pratt 1964, Yaari 1969) that, given appropriate regularity assumptions, these criteria yield the same partial ordering of the set of risk averse utility functions.

The classical theory described above has been extended to vector space settings with $O \subset \Re^n$ (Stiglitz 1969, Kihlstrom and Mirman 1974, Duncan 1977, Karni 1979, 1989), where n is a positive integer.[1] As in the classical setting, Jensen's inequality is used to characterize a risk averse utility function as strictly concave. The essential difference between the setting with $n = 1$ and the setting with $n > 1$ is that, in the former setting all increasing utility functions induce the same ordering on O, while in the latter setting this coincidence of orderings induced by increasing utility functions does not obtain. Kihlstrom and Mirman (1974) show that, if $n > 1$ and the risk aversion of utility functions u and v is comparable using definitions analogous to the above-mentioned notions, then u and v must induce the same ordering on O. This property, which we shall refer to as ordinal congruence, amounts to saying that in the vector outcome setting utility functions with comparable risk aversion must have the same level sets.

The technical approach used in Kihlstrom and Mirman (1974) to extend the classical theory to the setting with $n > 1$ is to reduce the problem to the classical setting and then exploit the classical results. Given a utility function, risk premia are defined for lotteries with 'skinny' supports, i.e., lotteries whose supports are contained in 1-dimensional affine subspaces intersecting O. They also construct Yaari-type acceptance sets using simple lotteries, i.e., lotteries with finite supports. Propositions 1 and 2 in Kihlstrom and Mirman (1974) demonstrate that the partial ordering of risk averse utility functions generated by (b) is equivalent to the partial ordering generated by (c) when the lotteries whose risk premia are considered belong to the 'skinny' class described above. Proposition 6 in that paper demonstrates that the partial ordering of risk averse utility functions generated by (a) is equivalent to the partial ordering generated by (c) when the lotteries considered for inclusion in the acceptance sets are simple. Some of these results have been generalized to arbitrary real vector spaces by Peters and Wakker (1987).

Our aim in this paper is to stretch the above-described theory in two directions. First, we wish to incorporate in our theory all lotteries in $\Delta(O)$, not only the simple lotteries and lotteries with skinny supports. Secondly, we wish to consider outcome spaces O that are more general than those considered in Kihlstrom and Mirman (1974). These objectives create the following dilemma. On the one hand, we wish to embed O in as general a vector space as possible, like in Peters and Wakker (1987). On the other hand, we need to compute expectations with respect to general lotteries over O, not only of real-valued utility functions, e.g., the integral on the right-hand-side of (1.1), but also of vector-valued mappings, e.g., the integral on the left-hand-side of (1.1). In the case of simple lotteries, computing these expectations is a straightforward matter of computing finite convex combinations of real numbers or vectors. For general lotteries, computing these expectations requires the outcome space to support the approximation and continuity arguments involved in the relevant method of integration. Another technical requirement of the theory is that the outcome space must be (partially) ordered so that the vector-valued risk premia generated by characterization (b) can be compared.

Our proposal for satisfying the above-listed desiderata is to embed O in a partially ordered real locally convex topological vector space. Given this setting, the integral on the right-hand-side of (1.1) will be the abstract Lebesgue integral, while the integral on the left-hand-side of (1.1) will be the Pettis integral. Using this setting, we show that characterizations (a), (b) and (c) are equivalent, with the definition of (b) modified to deal with the set-valued nature of risk premia in the vector outcome setting. Unlike in the proofs of analogous results in Kihlstrom and Mirman (1974), differentiability of the utility function is not required, and unlike the results in Peters and Wakker, our utility functions are real-valued, i.e., not allowed the values ∞ or $-\infty$.

In Duncan (1977), the definition of a risk premium vector is the first step in the derivation of a vector analogue of the Arrow-Pratt coefficient of absolute risk aversion. We generalize the definition of an Arrow-Pratt coefficient of absolute risk aversion to the setting of an ordered Hilbert space. In this setting, we show that characterizations (c) and (d) are equivalent.

Our generalization of the theory of comparative risk aversion suggests some natural applications which were hitherto beyond the scope of the formal theory. For instance, consider $O = \mathcal{C}(\Re_+, \Re)$, with an outcome $x \in O$ interpreted as the continuous sample path of a security's value. A lottery in this setting is a probability measure over the set of continuous sample paths. With appropriate assumptions, such measures describe diffusions, which are the formal elements of much of modern asset pricing theory.[2]

We begin our analysis by stating some preliminary definitions and results in Section 2. These results are used in Section 3 to establish a generalized version of Jensen's inequality that is appropriate for our setting and purposes. This inequality is used to characterize risk averse utility functions as strictly concave functions in our setting. In Section 4, Definitions 4.5, 4.9 and 4.10 define binary relations $\succeq^1$, $\succeq^2$ and $\succeq^3$ respectively on $\mathcal{U}$, which is the set of risk averse utility functions; these relations correspond to the classical notions of comparative risk aversion (a), (b) and (c) respectively. Theorem 4.15 is the first substantive result of this paper. It shows that, in an appropriate setting, $\succeq^1$, $\succeq^2$ and $\succeq^3$ are equivalent. In Section 5, we consider a more restrictive setting that allows the notion of differentiability to be used. In this context, we define a generalized notion of the Arrow-Pratt coefficient of absolute risk aversion. Using this notion, we define relation $\succeq^4$ on $\mathcal{U}$ corresponding to the classical notion (d) of comparative risk aversion. Theorem 5.5 is the second substantive result of this paper. It shows that, in an appropriate setting, $\succeq^3$ and $\succeq^4$ are equivalent.

2. Technical preliminaries

We start with two versions of the supporting hyperplane theorem tailored for our purposes. First, a convex set with nonempty interior is supported by a non-trivial hyperplane at any point in its frontier.

Lemma 2.1. *If*

(a) C is a convex subset of a topological vector space L with $\operatorname{Int} C \neq \emptyset$, *and*

(b) $x \in C - \operatorname{Int} C$,

then there exists a non-zero continuous linear functional $p : L \to \Re$ *such that* $p(x) \geq p(y)$ *for every* $y \in C$.

We use this result to generate a supporting (resp. strictly supporting) hyperplane at any given point in the graph of a concave (resp. strictly concave) function.

Lemma 2.2. *If*

(a) C *is a convex subset of a topological vector space* L,

(b) $u : C \to \Re_+$ *is concave, and*

(c) $x \in \operatorname{Int} C$,

then

(A) There exists $a \in \Re$ *and a continuous linear functional* $b : L \to \Re$ *such that* $a + b(x) = u(x)$ *and* $a + b(y) \geq u(y)$ *for every* $y \in C - \{x\}$.

(B) If, in addition, u *is strictly concave, then the inequality in (A) is strict.*

We shall employ, without explicit comment, the following conventions throughout this paper. First, a subset of a topological space is given the subspace topology. Secondly, a topological space is given the Borel σ-algebra, with $\mathcal{B}(Y)$ denoting the Borel σ-algebra of a topological space Y. Thirdly, the set of real numbers $\Re$ is given the Euclidean topology.

We endow the outcome space O with mathematical structure by making it a subset of a space X with the features described in the following assumption. This assumption applies throughout the rest of this paper with additional restrictions stated explicitly when required.

Assumption 2.3. O *is a nonempty subset of a real locally convex topological vector space* X. X^* *is the space of continuous real-valued linear functionals on* X.

