Journal of Risk and Uncertainty, 1: 355–387 (1988)

Ordinal Independence in Nonlinear Utility Theory

JERRY R. GREEN
Harvard University

BRUNO JULLIEN
Harvard University

Key words: ordinal independence, nonlinear utility theory, preferences,

Abstract

Individual behavior under uncertainty is characterized using a new axiom, ordinal independence, which is a weakened form of the von Neumann–Morgenstern independence axiom. It states that if two distributions share a *tail* in common, then this tail can be modified without altering the individual's preference between these distributions. Preference is determined by the tail on which the distributions differ. This axiom implies an appealing and simple functional form for a numerical representation of preferences. It generalizes the form of *anticipated utility*, and it explains some well-known forms of behavior, such as the Friedman–Savage paradox, that anticipated utility cannot.

Recent research in the theory of individual decision making under uncertainty has developed in three directions. All of these are outgrowths of and reactions to the empirical refutation of expected utility theory that is widely acknowledged.

First, there are attempts to describe the decision-making process by examining aspects other than the probability distribution over the ultimate payoffs. Research in this direction uses variations in the description of the temporal resolution of the uncertainty or of the payoffs themselves as important ingredients that can affect the individual's choice.[1]

Second, there are models that look only at the probability distribution of payoffs, and impose normative axioms on choices between distributions.[2] The present article falls in this category.

The final group of models also looks only at the probability of various consequences. Here, however, an attempt is made to keep normative axioms to a minimum and to see how much flexibility can be maintained while at the same time explaining observed phenomena. The pioneering paper in this line of work is Machina (1982).

This research was supported by National Science Foundation grant number IRI-85-07291. The authors thank Mark Machina and Larry Epstein for helpful advice and comments. Jerry Green acknowledges support from the John Simon Guggenheim Memorial Foundation. Bruno Jullien gratefully acknowledges financial support from the Sloan Foundation.

This article introduces an axiom related to, but weaker than, the independence axiom of expected utility theory. We substitute this weaker axiom and obtain a numerical representation for preferences over distributions of payoffs. Naturally, this family of functionals includes the linear functionals of the expected utility family. It also includes the *anticipated utility* representation of Quiggin (1982), Segal (1984), and Yaari (1987). It is disjoint from the quasilinear family of preferences studied by Dekel (1986), Chew and MacCrimmon (1979), and Gul (1988) except that both contain expected utility as a special case.

There are two reasons for being interested in this axiom and the resulting representation. The first is that although anticipated utility can account for some of the observed violations of expected utility theory, it cannot account for all of them. In particular, the famous phenomenon of Friedman and Savage (1948), in which an individual's risk preferences seemingly depend on his status quo level of wealth, cannot be explained within the anticipated utility framework but can be explained by ours.

Second, and perhaps more importantly, we find the axiom itself intuitively appealing—more so than the necessarily stronger anticipated utility axiom. It is of interest to learn its implications. Part of the reason for the intuitive appeal of this axiom is that it bears some resemblance to psychological concepts of editing. One way in which the comparison between two decision problems can be simplified is to eliminate from consideration some values of payoffs on which the two payoff distributions coincide, and to determine preferences over these distributions by looking at the part of the payoff space on which the conditional distributions differ. Thus the common part of the space is *edited out*. We apply this logic when the common part of the space is a half-line: If F and G coincide either above or below some point, then the preference between F and G is determined by their restriction to the complementary half-line, on which they differ.

1. Ordinal independence

The basic axiom introduced in this article is called *ordinal independence*. It applies to spaces of payoffs that are naturally ordered, such as the real numbers. For simplicity, we assume that payoffs are in a bounded interval of real numbers, $X = [\underline{x}, x]$. Let the space of probability distributions over X be denoted D. The elements of D will be identified with their cumulative distribution functions and will be denoted $F, G, H, \ldots$

Preferences on D will be described by a binary relation $\succcurlyeq$. We assume that $\succcurlyeq$ is complete, transitive and continuous.

Complete weak order

The binary relation $\succcurlyeq$ on D is a complete weak order: For all $F, G \in D$, either $F \succcurlyeq G$ or $G \succcurlyeq F$. And if $F \succcurlyeq G$ and $G \succcurlyeq H$, then $F \succcurlyeq H$.

Continuity

The binary relation is continuous in the weak topology on D.

Monotonicity

If F (first-order) stochastically dominates G, then $F \succcurlyeq G$.

To these standard conditions we add the axiom of ordinal independence which can be stated as follows:

Ordinal independence

If $F \succcurlyeq G$ and $F(x) = G(x)$ for $x \geqslant \hat{x}$ (resp. $x < \hat{x}$) and $\bar{F}(x) = \bar{G}(x)$ for $x \geqslant \hat{x}$ (resp. $x < \hat{x}$), then $\bar{F} \succcurlyeq \bar{G}$.

This condition is a limited type of independence axiom. Let H and H' be distributions with support bounded above by $\hat{x}$ and let F and G be distributions with support bounded below by $\hat{x}$. Then, if the decision maker is indifferent between $\alpha F + (1 - \alpha)H$ and $\alpha G + (1 - \alpha)H$, he is also indifferent between $\alpha F + (1 - \alpha)H'$ and $\alpha G + (1 - \alpha)H'$.

These substitutions preserve indifference if, and in general only if, the support of the conditional distribution being substituted lies entirely above or entirely below the support of the distributions conditional on the complementary event.

2. Representation theorem

In this section we present the principal representation theorem for preferences satisfying the assumptions discussed in section 1. We will relate the conclusions of this theorem to both anticipated utility and to expected utility, which are successively special cases.

We begin by stating the theorem and its principal corollary. These state two equivalent closed-form expressions for a numerical index of the preference relation.

Theorem 1

If $\succcurlyeq$ satisfies complete weak order, continuity, monotonicity, and ordinal independence, then there exists a function $\bar{\phi}$: $X \times [0,1] \rightarrow \mathrm{R}$ such that $\bar{\phi}(0,p) \equiv 0$, $\bar{\phi}$ is nondecreasing in x, and a measure μ on $[0,1]$ such that

$$V(F) \equiv \int_0^1 \bar{\phi}(z(p),p)d\mu(p)$$

is a numerical representation of $\succcurlyeq$, where

$$z(p) \equiv \inf \{x \in X | F(x) \geqslant p\}.$$

Moreover, μ has a continuous distribution function and $\bar{\phi}$ is continuous on its domain.

Corollary

Under the hypotheses of theorem 1, an alternative representation is

$$V(F) = \int \bar{h}(x,F(x))\, d\nu(x)$$

where $\bar{h}(x,1) \equiv 0$, and $\bar{h}$ is nonincreasing in p. Moreover, ν has continuous distribution function and $\bar{h}$ is continuous on its domain.

The meaning of the representations can be seen as follows. We can think of $\bar{\phi}(z(p),p)$ as itself being an integral of some function (this will be called μ_2 in the proof; see appendix), but let us here denote it by $\zeta(x,p)$:

$$\bar{\phi}(z(p),p) = \int_0^{z(p)} \zeta(x,p)dx.$$

Thus the numerical utility indicator is an integral of $\zeta(x,p)$ over the epigraph[3] of the distribution function as shown in figure 1, and with respect to a measure μ on [0,1].