$\Delta(O)$ is the set of probability measures on $(O, \mathcal{B}(O))$. $\mu \in \Delta(O)$ is said to be non-degenerate if there exists $B \in \mathcal{B}(O)$ such that $\mu(B) \in (0,1)$. We wish to define the mean of $\mu \in \Delta(O)$, i.e., give meaning to the integral $\int_O \mu(dy)\, y$. This will be done using the notion of a Pettis integral.

Definition 2.4. *Consider a probability space* $(\Omega, \mathcal{F}, P)$ *and an* $\mathcal{F}/\mathcal{B}(X)$ *measurable function* $x : \Omega \to X$. x *is Pettis integrable*[3] *over* Ω *if*

(a) $\int_\Omega P(d\omega)\, x^* \circ x(\omega)$ *exists for every* $x^* \in X^*$, *and*

(b) there exists $x_\Omega \in X$ *such that* $x^*(x_\Omega) = \int_\Omega P(d\omega)\, x^* \circ x(\omega)$ *for every* $x^* \in X^*$.

If x is Pettis integrable over Ω and x_Ω is unique, then we refer to x_Ω as the Pettis integral of x over Ω and denote it by $\int_\Omega P(d\omega)\, x(\omega)$.

Remark 2.5. *X^* is a total space of linear functionals on X (Dunford and Schwartz 1988, Corollary V.2.13), i.e., if $x \in X$ is such that $x^*(x) = 0$ for every $x^* \in X^*$, then $x = 0$. Therefore, if x is Pettis integrable over Ω, then x_Ω is uniquely determined.*

We shall use the Pettis integral to define $\int_O \mu(dy)\, I(y)$, where $\mu \in \Delta(O)$ and $I : O \to X$ is the identity function. This entails the following problem: find $x_O \in X$ such that $x^*(x_O) = \int_O \mu(dz)\, x^*(z)$ for every $x^* \in X^*$. Although this problem is a special case of Definition 2.4, the general problem of Definition 2.4 can be reduced to the problem of solving the special problem. To see the reason for this, note that $\int_\Omega P(d\omega)\, x^* \circ x(\omega) = \int_{x(\Omega)} P \circ x^{-1}(dz)\, x^*(z)$ for every $x^* \in X^*$. Thus, the integrals characterizing x_Ω can be calculated by integrating the identity function over $x(\Omega)$ using the image measure $P \circ x^{-1}$. We provide the details in Remark A.1 in the Appendix.

A second issue is whether there are general settings in which functions are Pettis integrable. In Remark A.2 of the Appendix, we identify a general class of settings in which Pettis integrals exist.

The following is a generalized version of the classical Jensen's inequality. We shall use it to characterize risk averse utility functions. The proof is an application of Lemma 2.2.

Theorem 2.6. *(Jensen's inequality)*[4] *If*

(a) O is convex,

(b) $\mu \in \Delta(O)$ is such that, given $(O, \mathcal{B}(O), \mu)$, the identity function $I : O \to X$ is Pettis integrable over O and $\int_O \mu(dy)\, y \in \operatorname{Int} O$, and

(c) $u : O \to \Re$ is concave, bounded below and $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable,

then

(A) $\int_O \mu(dy)\, u(y)$ exists and $u(\int_O \mu(dy)\, y) \geq \int_O \mu(dy)\, u(y)$.

(B) If u is strictly concave and μ is non-degenerate, then the inequality in (A) is strict.

If $X = \Re$, then the Pettis integral in (A) reduces to the generalized Lebesgue integral, thereby yielding the classical Jensen's inequality.

3. Risk averse utility functions

Definition 3.1. *$\Delta(O)_0$ is the set of $\mu \in \Delta(O)$ such that $m_\mu = \int_O \mu(dy)\, y$ exists and $m_\mu \in O$.*

A preference is risk averse if it strictly prefers the mean m_μ of a non-degenerate lottery μ to the lottery itself. More formally, we have the following definition.

Definition 3.2. *$u : O \to \Re$ is said to be risk averse if*

(a) u is $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable, and

for every non-degenerate $\mu \in \Delta(O)_0$,

(b) $\int_O \mu(dy)\, u(y)$ exists, and

(c) $u(m_\mu) > \int_O \mu(dy)\, u(y)$.

Note that the property of being risk averse is invariant across equivalent von Neumann-Morgenstern utility representations of the same preference.

Remark 3.3. *If $u : O \to \Re$ is risk averse and $v = a + bu$, with $a \in \Re$ and $b \in \Re_{++}$, then v is risk averse.*

The following result provides sufficient conditions for a utility function to be risk averse.

Theorem 3.4. *If*

(a) O is open and convex, and

(b) $u : O \to \Re$ is $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable, bounded below and strictly concave,

then

(A) $\int_O \mu(dy)\, u(y)$ exists for every $\mu \in \Delta(O)_0$, and

(B) u is risk averse.

Proof. (A) Consider $\mu \in \Delta(O)_0$. By definition and (a), $\int_O \mu(dy)\, y$ exists and $\int_O \mu(dy)\, y \in O = \operatorname{Int} O$. By Theorem 2.6(A), $\int_O \mu(dy)\, u(y)$ exists.

(B) Condition 3.2(a) is satisfied by hypothesis. Consider a non-degenerate $\mu \in \Delta(O)_0$. By (A), condition 3.2(b) is satisfied. Moreover, Theorem 2.6(B) implies $u(m_\mu) > \int_O \mu(dy)\, u(y)$, thereby satisfying condition 3.2(c). ∎

4. Comparing risk aversion without differentiability

We start by imposing an ordering structure on X.

Assumption 4.1. *$\geq$ is a partial order on X (i.e., $\geq$ is reflexive, transitive and antisymmetric) such that*

(a) $x \geq y$ implies $x + z \geq y + z$ for every $z \in X$, and

(b) $x \geq y$ implies $tx \geq ty$ for every $t \in \Re_{++}$.

Let $x > y$ if and only if $x \geq y$ and $x \neq y$. Let $X_+ = \{x \in X \mid x > 0\}$. There exists $e \in X_+$ such that, for every $x \in X$, there exists $t \in \Re_{++}$ such that $te > x$.

We say that $u : O \to \Re$ is increasing if for all $x, y \in O$, $x > y$ implies $u(x) > u(y)$. Let $\mathcal{U}$ be the set of functions $u : O \to \Re$ that are continuous, increasing, bounded below and strictly concave.

Lemma 4.2. *If $u \in \mathcal{U}$ and O is open and convex, then $\int_O \mu(dy)\, u(y)$ exists for every $\mu \in \Delta(O)_0$ and u is risk averse.*

Proof. The result follows from Theorem 3.4. ■

In this section, we shall study various methods of ordering the elements of $\mathcal{U}$. Given Lemma 4.2, these orderings can legitimately be interpreted as comparisons of risk aversion of the elements of $\mathcal{U}$ if O is open and convex.

Definition 4.3. *Given $x \in O$ and $u \in \mathcal{U}$, $A(x, u) = \{\mu \in \Delta(O)_0 \mid u(x) \leq \int_O \mu(dy)\, u(y)\}$ is called the acceptance set associated with x and u.*

$A(x, u)$ may be interpreted as the set of lotteries that utility function u would accept (i.e., not reject) given the *status quo* outcome x. Note that acceptance sets are determined by the preference and are invariant across equivalent von Neumann-Morgenstern representations of that preference.

Remark 4.4. *If $u \in \mathcal{U}$ and $v = a + bu$, where $a \in \Re$ and $b \in \Re_{++}$, then $v \in \mathcal{U}$ and $A(x, u) = A(x, v)$ for every $x \in O$.*

Our first notion of comparative risk aversion labels a risk averse von Neumann-Morgenstern utility u as more risk averse than a risk averse von Neumann-Morgenstern utility v if every lottery accepted by u is accepted by v. In the light of Remark 4.4, this amounts to comparing the risk aversion of the underlying preferences.

Definition 4.5. *For all $u, v \in \mathcal{U}$, $u \succeq^1 v$ if and only if $A(x, u) \subset A(x, v)$ for every $x \in O$.*

We now define the notion of risk premia associated with a lottery μ.

Definition 4.6. *Given $u \in \mathcal{U}$ and $\mu \in \Delta(O)_0$,*

$$\pi(u, \mu) = \left\{ \pi \in X \;\middle|\; m_\mu - \pi \in O \quad \wedge \quad u(m_\mu - \pi) = \int_O \mu(dy)\, u(y) \right\}$$

Unlike in the case of scalar outcomes, the set of risk premia, $\pi(u, \mu)$, is generally not a singleton set. Nor is it necessary that they be positive vectors. We note a simple condition on O that guarantees the non-emptiness of the set of risk premia.

Lemma 4.7. *If O is nonempty, convex and open in X, $u \in \mathcal{U}$ and $\mu \in \Delta(O)_0$, then $\pi(u,\mu) \neq \emptyset$.*

Proof. As O is convex, it is connected. As O is nonempty and connected, and u is continuous, $u(O) \subset \Re$ is nonempty and connected. Therefore, $u(m_\mu) > \int_O \mu(dy)\, u(y) \in u(O)$, i.e., there exists $x(u,\mu) \in O$ such that $u \circ x(u,\mu) = \int_O \mu(dy)\, u(y)$. Therefore, $\pi(u,\mu) \neq \emptyset$ as $m_\mu - x(u,\mu) \in \pi(u,\mu)$. ■

We also note that risk premia are determined by preferences over lotteries, i.e., they are invariant with respect to increasing affine transformations of a von Neumann-Morgenstern representation of the preference.

Remark 4.8. *Suppose $u \in \mathcal{U}$ and $v = a + bu$, where $a \in \Re$ and $b \in \Re_{++}$. Then, $v \in \mathcal{U}$ and $\pi(u,\mu) = \pi(v,\mu)$ for every $\mu \in \Delta(O)_0$.*

Our second notion of comparative risk aversion labels a risk averse von Neumann-Morgenstern utility u as more risk averse than a risk averse von Neumann-Morgenstern utility v if a risk premium for v can never exceed a risk premium for u. In the light of Remark 4.8, this amounts to comparing the risk aversion of the underlying preferences. In order to formalize this notion, define the binary relation $\geq^*$ on 2^X by: for all $A, B \in 2^X$,

$$A \geq^* B \qquad \Leftrightarrow \qquad \neg y > x, \quad \forall x \in A \; \forall y \in B$$

Definition 4.9. *For all $u, v \in \mathcal{U}$, $u \succeq^2 v$ if and only if $\pi(u,\mu) \geq^* \pi(v,\mu)$ for every $\mu \in \Delta(O)_0$.*

Our third notion of comparative risk aversion labels a risk averse von Neumann-Morgenstern utility u as more risk averse than a risk averse von Neumann-Morgenstern utility v if u is an increasing concave transformation of v. Given $B \subset \Re$, we say that $f : B \to \Re$ is increasing if $x, y \in B$ and $x < y$ implies $f(x) < f(y)$.

Definition 4.10. *For all $u, v \in \mathcal{U}$, $u \succeq^3 v$ if and only if $u = f \circ v$ for some $f : v(O) \to \Re$ that is increasing and concave.*

Theorem 4.15 is the first substantive result of this paper. It shows that in a very general setting, the three notions of comparative risk aversion listed above are equivalent. We start with a definition and some preliminary lemmas.

Definition 4.11. *$u : O \to \Re$ and $v : O \to \Re$ are said to be ordinally congruent if, for all $x, y \in O$, $u(x) \geq u(y)$ if and only if $v(x) \geq v(y)$.*

Lemma 4.12. *If*

(a) $u, v \in \mathcal{U}$ are ordinally congruent, and

(b) $x, y \in O$ such that $v(x) \neq v(y)$,

then there exists $a \in \Re$ and $b \in \Re_{++}$ such that $a + bv(x) = u(x)$ and $a + bv(y) = u(y)$. Define $w \in \mathcal{U}$ by $w = a+bv$. Then, $A(x, w) = A(x, v)$ for every $x \in O$ and $\pi(w, \mu) = \pi(v, \mu)$ for every $\mu \in \Delta(O)_0$.

Proof. Set

$$a = u(x) - bv(x) \qquad \text{and} \qquad b = \frac{u(x) - u(y)}{v(x) - v(y)}$$

As u and v are ordinally congruent by (a), we have $b > 0$. The other claims follow from Remarks 4.4 and 4.8. ■

Lemma 4.13. *If $u, v \in \mathcal{U}$ are ordinally congruent, then there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$; moreover, f is increasing.*

Proof. For $r \in v(O)$, let $f(r) = u \circ v^{-1}(\{r\})$; by ordinal congruence, u is constant over $v^{-1}(\{r\})$. It follows that $f \circ v(x) = u \circ v^{-1}(\{v(x)\}) = u(x)$ for every $x \in O$. It is routine to show that f is unique and increasing on $v(O)$. ■

Lemma 4.14. *If*

(a) O is connected,

(b) $u : O \to \Re$ is continuous and $v : O \to \Re$, and

(c) $f : v(O) \to \Re$ is an increasing function such that $u = f \circ v$,

then f is continuous.

Proof. Suppose f is discontinuous at $v^* \in v(O)$. Let $A = \{f(r) \mid r \in v(O) \cap (-\infty, v^*)\}$ and $B = \{f(r) \mid r \in v(O) \cap (v^*, \infty)\}$. If $v^* \in (\inf v(O), \sup v(O))$, then $A \neq \emptyset \neq B$ and either $f(v^*) < \inf B$ or $f(v^*) > \sup A$. If $v^* = \inf v(O)$, then $A = \emptyset$, $B \neq \emptyset$ and $f(v^*) < \inf B$. If $v^* = \sup v(O)$, then $A \neq \emptyset$, $B = \emptyset$ and $f(v^*) > \sup A$. So, there can be two cases: either $f(v^*) < \inf B$ and $B \neq \emptyset$, or $f(v^*) > \sup A$ and $A \neq \emptyset$. We show a contradiction in the first case. A contradiction can be derived for the second case in analogous fashion.

As $f(v^*) < \inf B$, there exists $\alpha \in \Re$ such that $f(v^*) < \alpha < \inf B$. Consider the sets $f \circ v(O) \cap (-\infty, \alpha)$ and $f \circ v(O) \cap (\alpha, \infty)$. Clearly, these sets are disjoint. As f is increasing, their union equals $f \circ v(O)$. Clearly, $f(v^*) \in f \circ v(O) \cap (-\infty, \alpha)$. Moreover,

$\emptyset \neq B \subset f \circ v(O) \cap (\alpha, \infty)$. Thus, these sets form a disconnection of $f \circ v(O) = u(O)$, contradicting the facts that u is continuous and O is connected. ■

We are now ready to prove our first main result.

Theorem 4.15. *If*

(a) O is nonempty, convex and open in X,

(b) $x + X_+ \subset O$ for every $x \in O$, and

(c) $u, v \in \mathcal{U}$ are ordinally congruent,

then $u \succeq^1 v \Leftrightarrow u \succeq^2 v \Leftrightarrow u \succeq^3 v$.