In anticipated utility theory the function $\zeta(x,p)$ becomes multiplicatively separable: $\zeta(x,p) = \zeta_1(x)\zeta_2(p)$. And, if we incorporate $\zeta_2(p)$ into the measure μ, we obtain an expression analogous to that of Segal (1984), Chew, Karni and Safra (1987), and Yaari (1987):

$$V(F) = \int u(x)d(g \cdot F(x)) = \int u(z(p))dg(p),$$

where

$$dg(p) = \zeta_2(p)d\mu(p),$$
$$u(x) = \zeta_1(x).$$

Expected utility theory is the further special case in which ζ_2 and μ, when so combined, produce a uniform distribution over [0,1]. Then, substituting $p = F(x)$, one can obtain the usual formula,

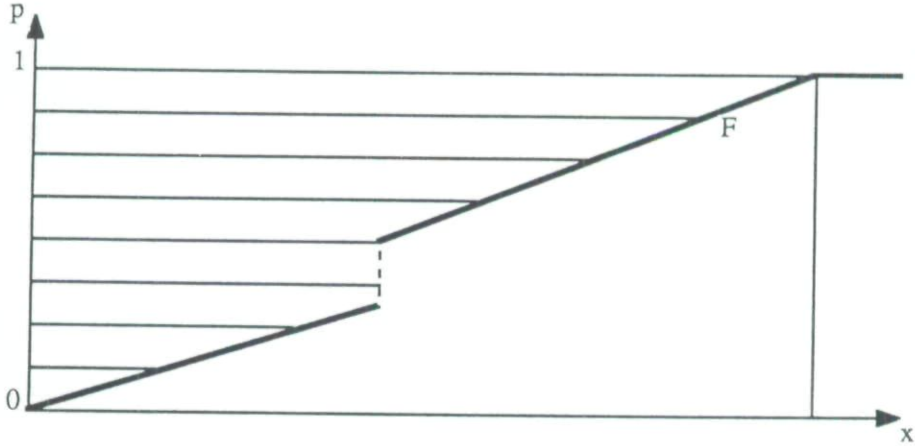

Fig. 1. The numerical utility indicator.

$$\int \zeta_1(z(p))dp = \int \zeta_1(x)dx.$$

The idea in the proof is to show that ordinal independence implies separability in the sense of Gorman (1968), reinterpreted in a certain fashion. The commodities that are separable from each other are the levels of payoff at various percentiles of the payoff distribution. Thus, a representation of preference will be an additively separable function of the payoffs at these levels. The relative importance of consumption at different percentiles of the payoff distribution is reflected in the measure μ. Intuitively, we are allowing the percentile of the distribution at which any given payoff x occurs to have two effects. It can affect the value of x, and independently various percentiles can be more or less important to preferences.

3. Representation with differentiability

The representation given in section 2, although simple, becomes substantially simpler under the following assumption.

Differentiability on Basic Distributions (DBD)

The certainty equivalent, $c(F_{\alpha,\hat{x}})$, of the distribution

$$\begin{aligned} F_\alpha &= \alpha \quad && \text{if } x < \hat{x}, \\ &= 1 && \text{if } x \geqslant \hat{x}, \end{aligned}$$

is an everywhere differentiable function of $\alpha \in [0,1]$, and of $\hat{x}$.

Under this condition, as we will now show, the measure μ on $[0,1]$ can be shown to be absolutely continuous with respect to Lebesgue measure, having density m. Hence, we can write the representation proven in theorem 1 as

$$V(F) = \int_0^1 \phi(z(p),p)dp,$$

where

$$\phi(z(p),p) = \bar{\phi}(z(p),p)m(p).$$

Before proving this assertion, let us consider the following examples in which, because of the failure of DBD, the measure μ will not have a density.

Example 1

For any F, let $A(F) = \{(z,p) \in X \times [0,1] | p \geqslant F(z)\}$. Let $V(F)$ be given by a measure $\bar{\mu}(A(F))$ as follows: The measure $\bar{\mu}$ will be a product measure with factors $\mu_2(dx)$ and $\mu(dp)$, where μ_2 can be an arbitrary positive measure, absolutely continuous with respect to Lebesgue measure, and μ has a point mass at $\bar{p}$.

The preference relation represented by V will fail to be continuous as one can see by considering a sequence of distributions F_α, where

$$\begin{aligned} F_\alpha(x) &= 0 \qquad x < x_1, \\ F_\alpha(x) &= \alpha \qquad x_1 \leqslant x < \bar{x}, \\ F_\alpha(x) &= 1 \qquad x = \bar{x}. \end{aligned}$$

Clearly, if $\alpha \rightarrow \bar{p}, F_k \rightarrow \bar{F}$. But $V(F_\alpha)$ will not converge to $V(\bar{F})$. Thus the preferences will not be continuous.

Example 2

Modify example 1 so that μ is a continuous, but not absolutely continuous, measure. For example, μ can be a probability measure concentrated on the Cantor ternary set (see Royden, 1963, ch. 2, problem 42). Then V and the preferences it represents will be continuous. However, preferences will not be differentiable at any $\bar{p}$ in the Cantor ternary set, since

$$\frac{dV(F_\alpha)}{d\alpha} = \mu_2(X)\frac{d\mu[\bar{p},1]}{d\bar{p}},$$

which clearly does not exist.

Thus, to avoid problems with differentiability, it will be necessary to avoid measures on $X \times [0,1]$ which are continuous but not absolutely continuous.

Theorem 2

Let $\succcurlyeq$ satisfy the hypothesis of section 2 and DBD. Then it can be represented by

$$V(F) = \int_0^1 \phi(z(p),p)dp = \int_{\underline{x}}^{\bar{x}} h(x,F(x))dx. \tag{1}$$

Proof: By theorem 1, $V(F) = \int_0^1 \bar{\phi}(z(p),p)d\mu(p)$.

Therefore it suffices to show that μ is absolutely continuous (with respect to Lebesgue measure). By DBD, the derivative of $V(F)$ at F_α will be $\phi(x,\alpha)d\mu(\alpha)/d\alpha$, which exists for all α only if μ is absolutely continuous. Q.E.D.

Throughout subsequent sections we use the representation (1).

It is useful at this point to relate the representation (1) to expected utility theory. If $\succcurlyeq$ obeys the independence axiom, and hence can be represented by a linear functional, $\int u(x)dF(x)$, then h takes the form

$$h(x,F(x)) = u'(x)(1 - F(x)). \tag{2}$$

Evaluating (1) by integrating (2) by parts, we have

$$\int_{\underline{x}}^{\bar{x}} h(x,F(x))dx = u(\underline{x}) + \int u(x)dF(x).$$

Thus, we can take $u(\underline{x}) = 0$ by normalization.

As we have pointed out, ϕ takes the form

$$\phi(z(p),p) = U(z(p)).$$

It is also fruitful to look at the form of the functional for particular distributions. Let us consider first a discrete distribution with support $(x_1, \ldots, x_n)$ and p_i = prob (x_i). Define the function:

$$\psi(x,q,p) = \frac{1}{p}\int_{q-p}^{q} \phi(x,s)ds, \qquad q \geqslant p \geqslant 0.$$

Then we have

$$V(F) = \sum_{i=1}^{n} \psi(x_i, F(x_i), p_i) p_i$$
$$= E_F \psi(x, F(x), \text{prob}\,(x)).$$

$V(F)$ is the expectation of some utility index, where the utility of a given payoff x depends on its probability of occurrence and the level of its cumulative.

Consider now a distribution F with a density function f. Then, using a simple change of variable,

$$F(F) = \int \phi(z(p), p) dp = \int \phi(x, F(x)) f(x) dx$$

or

$$V(F) = E_F \phi(x, F(x)).$$

Now the utility index depends solely on the level of the cumulative.

Notice that $\psi(x,q,0) = \phi(x, F(x))$, so that the interpretation given for discrete distributions extends to more general distributions.

Although none of the functionals (1) will be Frechet differentiable, except for those satisfying expected utility theory, the weaker hypothesis of Gateaux differentiability (see Chew, Karni, and Safra, 1987) can hold and is equivalent to the existence of $\partial h(x, F(x))/\partial F(x) \equiv h_2(x, F(x))$.

A functional V is said to be Gateaux differentiable if for each $F \in D$ there exists a linear functional $L(\cdot, F)$ such that

$$L(F' - F, F) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (V(\varepsilon F' + (1 - \varepsilon) F) - V(F)). \tag{3}$$

As Chew, Karni, and Safra have shown, the hypothesis of Frechet differentiability used by Machina can be weakened to Gateaux differentiability, and properties of the Gateaux derivative can usefully characterize global attitudes towards risk. In this sense the Gateaux derivative is a *local utility function*.