Proof. We start with some general observations. First, by Lemma 4.2, $\int_O \mu(dy)\, u(y)$ and $\int_O \mu(dy)\, v(y)$ exist for every $\mu \in \Delta(O)_0$; moreover, u and v are risk averse. Secondly, by Lemma 4.7, $\pi(u, \mu) \neq \emptyset \neq \pi(v, \mu)$ for $\mu \in \Delta(O)_0$. Thirdly, as O is convex, it is connected.

(i) Suppose $\neg u \succeq^2 v$. Then, $\neg \pi(u, \mu) \geq^* \pi(v, \mu)$ for some $\mu \in \Delta(O)_0$. This means $\pi_v > \pi_u$ for some $\pi_u \in \pi(u, \mu)$ and $\pi_v \in \pi(v, \mu)$. As u is risk averse, $u(m_\mu) > \int_O \mu(dy)\, u(y) = u(m_\mu - \pi_u)$; clearly, $m_\mu \neq m_\mu - \pi_u$. Using Lemma 4.12, we assume without loss of generality that $v(m_\mu - \pi_u) = u(m_\mu - \pi_u)$. As $\mu \in \Delta(O)_0$,

$$\int_O \mu(dy)\, v(y) = v(m_\mu - \pi_v) < v(m_\mu - \pi_u) = u(m_\mu - \pi_u) = \int_O \mu(dy)\, u(y)$$

where the inequality follows from Assumption 4.1 and the fact that v is increasing. It follows that $\mu \in A(m_\mu - \pi_u, u)$ and $\mu \notin A(m_\mu - \pi_u, v)$. It follows that $\neg u \succeq^1 v$. Thus, $u \succeq^1 v$ implies $u \succeq^2 v$.

(ii) Suppose $u \succeq^2 v$. This means $\pi(u, \mu) \geq^* \pi(v, \mu)$ for every $\mu \in \Delta(O)_0$. Let $x \in O$ and $\mu \in A(x, u)$. Let $\pi_u \in \pi(u, \mu)$ and $\pi_v \in \pi(v, \mu)$.

Suppose $v(x) = v(m_\mu - \pi_v) = \int_O \mu(dy)\, v(y)$. Then, $\mu \in A(x, v)$.

Suppose $v(x) \neq v(m_\mu - \pi_v)$. Using Lemma 4.12, we assume without loss of generality that $v(x) = u(x)$ and $v(m_\mu - \pi_v) = u(m_\mu - \pi_v)$.

Suppose $u(m_\mu - \pi_u) > u(m_\mu - \pi_v)$ and there exists $r \in X_+$ such that $u(m_\mu - \pi_u) = u(m_\mu - \pi_v + r)$. By definition, $\int_O \mu(dy)\, u(y) = u(m_\mu - \pi_u) = u(m_\mu - \pi_v + r)$. Thus, $\pi_v \in \pi(v, \mu)$ and $\pi_v - r \in \pi(u, \mu)$. As $\pi_v > \pi_v - r$, we have $\neg \pi(u, \mu) \geq^* \pi(v, \mu)$, a contradiction.

Suppose $u(m_\mu - \pi_u) > u(m_\mu - \pi_v)$ and there does not exist $r \in X_+$ such that $u(m_\mu - \pi_u) = u(m_\mu - \pi_v + r)$. By Assumption 4.1, $A = \{t \in \Re_+ \mid u(m_\mu - \pi_u) < u(m_\mu - \pi_v + te)\} \neq \emptyset$, and by hypothesis, $0 \in B = \{t \in \Re_+ \mid u(m_\mu - \pi_u) > u(m_\mu - \pi_v + te)\}$.

Thus, A and B are nonempty, $A \cap B = \emptyset$ and $A \cup B = \Re_+$. As u is continuous, A and B are open in $\Re_+$. Thus, A and B are a disconnection of $\Re_+$, a contradiction.

Therefore, $u(m_\mu - \pi_u) \leq u(m_\mu - \pi_v)$, and it follows that $v(x) = u(x) \leq \int_O \mu(dy)\, u(y) = u(m_\mu - \pi_u) \leq u(m_\mu - \pi_v) = v(m_\mu - \pi_v) = \int_O \mu(dy)\, v(y)$. Thus, $\mu \in A(x, v)$. Consequently, $A(x, u) \subset A(x, v)$ for every $x \in O$. It follows that $u \succeq^1 v$.

(iii) By Lemma 4.13, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$; moreover, f is increasing. By Lemma 4.14, f is continuous, and therefore, $\mathcal{B}(v(O))/\mathcal{B}(\Re)$ measurable. Suppose $u \succeq^3 v$. It follows that f is concave.

Consider $x \in O$ and $\mu \in A(x, u)$. By definition, $\mu \in \Delta(O)_0$. It follows that m_μ exists, $m_\mu \in O$, $\int_O \mu(dy)\, u(y)$ exists and u is risk averse. By definition, $\int_O \mu(dy)\, f \circ v(y) = \int_O \mu(dy)\, u(y) \geq u(x) = f \circ v(x)$.

Suppose there exists $v^* \in \Re$ such that $v(y) = v^*$ μ-a.s. Then,

$$f\left(\int_O \mu(dy)\, v(y)\right) = f(v^*) = \int_O \mu(dy)\, f(v^*) = \int_O \mu(dy)\, f \circ v(y) \geq f \circ v(x)$$

As f is increasing, we have $\int_O \mu(dy)\, v(y) \geq v(x)$. Thus, $\mu \in A(x, v)$.

Suppose there does not exist $v^* \in \Re$ such that $v(y) = v^*$ μ-a.s. As v is continuous and O is connected, $v(O)$ is a connected subset of $\Re$, i.e., $v(O)$ is an interval. Let $\bar{v} = \sup v(O)$ and $\underline{v} = \inf v(O)$. Clearly, $\int_O \mu(dy)\, v(y) = \int_{v(O)} \mu \circ v^{-1}(dz)\, z \leq \bar{v}$. If $\bar{v} \notin v(O)$, then $\bar{v} > v(y)$ for every $y \in O$ and $\int_O \mu(dy)\, v(y) = \int_{v(O)} \mu \circ v^{-1}(dz)\, z < \bar{v}$. If $\bar{v} \in v(O)$ and $\int_O \mu(dy)\, v(y) = \int_{v(O)} \mu \circ v^{-1}(dz)\, z = \bar{v}$, then $v = \bar{v}$ μ-a.s., a contradiction. Thus, $\int_{v(O)} \mu \circ v^{-1}(dz)\, z < \bar{v}$. By an analogous argument, $\int_{v(O)} \mu \circ v^{-1}(dz)\, z > \underline{v}$. Thus, $\int_{v(O)} \mu \circ v^{-1}(dz)\, z \in \operatorname{Int} v(O)$. As u is bounded below, so is f. As f is concave, Theorem 2.6 implies

$$\begin{aligned} f\left(\int_O \mu(dy)\, v(y)\right) = f\left(\int_{v(O)} \mu \circ v^{-1}(dz)\, z\right) &\geq \int_{v(O)} \mu \circ v^{-1}(dz)\, f(z) \\ &= \int_O \mu(dy)\, f \circ v(y) \\ &= \int_O \mu(dy)\, u(y) \\ &\geq f \circ v(x) \end{aligned}$$

As f is increasing, $\int_O \mu(dy)\, v(y) \geq v(x)$, which implies $\mu \in A(x, v)$. It follows that $u \succeq^1 v$.

(iv) By Lemma 4.13, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$; moreover, f is increasing. By Lemma 4.14, f is continuous. Suppose $\neg u \succeq^3 v$. Thus, f is not concave.

Then, there exist $v_1, v_2 \in v(O)$ and $t \in (0,1)$ such that $f(tv_1 + (1-t)v_2) < tf(v_1) + (1-t)f(v_2)$. As $v(O) \subset \Re$ is connected, $tv_1 + (1-t)v_2 \in v(O)$; moreover, as f is continuous, $f \circ v(O)$ is connected. Without loss of generality, suppose $v_1 < v_2$. As f is increasing, $f(v_1) < f(v_2)$. Thus, $f(tv_1 + (1-t)v_2) < tf(v_1) + (1-t)f(v_2) < f(v_2)$.

As $tv_1 + (1-t)v_2 \in v(O)$ and $v_2 \in v(O)$, we have

$$[f(tv_1 + (1-t)v_2), f(v_2)] \subset f \circ v(O)$$

Consequently, there exists $v^* \in v(O)$ such that

$$f(tv_1 + (1-t)v_2) < f(v^*) < tf(v_1) + (1-t)f(v_2)$$

By definition, there exist $x_1, x_2, y \in O$ such that $v(x_1) = v_1$, $v(x_2) = v_2$ and $v(y) = v^*$. Thus,

$$\begin{aligned} u(y) = f \circ v(y) = f(v^*) &< tf(v_1) + (1-t)f(v_2) \\ &= tf \circ v(x_1) + (1-t)f \circ v(x_2) \\ &= tu(x_1) + (1-t)u(x_2) \end{aligned}$$

As f is increasing, we have

$$tv(x_1) + (1-t)v(x_2) = tv_1 + (1-t)v_2 < v^* = v(y)$$

Define $\mu \in \Delta(O)$ by $\mu = t\delta_{x_1} + (1-t)\delta_{x_2}$, where δ_{x_i} is the Dirac delta measure at x_i. Then, $\int_O \mu(dx)\, u(x) = tu(x_1) + (1-t)u(x_2) > u(y)$. Consequently, $\mu \in A(y,u)$. Also, $\int_O \mu(dx)\, v(x) = tv(x_1) + (1-t)v(x_2) < v(y)$, i.e., $\mu \notin A(y,v)$. It follows that $\neg u \succeq^1 v$. Thus, $u \succeq^1 v$ implies $u \succeq^3 v$. ■

In Theorem 4.15 and the lemmas leading up to it, we have used the notion of ordinal congruence. The following result formalizes in our setting the observation, Proposition 5 of Kihlstrom and Mirman (1974), that ordinal congruence is necessary in order to compare the risk aversion of utility functions.

Theorem 4.16. *If*

(a) O is convex,

(b) $x + X_+ \subset O$ for every $x \in O$, and

(c) $u, v \in \mathcal{U}$ are not ordinally congruent,

then $\neg u \succeq_1 v$ and $\neg v \succeq_1 u$.

Proof. As u and v are not ordinally congruent, there exist $x, y \in O$ such that $u(x) \geq u(y)$ and $v(x) < v(y)$. Using (b) and (c), we may, without loss of generality, assume that

$u(x) > u(y)$ and $v(x) < v(y)$. Consider $\mu = \delta_x/2 + \delta_y/2$. Clearly, $\mu \in A_v(x)$ and $\mu \notin A_u(x)$. Thus, $\neg v \succeq_1 u$. Similarly, $\mu \notin A_v(y)$ and $\mu \in A_u(y)$. Thus, $\neg u \succeq_1 v$. ■

5. Comparing risk aversion with differentiability

The fourth notion of comparative risk aversion, which is the most important from the point of view of applications and computations, is in terms of the size of the Arrow-Pratt coefficient of absolute risk aversion. As this notion involves differentiable utilities, it cannot be defined in as general a setting as that of Section 4. However, the Arrow-Pratt coefficient defined for $X = \Re$ and the generalized measure of Duncan defined for $X = \Re^n$ can be generalized to the setting where X is a Hilbert space with appropriate ordering structure. Once this is done, the classical result can be re-formulated. We specialize Assumptions 2.3 and 4.1 by giving X and $\geq$ more specific forms.

Assumption 5.1. *$(X, \langle .,. \rangle)$ is a real Hilbert space with $X \neq \{0\}$, inner product $\langle .,. \rangle$, and a Hilbert basis $\{b_i \mid i \in I\}$.*[5] *For $x \in X$, we say that $x \geq 0$ if $\langle x, b_i \rangle \geq 0$ for every $i \in I$; we say that $x > 0$ if $x \geq 0$ and $x \neq 0$. Let $X_+ = \{x \in X \mid x > 0\}$. For $x, y \in X$, we say that $x \geq y$ if $x - y \geq 0$. Finally, there exists $e \in X_+$ such that, for every $x \in X$, there exists $t \in \Re_{++}$ such that $te > x$.*

Note that $\{b_i \mid i \in I\} \subset X_+$ as the family is orthonormal. Therefore, $\|b_i\| = \langle b_i, b_i \rangle^{1/2} = 1$ for every $i \in I$. It is also trivial to check that $\geq$ is a partial order on $(X, \langle .,. \rangle)$ that satisfies Assumption 4.1; the antisymmetry of $\geq$ follows immediately from the fact that $\{b_i \mid i \in I\}$ is a total family in $(X, \langle .,. \rangle)$. Also, it is easily confirmed that $X_+ \cup \{0\}$ is a convex cone.

Consider a twice differentiable $u \in \mathcal{U}$. The (Fréchet) derivative of u is a mapping $Du : O \to \mathcal{L}(X, \Re)$, where $\mathcal{L}(X, \Re)$ is the space of continuous linear real-valued functionals on X. The second derivative of u is the derivative of Du, i.e., $D^2u : O \to \mathcal{L}(X, \mathcal{L}(X, \Re))$, where $\mathcal{L}(X, \mathcal{L}(X, \Re))$ is the space of continuous linear maps from X to $\mathcal{L}(X, \Re)$. As X is a Hilbert space, $\mathcal{L}(X, \Re)$ is isomorphic to X (Lang 1993, Theorem V.2.1). Thus, we may set $D^2u : O \to \mathcal{L}(X, X)$.

We define the generalized Arrow-Pratt coefficient of absolute risk aversion of $u \in \mathcal{U}$ by

$$a_u(x) = \frac{-D^2u(x)Du(x)}{\|Du(x)\|^2}$$

In the case $X = \Re$, this formula reduces to the classical Arrow-Pratt coefficient. As $D^2u(x) \in \mathcal{L}(X, X)$ and $Du(x) \in X$, it follows that $a_u(x) \in X$.

Definition 5.2. *For all $u, v \in \mathcal{U}$, $u \succeq^4 v$ if and only if $a_u(x) \geq a_v(x)$ for every $x \in O$.*

Consider $u, v \in \mathcal{U}$ that are ordinally congruent. By Lemma 4.13, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$; moreover, f is increasing. By Lemma 4.14, f is continuous. We now establish some other regularity properties of f.

Lemma 5.3. *If*

(a) $O \subset X$ is nonempty and open,

(b) $x + X_+ \subset O$ for every $x \in O$, and

(c) $u, v \in \mathcal{U}$ are twice differentiable and ordinally congruent,

then there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$; moreover, f is increasing and twice differentiable.

Proof. By Lemma 4.13, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$. Moreover, f is increasing and by Lemma 4.14, f is continuous. We show that f is twice differentiable.