The Gateaux derivative of V at F is the function (3) evaluated for each $x \in X$ at the distribution F'_x which is a point mass concentrated at x. Under the hypothesis of theorem 2, and assuming Gateaux differentiability, the Gateaux derivative $U_F(x)$ of V at F is

$$U_F(x) = \int_{\underline{x}}^{x} - h_2(s, F(s)) ds = \int_{\underline{x}}^{x} \phi_1(s, F(s)) ds$$

and

$$L(F' - F,F) = \int_{\underline{x}}^{\bar{x}} U_F(x)\{dF'(x) - dF(x)\}.$$

We make extensive use of this formula in sections 6 and 7 below.

4. Risk aversion

Throughout the remaining part of this article, we will restrict ourselves to smooth preferences and assume the following:

Smoothness condition

$\phi_1 = -h_2$ is continuous and positive,

$\phi_{12} = -h_{22}$ exists everywhere and is continuous.

One of the main results in Machina's (1982) original paper was that one could use local utility functions to compare the degree of risk aversion of two individuals, extending the analysis of Arrow–Pratt. It is easily seen that the proof of the relevant theorem (his theorem 4) uses integrals along lines and requires only that the preference functional V be Gateaux differentiable (see for example Chew, Karni, and Safra, 1987). It follows that the result applies to our preference functional.

A first application is an easy characterization of risk aversion.

Theorem 3

An individual (ϕ,h) is risk-averse if and only if for all x,p

$$\phi_{11}(x,p) = -h_{21}(x,p) \leqslant 0,$$

$$\phi_{12}(x,p) = -h_{22}(x,p) \leqslant 0.$$

Proof: Take $\phi^*(x,p) = x$, the risk-neutral preferences in theorem 4. An individual is risk-averse if ϕ is concave in x and h convex in p. We can see how the ϕ-form and the h-form are dual one to another. For any cumulative F we can interpret its inverse cumulative z_F as a cumulative (normalize $\underline{x}$ to 0 and $\bar{x}$ to 1). If a cumulative G is a mean-preserving spread of a cumulative F, then z_F is a mean-preserving spread of z_G. Therefore the concavity of ϕ in x (which corresponds to the concavity of the utility function in expected utility theory) transforms into the convexity of h in p.

The characterization of risk aversion is just a special case of a more general result on comparative risk aversion.

Definition

A distribution F is said to differ from a distribution G by a simple compensated spread if $V(F) = V(G)$ and if there exists x^* such that $F(x) \geqslant G(x)$ for $x \leqslant x^*$ and $F(x) \leqslant G(x)$ for $x \geqslant x^*$.

We will say that an individual A is more risk-averse than an individual B if, whenever F differs from G by a simple compensated spread from the point of view of B, then A prefers G to F.

Machina proves the equivalence between several definitions of increasing risk aversion and the fact that for all F, the utility function of A at F is a concave transform of the utility function of B at F. We will interprete this result directly using the functions ϕ and h.

Theorem 4

An individual (ϕ,h) is more risk averse than an individual (ϕ^*,h^*) if and only if for all x,p

$$-\frac{\phi_{11}(x,p)}{\phi_1(x,p)} \geqslant -\frac{\phi^*_{11}(x,p)}{\phi^*_1(x,p)} \qquad \text{and } -\frac{\phi_{12}(x,p)}{\phi_1(x,p)} \geqslant -\frac{\phi^*_{12}(x,p)}{\phi^*_1(x,p)}.$$

Proof: See appendix.

The extended Arrow–Pratt measure of absolute risk aversion is now composed of two points: $-\phi_{11}/\phi_1$ and $-\phi_{12}/\phi_1$. It reduces to the usual measure when preferences are linear since then $\phi_1(x,p) = u'(x)$ and the measure is $(-u''(x)/u'(x),0)$.

4.1. Interpretation in terms of risk premium

4.1.1. The wealth premium. Consider an individual (ϕ,h) who is given the choice between the lotteries over final levels of wealth:

$$A \begin{cases} \underline{x} & \text{with probability } p_0 \\ x_0 - \varepsilon & \text{with probability } p/2 \\ x_0 + \varepsilon & \text{with probability } p/2 \\ \bar{x} & \text{with probability } 1 - p_0 - p \end{cases}$$

$$B \begin{cases} \underline{x} & \text{with probability } p_0 \\ x_0 - \pi & \text{with probability } p \\ & \\ \bar{x} & \text{with probability } 1 - p - p_0 \end{cases}$$

The level of π that makes the individual indifferent between the two lotteries is the wealth premium that the individual is willing to pay to avoid the risk ε. A straightforward calculus shows that the limit $\tilde{\pi}$ of the premium π when p goes to zero is given by

$$\tfrac{1}{2}\{\phi(x_0 - \varepsilon, p_0) + \phi(x_0 + \varepsilon, p_0)\} = \phi(x_0 - \tilde{\pi}, p_0).$$

So $\tilde{\pi}$ is the risk premium associated to the lottery $(\varepsilon, -\varepsilon, \frac{1}{2}, \frac{1}{2})$ when the initial wealth is x_0 and the individual maximizes an expected utility with utility function $\phi(x, p_0)$. As is well known, it is approximated by

$$\tilde{\pi} \simeq -\frac{1}{2}\frac{\phi_{11}(x_0, p_0)}{\phi_1(x_0, p_0)}\left(\frac{\varepsilon^2}{2}\right).$$

When p is small, ϕ is almost linear in p around p_0. Therefore everything is similar to the case of expected utility.

4.1.2. The probability premium. Consider an individual (ϕ, h) who is now given the choice between the following lotteries:

$$A \begin{cases} x & \text{with probability } p_0 - \varepsilon \\ x_0 + \dfrac{x}{2} & \text{with probability } 2\varepsilon \\ x_0 + x & \text{with probability } 1 - p_0 - \varepsilon \end{cases}$$

$$B \begin{cases} x_0 & \text{with probability } p_0 - q \\ x_0 + x & \text{with probability } 1 - p_0 + q \end{cases}$$

The level of q that makes the individual indifferent between the two lotteries is the probability premium that the individual is willing to accept before giving up the extra gamble ε (notice that now, when $q = 0$, A is less risky than B). This example is dual to the previous one. If we reinterpret the inverse cumulative functions as cumulative functions, the two problems are the same except that now h replaces ϕ. The limit $\tilde{q}$ of the premium q when x goes to zero is given by

$$\frac{h(x_0, p_0 - \varepsilon) + h(x_0, p_0 + \varepsilon)}{2} = h(x_0, p_0 - \tilde{q}).$$

So when ε is small, $\tilde{q}$ can be approximated by

$$\tilde{q} \simeq -\frac{1}{2}\frac{h_{22}(x_0,p_0)}{h_2(x_0,p_0)}\frac{\varepsilon^2}{2} = -\frac{1}{2}\frac{\phi_{12}(x_0,p_0)}{\phi_1(x_0,p_0)}\frac{\varepsilon^2}{2}.$$

Notice that under expected utility, $\tilde{q}$ is exactly zero. When x is close to 0, the utility function is almost linear and the agent is risk-neutral. With nonlinear preferences, the marginal utility of wealth becomes almost constant in the relevant range, but there is another dimension to risk aversion. Preferences can be represented by Yaari's dual preferences.

5. The Allais paradox and the common-ratio effect

The Allais paradox and other related observations have been extensively examined in the existing literature. We refer the reader to Kahneman–Tversky (1979) and MacCrimmon–Larsson (1979) for detailed exposition and discussion. We will restrict most of our discussion to the case of three outcome distributions $(x_1,x_2,x_3,p_1,p_2,p_3)$ with $x_1 < x_2 < x_3$. There is a very convenient graphical representation of such a distribution introduced by Machina: for given outcomes $x_1 < x_2 < x_3$, we can represent a distribution in the plane by using p_1 and p_3, the probabilities of the low and the large outcomes. Diagram 1 illustrates the common-ratio effect.