Fix $x \in O$. As O is open, there exists $t \in (0, 1]$ such that $x - te \in O$; otherwise, $x - e/n \in X - O$ for every $n \in \mathcal{N}$, which implies $x \in X - O$ as $X - O$ is closed, a contradiction. Define $\bar{e} = te > 0$. Let $E = \{y \in O \mid x - \bar{e} < y < x + \bar{e}\}$. Define $w : (0, 2) \to \Re$ by $w(r) = v(x - \bar{e} + r\bar{e})$. As $\bar{e} > 0$ and v is increasing, w is increasing. As v is continuous, w is continuous. Moreover, (c) implies that w is twice differentiable. Let w^{-1} be the function inverse of w. Clearly, w^{-1} is increasing, and by the inverse function theorem, twice differentiable.

Also note that $w((0, 2)) = v(E)$. Clearly, $w((0, 2)) \subset v(E)$. Suppose there exists $r \in v(E) - w((0, 2))$. Consider the sets $w^{-1}((-\infty, r))$ and $w^{-1}((r, \infty))$. Clearly, these sets are nonempty and disjoint. As w is continuous, these sets are open subsets of $(0, 2)$. As $r \notin w((0, 2))$, we have $(0, 2) \subset w^{-1}((-\infty, r)) \cup w^{-1}((r, \infty))$. Thus, $(0, 2)$ is disconnected, a contradiction.

Define $\phi : v(E) \to \Re$ by $\phi(r) = u(x - \bar{e} + w^{-1}(r)\bar{e})$. Consider $y \in E$. Then, $v(y) \in v(E)$ and $\phi \circ v(y) = u(x - \bar{e} + w^{-1} \circ v(y)\bar{e})$. As $w((0, 2)) = v(E)$, there exists $r \in (0, 2)$ such that $v(y) = w(r) = v(x - \bar{e} + r\bar{e})$. Consequently,

$$\phi \circ v(y) = u(x - \bar{e} + w^{-1} \circ w(r)\bar{e}) = u(x - \bar{e} + r\bar{e}) = u(y)$$

where the last equality follows from ordinal congruence of u and v and the fact that $v(y) = v(x - \bar{e} + r\bar{e})$. Thus, f coincides with ϕ on $v(E)$. As u and w^{-1} are twice differentiable, so is ϕ, and therefore, f. ■

Lemma 5.4. *If*

(a) $(X, \langle .,. \rangle, \geq)$ *satisfies Assumption 5.1,*

(b) $O \subset X$ *is nonempty and open,*

(c) $x + X_+ \subset O$ *for every* $x \in O$*, and*

(d) $u \in \mathcal{U}$ *is differentiable,*

then $Du(x) > 0$ *for every* $x \in O$*.*

Proof. Let $\{b_i \mid i \in I\}$ be the Hilbert basis used to define $\geq$. Then, $\{b_i \mid i \in I\} \subset X_+$. Consider $x \in O$, $t \in (0,1)$ and $i \in I$. Using (c), $x + b_i \in O$. As $b_i > 0$ and $\geq$ is a partial order, $tb_i > 0$. Using (c), $x + tb_i \in O$. As $u \in \mathcal{U}$, $u(x+b_i) > u(x)$ and u is strictly concave. Therefore, $u(x + tb_i) - u(x) > t[u(x + b_i) - u(x)]$ for every $t \in (0,1)$. Therefore, (b), (d) and the fact that $\|b_i\| = 1$ imply

$$t[u(x+b_i) - u(x)] < u(x+tb_i) - u(x) = t\langle Du(x), b_i\rangle + t\|b_i\| r(t\|b_i\|) = t\langle Du(x), b_i\rangle + tr(t)$$

where $\lim_{t \downarrow 0} r(t) = 0$. Dividing by t and taking limits as $t \downarrow 0$, we have $\langle Du(x), b_i\rangle \geq u(x + b_i) - u(x) > 0$. As this holds for every $i \in I$, (a) implies $Du(x) \geq 0$. Clearly, $Du(x) \neq 0$. Thus, $Du(x) > 0$. ■

Theorem 5.5. *If*

(a) $(X, \langle .,. \rangle, \geq)$ *satisfies Assumption 5.1,*

(b) $O \subset X$ *is nonempty, convex and open in* X*,*

(c) $x + X_+ \subset O$ *for every* $x \in O$*, and*

(d) $u, v \in \mathcal{U}$ *are twice differentiable and ordinally congruent,*

then $u \succeq^3 v \Leftrightarrow u \succeq^4 v$*.*

Proof. As the conditions of Lemma 5.3 are satisfied, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$. Moreover, f is increasing and twice differentiable.

Using the chain rule, we have $Du(x) = Df(v(x)) \circ Dv(x)$ for every $x \in O$, i.e.,

$$\langle Du(x), y\rangle = Df(v(x))\langle Dv(x), y\rangle = \langle Df(v(x))Dv(x), y\rangle$$

for every $x \in O$ and $y \in X$. Thus, $Du(x) = Df(v(x))Dv(x)$ for every $x \in O$. By Lemma 5.4, $Du(x) > 0$ and $Dv(x) > 0$ for every $x \in O$. Thus, $Df(v(x)) > 0$ for every $x \in O$. By an analogous argument,

$$D[Df(v(x))] = D[Df \circ v(x)] = D^2 f(v(x))Dv(x)$$

Using the product formula to differentiate the identity $Du(x) = Df(v(x))Dv(x)$, we have

$$\begin{aligned} D^2u(x)y &= \langle D[Df(v(x))], y\rangle Dv(x) + Df(v(x))D^2v(x)y \\ &= \langle D^2f(v(x))Dv(x), y\rangle Dv(x) + Df(v(x))D^2v(x)y \\ &= D^2f(v(x))\langle Dv(x), y\rangle Dv(x) + Df(v(x))D^2v(x)y \end{aligned}$$

for every $y \in X$. By Lemma 5.4, $\|Du(x)\| > 0$ and $\|Dv(x)\| > 0$. Setting $y = Du(x)$, dividing both sides by $\|Du(x)\|^2$, and using the formula $Du(x) = Df(v(x))Dv(x)$, we have

$$\frac{D^2u(x)Du(x)}{\|Du(x)\|^2} = \frac{D^2f(v(x))Df(v(x))\|Dv(x)\|^2Dv(x)}{Df(v(x))^2\|Dv(x)\|^2} + \frac{Df(v(x))^2D^2v(x)Dv(x)}{Df(v(x))^2\|Dv(x)\|^2}$$

which simplifies to

$$\frac{D^2u(x)Du(x)}{\|Du(x)\|^2} = \frac{D^2f(v(x))Dv(x)}{Df(v(x))} + \frac{D^2v(x)Dv(x)}{\|Dv(x)\|^2}$$

Re-arranging and using the definition of the generalized Arrow-Pratt coefficient, we have

$$a_v(x) - a_u(x) = \frac{D^2f(v(x))Dv(x)}{Df(v(x))} \tag{5.6}$$

Suppose $u, v \in \mathcal{U}$ and $u \succeq^3 v$. By definition, there exists an increasing and concave function $f : v(O) \to \Re$ such that $u = f \circ v$. By Lemma 5.3, f is twice differentiable. By the above argument, $Df > 0$. By Lemma 5.4, $Dv(x) > 0$. As f is concave and twice differentiable, $D^2f \leq 0$. Thus, (5.6) implies that $a_v(x) - a_u(x) \leq 0$ for every $x \in O$, i.e., $u \succeq^4 v$.