The sure outcome C is preferred to the lottery D. A and B are obtained from C and D by mixing the bad outcome with probability q. In many observations B is preferred to A, contradicting the prediction of expected utility theory. Notice that

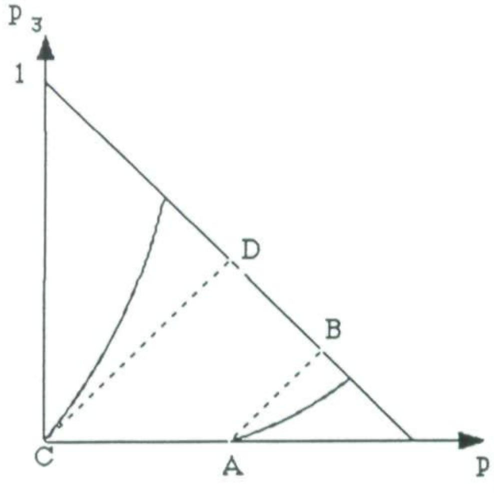

Diagram 1. The common-ratio effect.

under expected utility the isopreference curves are parallel straight lines. Similarly, diagram 2 illustrates the Allais paradox.

Now A and B are obtained from C and D by transferring a probability q from the medium outcome x_2 to the low outcome x_1. As before, an expected utility maximizer who prefers C to D must prefer A to B.

Machina (1982) proposes to generalize both paradoxes under the following behavioral assumption:

Generalized Common Ratio Effect (GCRE): Let $F_A, F_B, F_C, F_D \in D$ be such that F_C and F_D respectively stochastically dominate F_A and F_B, and $F_D - F_C = \lambda(F_B - F_A)$ for some $\lambda > 0$. Then, if F_B differs from F_A by a simple compensated spread, $V(F_D) \leqslant V(F_C)$. Similarly, if F_D differs from F_C by a simple compensated spread, then $V(F_B) \geqslant V(F_A)$.

When the distributions have a support composed only of three outcomes $x_1 < x_2 < x_3$, the GCRE has the interpretation illustrated in diagram 3.

A distribution stochastically dominates another distribution if it lies above and on the left. If we choose A, B, C as shown and $A \sim B$, then C must be preferred to any distribution on the segment $[E,F]$. By choosing B close to A, we see that it implies that the slope of the isopreference curve at C be greater than the slope of the isopreference curve at A. The slope of the isopreference curve at some point (p_1,p_3) is given by

$$\frac{dp_3}{dp_1}|V = \frac{\phi(x_2,p_1) - \phi(x_1,p_1)}{\phi(x_3,1-p_3) - \phi(x_2,1-p_3)}.$$

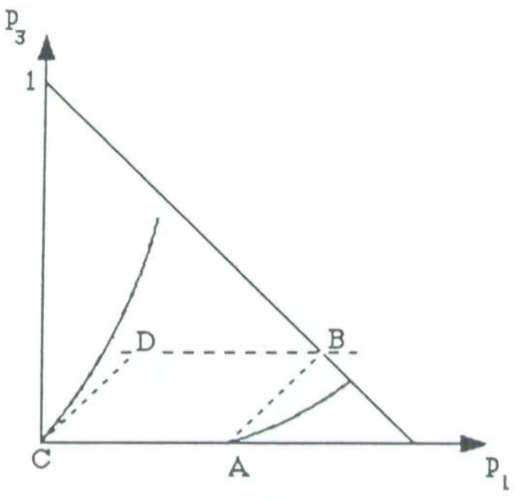

Diagram 2. The Allais paradox.

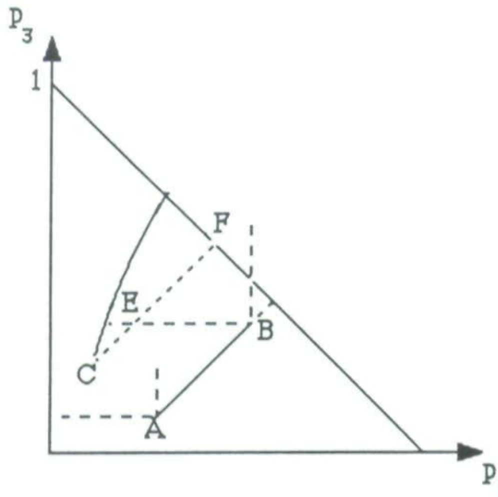

Diagram 3. Three-outcome generalized common-ratio effect.

The GCRE implies that this ratio be nonincreasing with p_1 and nondecreasing with p_3.

Theorem 5

An individual (ϕ,h) verifies the GCRE if and only if she is an expected utility maximizer.

Proof: Choose any $x_1 < x_2 < x_3 < x_4$, then $\phi(x_2,p_1) - \phi(x_1,p_1)/(\phi(x_3,1-p) - \phi(x_2,1-p))$ is nondecreasing with p and $\phi(x_3,p) - \phi(x_2,p)/(\phi(x_4,1-p_3) - \phi(x_3,1-p_3))$ is nonincreasing with p. This is only possible if $\phi(x_3,p) - \phi(x_2,p)$ is independent of p, or $\phi_1(x,p) = u'(x)$. Q.E.D.

The theorem generalizes the same result obtained by Segal (1984) for anticipated utility; it shows that this is not linked to the multiplicative separability. Segal then proposed a weaker behavioral assumption inspired by the Allais paradox:

Generalized Allais Paradox (GAP): Let F_A, F_B, F_C, F_D be such that F_C and F_D respectively stochastically dominate F_A and F_B, and $F_D - F_C = F_B - F_A$. Let $V(F_A) = V(F_B)$ and x^* be such that for $x \geqslant x^*$, $F_A(x) \geqslant F_B(x)$ and for $x \leqslant x^*$, $F_A(x) \leqslant F_B(x)$. If for $x \geqslant x^*$, $F_C(x) = F_A(x)$, then $V(F_C) \geqslant V(F_D)$.

In this version C and D are obtained by shifting in the same amount probabilities from low to medium outcomes, as in the Allais paradox. In the case of

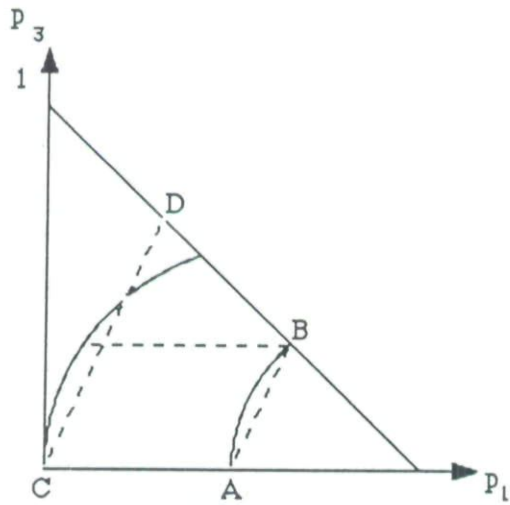

Diagram 4. Isopreference curves for $\phi_{12} < 0$.

three outcomes, it means that A and C are on the same horizontal line, and so are B and D.

Theorem 6

The preferences (ϕ,h) verify the GAP if and only if for all x,p

$$\phi_{12}(x,p) = -h_{22}(x,p) \leqslant 0.$$

Proof: See appendix.

The result is interesting because it shows that the type of behavior characterizing the Allais paradox is not only compatible with but is implied by risk aversion.[4] It suggests that the best candidates to explain some risk-loving behavior while staying consistent with experimental observations are preferences with h convex in p but ϕ not concave in x.