Conversely, suppose $u \succeq^4 v$. Then, $a_v(x) - a_u(x) \leq 0$ for every $x \in O$. By Lemma 5.3, there exists a unique function $f : v(O) \to \Re$ such that $u = f \circ v$, and f is increasing and twice differentiable. As $Df > 0$ and $Dv > 0$, it follows from (5.6) that $D^2f(v(x)) \leq 0$ for every $x \in O$. Thus, f is concave on $v(O)$ and $u \succeq^3 v$. ■

Appendix

Proof of Lemma 2.1. By hypothesis, $\{x\}$ and $\operatorname{Int} C$ are nonempty and disjoint sets. Theorem V.2.1 in Dunford and Schwartz (1988) implies that $\operatorname{Int} C$ is convex and $C \subset \overline{\operatorname{Int} C}$. By the separating hyperplane theorem (Dunford and Schwartz 1988, Theorem V.2.8), there exists a non-zero continuous linear functional $p : L \to \Re$ such that $p(x) \geq p(y)$ for every $y \in \operatorname{Int} C$. Consider $y \in C - \operatorname{Int} C$. As $C \subset \overline{\operatorname{Int} C}$, we have $y \in \overline{\operatorname{Int} C}$; consequently, there exists a net $(y_i)_{i \in I} \subset \operatorname{Int} C$ converging to y. As $y_i \in \operatorname{Int} C$, we have $p(x) \geq p(y_i)$. As $\lim_i y_i = y$ and p is continuous, we have $p(x) \geq p(y)$, as required. ■

Proof of Lemma 2.2. (A) Given the product topology, $L \times \Re$ is a topological vector space. (a) and (b) imply that $H = \{(y, z) \in C \times \Re \mid z \leq u(y)\}$ is a convex set. $(x, u(x)) \in H - \operatorname{Int} H$ as $(x, u(x) + 1/n) \in (C \times \Re) - H$ for every $n \in \mathcal{N}$. As (b) implies that $(x, -1) \in \operatorname{Int} H$, we have $\operatorname{Int} H \neq \emptyset$.

Lemma 2.1 implies that there exists a non-zero continuous linear functional $p : L \times \Re \to \Re$ and $\gamma \in \Re$ such that $p(x, u(x)) = \gamma \geq p(y, z)$ for every $(y, z) \in H$. Linearity of p implies that $p(y, z) = p(y, 0) + p(0, z)$ for every $(y, z) \in L \times \Re$. Define $\alpha : L \to \Re$ by $\alpha(y) = p(y, 0)$ for $y \in L$, and let $\beta \in \Re$ be such that $p(0, z) = \beta z$ for $z \in \Re$. Clearly, α is a continuous linear functional on L. It follows from the definitions of α and β that

$$\alpha(x) + \beta u(x) = \gamma \geq \alpha(y) + \beta z$$

for every $(y, z) \in H$. In particular, $\gamma \geq \alpha(y) + \beta u(y)$ for every $y \in C$.

Suppose $\beta = 0$. As p is non-zero, α is non-zero and $\alpha(x) = \gamma \geq \alpha(y)$ for every $y \in C - \{x\}$. As α is non-zero, there exists $x' \in L$ such that $\alpha(x') > 0$. Consequently, $x' \neq 0$. By (c), x is an interior point of C. Therefore, x is an internal point of C (Dunford and Schwartz 1988, Theorem V.2.1). Then there exists $r > 0$ such that $x + rx' \in C - \{x\}$. Therefore, $\gamma \geq \alpha(x + rx') = \alpha(x) + r\alpha(x') > \alpha(x) = \gamma$, which is a contradiction.

Suppose $\beta < 0$. Then $(x, -n) \in H$ for every $n \in \mathcal{N}$. As $\beta < 0$, there exists $N \in \mathcal{N}$ such that $-\beta N > \gamma - \alpha(x)$. Thus, $\alpha(x) - \beta N > \gamma$, which contradicts $(x, -N) \in H$.

Thus, $\beta > 0$. Set $a = \gamma/\beta$ and $b = -\alpha/\beta$. The result follows.

(B) Suppose there exists $y \in C - \{x\}$ such that $a + b(y) = u(y)$. Then, $x/2 + y/2 \in C$ and $a + b(x/2 + y/2) = [a + b(x)]/2 + [a + b(y)]/2 = u(x)/2 + u(y)/2 < u(x/2 + y/2)$, which is a contradiction. ■

Remark A.1. *Consider a probability space $(\Omega, \mathcal{F}, P)$ and an $\mathcal{F}/\mathcal{B}(X)$ measurable function $x : \Omega \to X$. Set $O = x(\Omega)$. If the identity function $I : O \to X$ is Pettis integrable over O with respect to the probability space $(O, \mathcal{B}(O), P \circ x^{-1})$, then x is Pettis integrable over Ω with respect to $(\Omega, \mathcal{F}, P)$.*

Proof. Suppose $I : O \to X$ is Pettis integrable over O with respect to the probability space $(O, \mathcal{B}(O), P \circ x^{-1})$. Consider $x^* \in X^*$. By definition, $x^* \circ I : O \to \Re$ is $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable. Therefore, $x^* \circ x = x^* \circ I \circ x$ is $\mathcal{F}/\mathcal{B}(\Re)$ measurable. By definition, $\int_O P \circ x^{-1}(dz)\, x^*(z)$ exists. Consequently, $\int_\Omega P(d\omega)\, x^* \circ x(\omega) = \int_O P \circ x^{-1}(dz)\, x^*(z)$ exists. Finally, there exists $x_O \in X$ such that $x^*(x_O) = \int_O P \circ x^{-1}(dz)\, x^*(z) = \int_\Omega P(d\omega)\, x^* \circ x(\omega)$ for every $x^* \in X^*$. Set $x_\Omega = x_O$. It follows that x is Pettis integrable over Ω. ■

Remark A.2. *Let Y be a separable Banach space with closed unit sphere S. Let $X = Y^*$ and endow it with the Y topology. Let O be the closed unit sphere of X, i.e., $O = \{x \in X \mid \sup_{y \in S} |x(y)| \leq 1\}$. We show that the identity function $I : O \to X$ is Pettis integrable with respect to the probability space $(O, \mathcal{B}(O), \mu)$.*

Proof. It is easily confirmed that O is convex. By Alaoglu's theorem (Dunford and Schwartz 1988, Theorem V.4.2), O is compact. As Y is separable, O is metrizable (Dunford and Schwartz 1988, Theorem V.5.1). Consequently, O is separable.

Define the evaluation mapping $e : X \times Y \to \Re$ by $e(x, y) = x(y)$. Clearly, $\{e(., y) \mid y \in Y\}$ is a total space of linear functionals on X. As this is the space of linear functionals used to define the Y topology on X, $\{e(., y) \mid y \in Y\}$ is the set of continuous linear functionals on X (Dunford and Schwartz 1988, Theorem V.3.9). Consider $y \in Y$. As $e(., y)$ is continuous, $e(., y)$ is $\mathcal{B}(X)/\mathcal{B}(\Re)$ measurable. Also, I is $\mathcal{B}(O)/\mathcal{B}(X)$ measurable. Therefore, $e(., y) \circ I$ is $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable. Moreover, I and $e(., y)$ are continuous. As O is compact, $e(., y) \circ I$ is bounded. Therefore, $\int_O \mu(dz)\, e(., y) \circ I(z) = \int_O \mu(dz)\, e(z, y)$ exists.

Finally, we show that there exists $x \in O$ such that $e(x, .) = \int_O \mu(dz)\, e(z, .)$. Define $F : O \to \Re^Y$ by $F(x) = e(x, .)$, where $\Re^Y$ is given the product topology. Thus, our problem is to show that $F(x) = \int_O \mu(dz)\, e(z, .)$ for some $x \in O$.

As F is linear and continuous, $F(O)$ is convex and compact. As $\Re$ is Hausdorff, so is $\Re^Y$, and therefore $F(O)$ is closed. First, consider $\mu \in \Delta(O)$ with $\operatorname{supp} \mu < \infty$. It follows that $\int_O \mu(dz)\, e(z, .) = e(\int_O \mu(dz)\, z, .)$. As O is convex, $\int_O \mu(dz)\, z \in O$. Thus, $\int_O \mu(dz)\, e(z, .) = F(\int_O \mu(dz)\, z) \in F(O)$. Now consider an arbitrary $\mu \in \Delta(O)$. Endow $\Delta(O)$ with the weak* topology. As O is separable metric, $\{\mu \in \Delta(O) \mid \operatorname{supp} \mu < \infty\}$

is dense in $\Delta(O)$ (Parthasarathy 1967, Theorem II.6.3). It follows that there exists a net $(\mu_j)_{j\in J} \subset \Delta(O)$ such that $\operatorname{supp}\mu_j < \infty$ for each $j \in J$ and $\lim_j \mu_j = \mu$. By the definition of weak* convergence, $\int_O \mu(dz)\, e(z,.) = \lim_j \int_O \mu_j(dz)\, e(z,.)$. By the above argument, $\int_O \mu_j(dz)\, e(z,.) \in F(O)$ for every $j \in J$. As $F(O)$ is closed, $\int_O \mu(dz)\, e(z,.) = \lim_j \int_O \mu_j(dz)\, e(z,.) \in F(O)$. ■

Proof of Theorem 2.6. (A) Let $\underline{u} \in \Re$ be a lower bound for $u(O)$. Define $v : O \to \Re$ by $v(y) = u(y) - \underline{u}$. Clearly, v is concave, $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable and $v(y) \geq 0$ for every $y \in O$. Clearly, it suffices to show that (A) and (B) hold for v instead of u.

Let $\int_O \mu(dy)\, y = x$. Assumptions (a), (b) and (c) ensure that the assumptions of Lemma 2.2 are satisfied. Applying Lemma 2.2, there exists $a \in \Re$ and $x^* \in X^*$ such that $a + x^*(x) = v(x)$ and $a + x^*(y) \geq v(y)$ for $y \in O - \{x\}$. As v is $\mathcal{B}(O)/\mathcal{B}(\Re)$ measurable, and $a + x^*(.) \geq v(.) \geq 0$, and $\int_O \mu(dy)\, [a + x^*(y)] = a + \int_O \mu(dy)\, x^*(y)$ exists by (b), it follows that $\int_O \mu(dy)\, v(y)$ exists. It follows that

$$v(x) = a + x^*(x) = a + \int_O \mu(dy)\, x^*(y) = \int_O \mu(dy)\, [a + x^*(y)] \geq \int_O \mu(dy)\, v(y)$$

The second equality follows from (b).

(B) As μ is non-degenerate, there exists $B \in \mathcal{B}(O)$ such that $\mu(B) \in (0,1)$. Without loss of generality, let $x \in O - B$. By Lemma 2.2(B), $a + x^*(y) > u(y)$ for every $y \in B$. As $\mu(B) > 0$, we have

$$a\mu(B) + \int_B \mu(dy)\, x^*(y) = \int_B \mu(dy)\, [a + x^*(y)] > \int_B \mu(dy)\, v(y)$$

and

$$a\mu(O - B) + \int_{O-B} \mu(dy)\, x^*(y) = \int_{O-B} \mu(dy)\, [a + x^*(y)] \geq \int_{O-B} \mu(dy)\, v(y)$$

Adding these inequalities, we have $v(x) = a + x^*(x) = a + \int_O \mu(dy)\, x^*(y) > \int_O \mu(dy)\, v(y)$. The second equality follows from (b). ■

Notes

1. A distinct line of research (Grant, Kajii and Polak 1992a and 1992b, Spence and Zeckhauser 1972) studies the relationship between multivariate risk (lotteries over commodity bundles) and univariate risks (lotteries over wealth) when the two are linked by a budget constraint. It considers questions such as: In what ways can preferences over multivariate lotteries generate preferences over univariate lotteries? How do the properties of preferences in the multivariate setting map to the properties of preferences in the univariate setting? What can be inferred about preferences in the multivariate setting if we know the properties of preferences in the univariate setting?

2. For instance, the Wiener measure on the sample space of continuous real-valued functions defined on the non-negative real numbers results in the coordinate process becoming the Wiener process. This process is the building-block for geometric Brownian motion, a process routinely used to model price movements in the theory of asset-pricing. Itô and McKean (1965) is a classic reference for the mathematics of diffusions. For an introduction to the economic applications, see Duffie (1988).

3. See Pettis (1938) for the original statement of the theory. We are using a less restrictive version of the original definition.

4. For other general variants of this inequality, see for instance, Perlman (1974).

5. See Lang (1993) for details. Existence of a Hilbert basis follows from Corollary V.1.7 in Lang (1993).

References

Arrow, K. J. 1971. *Essays in the Theory of Risk Bearing.* Markham, Chicago.

Duffie, D. 1988. *Security Markets.* Academic Press, New York.

Duncan, G. T. 1977. A matrix measure of multivariate local risk aversion. *Econometrica* **45**, 895-904.

Dunford, N. and J. T. Schwartz. 1988. *Linear Operators, Part I: General Theory*, Wiley Interscience.

Grant, S., Kajii, A. and B. Polak. 1992a. Many good choice axioms: when can many-good lotteries be treated as money lotteries? *Journal of Economic Theory* **56**, 313-337.

Grant, S., Kajii, A. and B. Polak. 1992b. Many good risks: an interpretation of multivariate risk and risk aversion without the independence axiom. *Journal of Economic Theory* **56**, 338-351.

Itô, K. and H. P. McKean, Jr. 1965. *Diffusion Processes and Their Sample Paths.* Springer, Berlin.

Karni, E. 1979. On multivariate risk aversion. *Econometrica* **47**, 1391-1401.

Karni, E. 1989. Generalized expected utility analysis of multivariate risk aversion. *International Economic Review* **30**, 297-305.

Kihlstrom, R. M. and L. J. Mirman. 1974. Risk aversion with many commodities. *Journal of Economic Theory* **8**, 361-368.

Lang, S. 1993. *Real and Functional Analysis.* Springer-Verlag, New York.

Parthasarathy, K. R. 1967. *Probability Measures on Metric Spaces.* Academic Press, New York.

Perlman, M. D. 1974. Jensen's inequality for a convex vector-valued function on an infinite-dimensional space. *Journal of Multivariate Analysis* **4**, 52-65.

Peters, H. J. M. and P. P. Wakker. 1987. Convex functions on non-convex domains. *Economics Letters* **22**, 251-255.

Pettis, B. J. 1938. On integration in vector spaces. *Transactions of the American Mathematical Society* **44**, 277-304.

Pratt, J. W. 1964. Risk aversion in the small and in the large. *Econometrica* **32**, 122-136.

Spence, M. and R. Zeckhauser. 1972. The effect of the timing of consumption decisions and the resolution of lotteries on the choice of lotteries. *Econometrica* **40**, 401-403.

Yaari, M. E. 1969. Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory* **1**, 315-329.