Remark: The GAP does not imply the common ratio effect. The reason is that if $\phi_{12} < 0$, the isopreference curves are concave so that the GAP does not prevent the sitaution depicted in diagram 4.

However it is clear that they are not contradictory.

6. The Friedman–Savage hypothesis and the boundedness of preferences

In their seminal article, Friedman–Savage (1948) pointed out that many individuals were simultaneously purchasing lottery tickets and insurance. They proposed a von Neumann–Morgenstein utility function concave for low outcomes (and hence risk-averse) and convex for large outcomes (risk-loving). There were several difficulties with this utility representation. First, the utility function was unbounded, therefore inducing unbounded preferences and subject to the *St. Petersburg Paradox*. In addition, the willingness to pay for $1/k$ chance of winning $\$k\varepsilon$ was increasing with k. This led the authors to add a terminal concave part at very large outcomes. Finally, the utility representation could not explain why people purchase lottery tickets and insurance regardless of their initial wealth. These points are discussed in great detail by Machina (1982). As was pointed out by Machina, when the preference functional is nonlinear in the distributions, preferences may be bounded even though the local utility functions are unbounded. This could explain at the same time the observed gambling behaviors and their relative invariance to the initial wealth since the inflection point of the local utility function would depend on the initial wealth. Machina's analysis relies on Frechet differentiability and therefore cannot be applied directly to our preference relation. The reason is that when V is Frechet differentiable, the derivative of V at some point characterizes the local behavior of V in a precise sense; however, this is not true if V is Gateaux differentiable. In the latter case, we cannot uniformly approximate V by its derivative in some neighborhood.

Let us first mention that the boundedness of preferences guarantees that the agent will not purchase a lottery ticket with too high a prize.

Theorem 7

Suppose that V is bounded; then for all $w > 0$ and $\varepsilon > 0$, the sure outcome w is preferred to the lottery $(w - \varepsilon, w + (1 - p/p)\varepsilon, 1 - p, p)$ for p small enough.

Proof: See appendix.

The theorem tells us that a decision maker will not purchase a fair gamble offering $k\varepsilon$ with probability $1/k$ if k is too large.

We now turn to the problem of the unbounded utility functions. It appears quickly that the nonseparability of the function ϕ and h is crucial.

Theorem 8

If ϕ is separable, i.e. $\phi(x,p) = u(x)g(p)$, and $V(\cdot)$ is bounded, then the local utility function at any distribution F is bounded.

Proof: Let G_W be the cumulative distribution of a point mass at W.

$$V(G_W) = u(w)\int_0^1 g(p)dp, \qquad \text{so } u(\cdot) \text{ is bounded;}$$

$$U_F(x) = \int_0^x u'(s)g(F(s))ds < \sup_p g \sup_x u. \qquad \text{Q.E.D.}$$

When ϕ is not separable, we can easily construct preferences that are bounded with unbounded utility functions, using the following lemma.

Lemma (8.1)

If the function $p \in]0,1[\rightarrow \sup_x \phi(x,p)$ is L^1, then $V(\cdot)$ is bounded.

Proof: $V(F) = \int_0^1 \phi(z(p),p)dp < \int_0^1 \sup_x \{\phi(x,p)\}dp$.

So to exhibit some bounded preference functional V with an unbounded local utility function U_F at all distribution F, we choose ϕ verifying the condition of lemma (8.1) and such that $\phi(x,1)$ is unbounded. If we choose $\phi(x,0)$ to be concave and $\phi(x,1)$ to be convex, the utility functions will have the desired concave–convex shape (see figures 2 and 3).

Note that at the same time we solve the problem of the relative invariance of gambling behavior to initial wealth, since the inflection point (or region) of the utility function will change with the initial distribution. An appealing property is that for a nonrandom wealth w, the inflection point is exactly at w.

As we pointed out above, the derivative of V at F does not characterize completely the local behavior of V. When talking about lottery or insurance, we do not consider unidirectional pertubations of the initial wealth, so that a decision maker may not want to purchase a lottery ticket or an insurance contract even though the local utility function has the right shape. The strategy we adopted is to characterize the conditions under which he or she would purchase a lottery involving a fixed gain or an insurance against a fixed loss if the probability of the event considered is small enough, in the spirit of the local analysis.

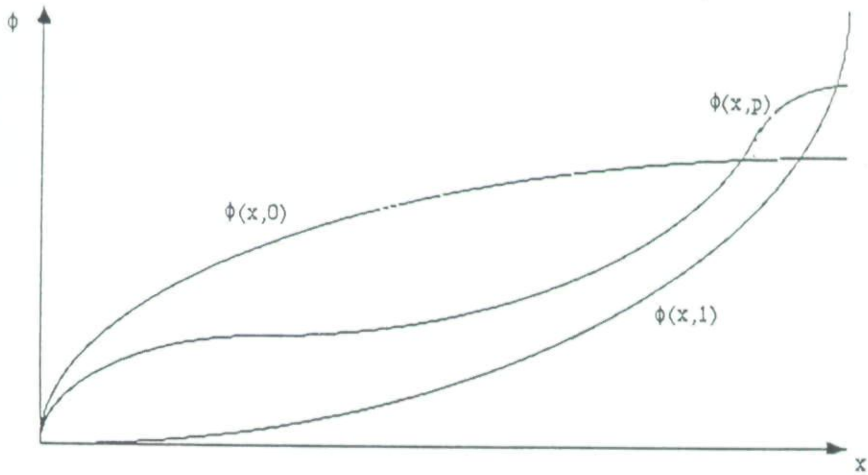

Fig. 2. Concave-convex utility function when $\phi(x,0)$ is concave and $\phi(x,1)$ is convex.

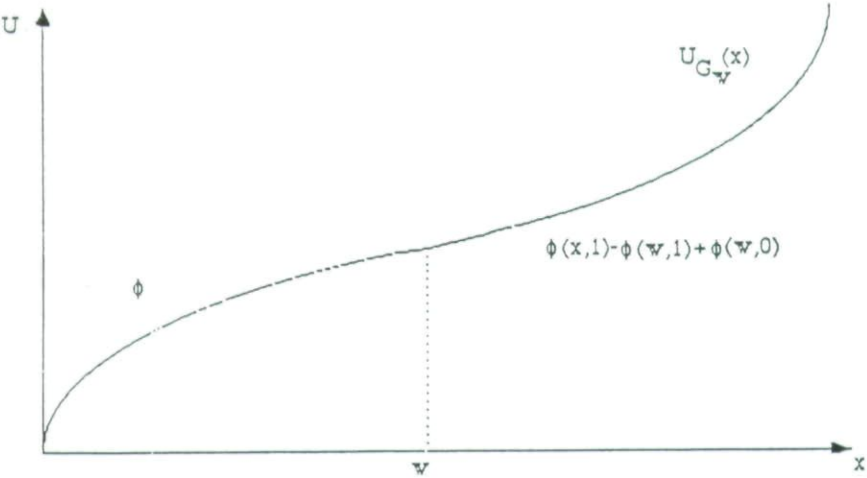

Fig. 3. Concave-convex utility function solving the problem of the relative invariance of gambling behavior to initial wealth.

Lemma (8.2)

An individual with initial distribution F will accept the lottery (ε, $(-p/1-p)\varepsilon$, p, $1-p$) when p is small if

$$E_F\left\{\frac{U_F(x+\varepsilon)-U_F(x)}{\varepsilon}\right\} > \int_0^1 \phi_1(z(p),p)dp \tag{4}$$

and only if the weak inequality holds.

Proof: See appendix.

For the case of insurance, the initial distribution must include the loss but the result is similar.

Lemma (8.3)

An individual with an initial distribution $G(x) = pF(x+\varepsilon) + (1-p)F(x)$, i.e., wealth w with distribution $F(w)$ plus an additional p chance of losing ε independently of w, will insure against the loss ε when p is small enough if

$$E_F\left\{\frac{U_F(x)-U_F(x-\varepsilon)}{\varepsilon}\right\} > \int_0^1 \phi_1(z(p),p)dp, \tag{5}$$

and only if the weak inequality holds.

The conditions when V is Frechet differentiable are the same except that $\int_0^1 \phi_1(z(p),p)dp$ is to be replaced by $E_F U'_F(x)$. If we assume that $\phi_{12}(x,p) \leqslant 0$, then $\int_0^1 \phi_1(z(p),p)dp \geqslant \int_0^1 \phi_1(x,F(x))\, dF(x) = E_F U'_F(x)$, so that the conditions (4) and (5) are stronger. However, when the distribution F has no point-mass, $\int_0^1 \phi_1(z(p),p)dp = \int_0^1 \phi_1(x,F(x))\, dF(x)$. Therefore it is natural to impose as a first requirement that the local utility function at a smooth distribution have a concave–convex shape. If we want preferences to be bounded, this rules out separable forms $\phi(x,p) = u(x)g(p)$.

There is a more fundamental reason to exclude separable forms. As was done by Friedman and Savage, a separable form could be reconciled with bounded preferences by adding a terminal concave section to the utility functions. But a separable form cannot explain the invariance of gambling behavior to initial wealth. The local utility function for a fixed initial wealth w when $\phi(x,p) = u(x)g(p)$ is given by

$$\begin{aligned} U_{G_w}(x) &= u(x)g(0) && \text{if } x \leqslant w, \\ &= u(x)g(1) - u(w)\{g(1) - g(0)\} && \text{if } x \geqslant w. \end{aligned}$$

We see that the shape of the utility function at some level x is independent of w. It is impossible that the utility be concave–convex with an inflexion point close to w for all G_w, since the inflexion point must be independent of w.

Remark: If $\phi_{12}(w,p) \leqslant 0$ and $\phi(x,0)$ is concave, when the initial distribution is G_w, (5) is verified for all $\varepsilon \leqslant w$, while (4) is not verified for ε small. This is consistent with the existence of lotteries with substantial prizes only.

Using lemmas (8.1) and (8.2), we see that to build an example of bounded preferences compatible with the simultaneous purchase of lottery tickets and insurance, one can do the construction illustrated in figures 2 and 3 and choose $\phi(x,1)$ such that $\lim_{x\to+\infty} \phi_1(x,1) = +\infty$. Then the individual will purchase a lottery (ε,p) when ε is large and p is small enough. This does not contradict theorem 7 because the probability p has to be chosen after ε. In theorem 7 we fix the premium and increase the prize, while now we fix the prize and decrease the premium. The next section will give specific examples.

7. Integrated solutions

We want now to determine whether the theory enables us to reconcile the behaviors discussed in section 5 and 6 with the same preference functional. More precisely, can we find ϕ and h such that preferences are bounded, the GAP is verified, and, at least for all G_w, the relations (4) and (5) are verified for some $\varepsilon > 0$? It turns out that if we do not restrict the levels of wealth considered, the search is hopeless. We show below that the GAP and the purchase of some lottery tickets at all levels of wealth (as we defined it) are incompatible with bounded preferences.

Theorem 9

Suppose that for all x, p, $\phi_{12}(x,p) < 0$ and that for all distributions G_w there exists $\varepsilon > 0$ such that relation (4) is verified; then preferences are unbounded.

Proof: See appendix.

The intuition behind the results is that if $\phi_{12}(x,p) < 0$, the slope of $\phi(.,p)$ decreases with p so that when $\phi(.,1)$ is unbounded, so will be $\phi(.,p)$. But if $\phi(.,p)$ is unbounded for all p, preferences must be unbounded.

The result is not so disturbing because wishing to reconcile everything at all levels of wealth appears a little excessive. After all, we are talking about the initial wealth of the individual, and initial wealth is bounded. We will show in the following examples that we can still go very far in the search for an integrated solution. The problem comes from the behavior when the initial wealth is very large. If we assume that either the GAP or relation (4) is verified only for bounded levels of initial wealth, the other requirements can be verified for all level of wealth.

Example 1

Suppose that $\phi(x,p) = x + (1 - 2\sqrt{p})\, x/(1 + x)$. Then

$$\phi_1(x,p) = 1 + \frac{(1 - 2\sqrt{p})}{(1 + x)^2} > 0,$$

$$\phi_{12}(x,p) = -\frac{1}{\sqrt{p}(1 + x)^2} < 0.$$

$\phi(.,p)$ is concave for $p < 1/4$, convex for $p > 1/4$ (see figure 4).

For $F = G_w$,

$$\int_0^1 \phi_1(w,p)dp = 1 - \frac{1}{3}\frac{1}{(1 + w)^2},$$

while

$$\frac{U_{G_w}(w + \varepsilon) - U_{G_w}(w)}{\varepsilon} = 1 + \frac{1}{\varepsilon}\left[\frac{w}{1 + w} - \frac{w + \varepsilon}{1 + w + \varepsilon}\right],$$

$$\frac{U_{G_w}(w - \varepsilon) - U_{G_w}(w)}{\varepsilon} = 1 + \frac{1}{\varepsilon}\left[\frac{w}{1 + w} - \frac{w - \varepsilon}{1 + w - \varepsilon}\right].$$

Therefore relation (4) is verified for $\varepsilon > (1 + w)2$, while relation (5) is verified for all $\varepsilon < w$. One appealing aspect is that the purchase of insurance appears to be more general than the purchase of lottery tickets. However, preferences are unbounded:

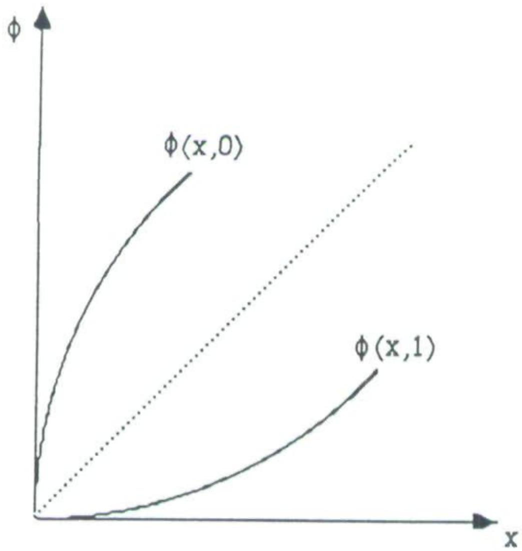

Fig. 4. Example 1.

$$V(G_w) = \int_0^1 \left\{ w + (1 - 2\sqrt{p}) \frac{w}{1+w} \right\} dp = w - \frac{1}{3} \frac{w}{1+w} \xrightarrow[w \to +\infty]{} +\infty.$$

Example 2 (see Figure 5)

Choose k large enough:

$$\begin{aligned} \phi(x,p) &= x + (1 - 2\sqrt{p}) \frac{x}{1+x} && \text{if } x \leqslant k, \\ &= \phi(k,p) + \phi_1(k,p) \frac{x-k}{1+x-k} && \text{if } x \geqslant k. \end{aligned}$$

Now the preferences are bounded, since

$$\phi(x,p) < \phi(k,p) + \phi_1(k,p) < k + \frac{k}{1+k} + 1 + \frac{1}{(1+k)^2}.$$

ϕ_1 is defined everywhere and continuous:

$$\phi_1(x,p) = 1 + \frac{1 - 2\sqrt{p}}{(1+x)^2} \qquad \text{if } x \leqslant k,$$

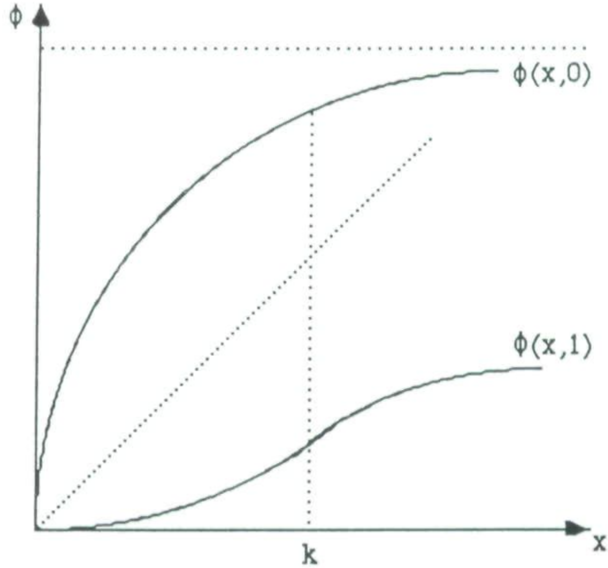

Fig. 5. Example 2.

$$\phi_1(x,p) = \phi_1(k,p)\,\frac{1}{(1 + x - k)^2} \qquad \text{if } x \geqslant k.$$

$\phi_{12}(x,p)$ is defined everywhere and $\phi_{12}(x,p) < 0$:

$$\phi_{12}(x,p) = -\,\frac{1}{\sqrt{p}(1 + x)^2} \qquad \text{if } x \leqslant k,$$

$$= -\,\frac{1}{\sqrt{p}(1 + k)^2(1 + x - k)^2} \qquad \text{if } x \geqslant k.$$

So the GAP is verified. Provided that $w < \frac{1}{3}(k - 2)$, an individual with initial wealth w will purchase some lottery tickets and insure all small risks (relation (4) is verified for some ε and relation (5) for all $\varepsilon < w$).

Example 3 (see figure 6)

Define $x(p) = k/\sqrt{p}(1 - p)$, k large.

$$\phi(x,p) = x + (1 - 2\sqrt{p})\,\frac{x}{1 + x} \qquad \text{if } x \leqslant x(p),$$

$$= \phi(x(p),p) + \phi_1(x(p),p)\,\frac{x - x(p)}{1 + x - x(p)} \qquad \text{if } x \geqslant x(p).$$

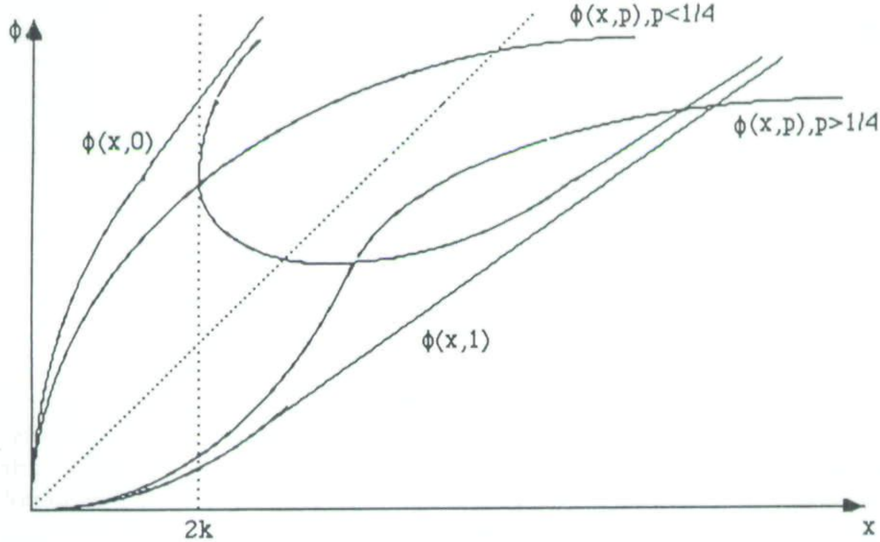

Fig. 6. Example 3.

As in example 2, ϕ_1 is defined, continuous and positive everywhere. As long as $x < x(p)$, $\phi_{12}(x,p)$ is defined and negative so that the GAP is verified at least for distributions with upper bound less than $2k$.

For a fixed wealth w, the utility is

$$U_{G_w}(x) = x + \frac{x}{1+x} \qquad \text{if } x \leqslant w,$$

$$= x - \frac{x}{1+x} + 2\frac{w}{1+w} \qquad \text{if } x \geqslant w.$$

The utility function at G_w is the same as in example 1 (only $\phi(.,0)$ and $\phi(.,1)$ matter). Notice that[5] $\phi_1(x,p) \leqslant 1 + (1 - 2\sqrt{p})/(1 + x)$. Therefore, the results of example 1 hold, and an individual with initial wealth w will purchase some lottery tickets and insure all small risks.

Notes

1. Kahneman and Tversky (1979, 1984), Loomes and Sugden (1982), Segal (1987).
2. Chew and MacCrimmon (1979), Chew (1983), Segal (1984), Dekel (1986), and others.
3. The idea of using a measure on the epigraph of F as a representation of preferences is due to Segal (1984). See also Chew and Epstein (1987).
4. In the case of a multiplicatively separable form $\phi_1(x,p) = u'(x)f'(1-p)$, the condition reduces to f convex as found by Segal (1984).

5. For all $\alpha \in [-1,1], x > k, 1 + \dfrac{\alpha}{(1+x)^2} > \left(1 + \dfrac{\alpha}{(1+k)^2}\right) \dfrac{1}{(1+x-k)^2}$.

Appendix

Proof of Theorem 1

We begin by considering the set of all distributions with equally unlikely outcomes, D^E, and, for each n, the subset $D_n^E \subset D^E$ with n outcomes each of which has a probability that is a multiple of $1/n$.

For $F \in D_n^E$, let us list the mass-points of F in nondecreasing order as $x_1^F, \ldots, x_n^F$. By the axioms of ordering and continuity on D_n^E, we can represent preferences by a numerical indicator $V_n^E(F) \equiv U_n(x_1, \ldots, x_n)$. The domain of U_n is the n-fold Cartesian product of X, subject to the constraints that $x_1 \leqslant \ldots \leqslant x_n$. Let us denote this space X^n. The subsets of components $\{i|1 \leqslant i \leqslant j\}$ and $\{i|j \leqslant i \leqslant n\}$ are separable in Gorman's sense, by virtue of the ordinal independence assumption.

Lemma

If $\succcurlyeq$ satisfies ordering, continuity, and ordinal independence on D_n^E, then there exist u_i^n, $i = 1, \ldots, n$, such that

$$\sum_{i=1}^{n} u_i^n(x_i^F) \tag{A.1}$$

is a numerical representation of $\succcurlyeq$. Moreover, u_i^n is continuous and nondecreasing.

Proof of lemma: int X^n can be written as the union of open rectangles $\{S_k\}_{k=1,\ldots}$ where

$$S_k \subseteq \mathbb{R}^n$$

and where, for any k, $S_k \cap (\cup_{j=1}^{k-1} S_j) \neq \phi$.

Apply Gorman's theorem to S_1, obtaining a representation

$$U_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} u_i(x_i), \qquad x \in S_1,$$

where u_i are continuous and nondecreasing, by virtue of the continuity and monotonicity axioms.

Now apply Gorman's theorem to S_2, obtaining

$$U_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} \tilde{u}_i(x_i), \qquad x \in S_2.$$

The functions u_i are unique up to common affine transformations. Therefore there will be a unique set of functions $\tilde{u}_i$ that agree with the u_i on their common domain $S_1 \cap S_2$. With only a slight abuse of notation we can use u_i to denote these functions throughout the domain $S_1 \cup S_2$. Continuing this procedure, and extending continuously to the boundary of X^n, we obtain (A.1). Q.E.D.

Define

$$\psi^n: X \times \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\} \to \mathrm{R}$$

by

$$\psi^n\left(x, \frac{k}{n}\right) = \sum_{i=1}^{k} u_i^n(x) \tag{A.2}$$

and

$$\psi^n(x, 0) = 0.$$

Since

$$\psi^n\left(x_i^F, \frac{i}{n}\right) - \psi^n\left(x_i^F, \frac{i-1}{n}\right) = u_i^n(x_i^F) \qquad \text{for } i = 1, \ldots, n,$$

we have

$$\sum_{i=1}^{n} \psi^n\left(x_i^F, \frac{i}{n}\right) - \psi^n\left(x_i^F, \frac{i-1}{n}\right) = \sum_{i=1}^{n} u_i^n(x_i^F),$$

which also represents $\geqslant$ on D_n^E.

Note that if $m = jn$ for some integer j, then

$$\psi^m\left(x, \frac{k}{n}\right) = \psi^n\left(x, \frac{k}{n}\right).$$

Therefore, applying the above definition for all n, we obtain a function

$$\psi: X \times \mathrm{Q} \to \mathrm{R}$$

where Q is the set of all rational numbers in [0,1].

Define $\tilde{\psi}(x,q) = \psi(x,q) - \psi(\bar{x},q)$.

Lemma

If $F \in D^E$ has mass points with rational probabilities, then $V(F) = \Sigma_{i=1}^n \tilde{\psi}(x_i^F,q_i) - \tilde{\psi}(x_i^F,q_{i-1})$ represents $\succcurlyeq$, where q_i is the cumulative at x_i^F.

Proof of lemma: Let $F \in D^E$ have mass points at $x_1^F, \ldots, x_n^F$, in increasing order,

$$\sum_{i=1}^n \tilde{\psi}(x_i^F,q_i) - \tilde{\psi}(x_i^F,q_{i-1})$$

$$= \sum_{i=1}^n c(x_i^F,q_i) - \psi(x_i^F,q_{i-1}) - \sum_{i=1}^n (\psi(\bar{x},q_i) - \psi(\bar{x},q_{i-1})). \tag{A.3}$$

However, the last summation is invariant to $(q_1, \ldots, q_n)$ and hence to F, because it represents the utility of a unit point mass at $\bar{x}$. Thus the left-hand side of (A.3) is also a representation of $\succcurlyeq$.

Lemma

$\tilde{\psi}$ is continuous.

Proof of lemma

Let F be the distribution: x_1 with probability q and $\bar{x}$ with probability $1 - q$ (denoted $(x_1,\bar{x},q,1 - q)$) where q is rational. We have

$$V(F) = \tilde{\psi}(x_1,q) - \tilde{\psi}(x,0) + \tilde{\psi}(\bar{x},1) - \tilde{\psi}(\bar{x},q),$$

which equals

$$V(F) = \psi(x_1,q) - \psi(\bar{x},q)$$

or

$$\tilde{\psi}(x_1,q).$$

$\tilde{\psi}(x_1,q)$ is continuous in x_1 from the continuity of the u_n^i's.

Suppose now that a sequence of rational numbers q_k converges to q and that $\tilde{\psi}(x_1,q_k)$ does not converge to $\tilde{\psi}(x_1,q)$, say it converges to $\tilde{\psi}(x_1,q) - \alpha$, where $\alpha > 0$. There exists some $\varepsilon > 0$ such that

$$\tilde{\psi}(x_1,q) - \frac{\alpha}{2} < \tilde{\psi}(x_1 - \varepsilon,q) < \tilde{\psi}(x_1,q).$$

For k large enough, the distribution $(x_1 - \varepsilon,\bar{x},q,1 - q)$ is strictly preferred to the distribution $(x_1,\bar{x},q_k,1 - q_k)$, which contradicts the continuity assumption. Q.E.D.

Since $V(F)$ is continuous on the distributions with rational probabilities, it has a continuous extension $\tilde{V}(F)$ on D.

Since $\tilde{\psi}$ is continuous, it has a unique continuous extension to $X \times [0,1]$, denoted $\bar{\psi}$. Moreover, $V(F) = \Sigma_{i=1}^{n}[\bar{\psi}(x_i,p_i) - \bar{\psi}(x_i,p_{i-1})]$ where F is the distribution with support $\{x_1,\ldots,x_n\}$ and cumulative p_i at x_i.

The function $\bar{\psi}$ induces a continuous measure on the Borel sets of $X \times [0,1]$ as follows: Consider the function μ induced on rectangles $[x_1,x_2] \times [p_1,p_2]$ by

$$\mu([x_1,x_2] \times [p_1,p_2]) = \bar{\psi}(x_2,p_2) - \bar{\psi}(x_2,p_1) - \bar{\psi}(x_1,p_2) + \bar{\psi}(x_1,p_1),$$

and let $\bar{\mu}$ be its Lebesgue extension to all Borel sets in $X \times [0,1]$.

Now as $\bar{\psi}$ is continuous, $\bar{\mu}$ will be continuous, in the sense that it has no point-mass.

By construction of the Lebesgue extension, the measure is set-continuous: that is, if $A_n \to A$ then $\mu(A_n) \to \mu(A)$ where $A_n \to A$ means $\lim\sup A_n = \lim\inf A_n = A$.

Let us define now $A_F = \{(x,p)|p \geqslant F(x)\}$. It is easy to construct, for each distribution F, a sequence of simple distributions such that $A_n \to A_F$ (A_n and A are the upper set), and $A_n \subseteq A_{n-1}$ and $A_F \subseteq A_n$. For a simple distribution,

$$\begin{aligned}
\bar{\mu}(A_n) &= \sum_{i=1}^{n} [\bar{\psi}(x_i,p_i) - \bar{\psi}(x_i,p_{i-1})] - \sum_{i=1}^{n} [\bar{\psi}(0,p_i) - \bar{\psi}(0,p_{i-1})] \\
&= \tilde{V}(F_n) - \sum_{i=1}^{n} \tilde{\psi}(0,p_i) + \sum_{i=1}^{n-1} \tilde{\psi}(0,p_i) \\
&= \tilde{V}(F_n) - \tilde{\psi}(0,1),
\end{aligned}$$

so for any F, $\bar{\mu}(A_F) = \tilde{V}(F) - \tilde{\psi}(0,1)$.

We can take $\bar{\mu}(A_F)$ as representation of the preference.

Now decompose $\bar{\mu}(dp,dx)$ as a marginal measure $\mu(dp)$ and a conditional distribution $\mu_2(p,dx)$.

$$\bar{\mu}(A) = \int \mu(dp) \int \mu_2(p,dx),$$

where the integral is taken over $(p,x) \in A_F$.

Write

$$\{x|(p,x) \in A_F\} = [0,z(p)];$$

define

$$\bar{\phi}(z(p),p) = \int_0^{z(p)} \mu_2(p,dx).$$

Then

$$\bar{\mu}(A_F) = \int \bar{\phi}(z(p),p)\mu(dp).$$

The corollary is obtained by decomposing the measure into the marginal measure $\nu(dx)$ and the conditional measure $\nu_2(p,dx)$.

Proof of Theorem 4: (ϕ,h) is more risk-averse than (ϕ^*,h^*) if and only if the function $x \to \int_0^x \phi_1(s,F(s))ds$ is a concave transform of the function $x \to \int_0^x \phi_1^*(s,F(s))ds$, for all F.

From Pratt's characterization of concave transforms, this is equivalent to

$$\forall F,\ \forall x_1,x_2,x_3, \qquad x_1 < x_2 < x_3.$$

$$\frac{\int_{x_1}^{x_2} \phi_1(s,F(s))ds}{\int_{x_1}^{x_2} \phi_1^*(s,F(s))ds} \geqslant \frac{\int_{x_2}^{x_3} \phi_1(s,F(s))ds}{\int_{x_2}^{x_3} \phi_1^*(s,F(s))ds}. \tag{A.4}$$

Taking the limit when $x_2 \to x_1$ and the limit when $x_2 \to x_3$, we find, when F is continuous,

$$\frac{\phi_1(x_1,F(x_1))}{\phi_1^*(x_1,F(x_1))} \geqslant \frac{\int_{x_1}^{x_3} \phi_1(s,F(s))ds}{\int_{x_1}^{x_3} \phi_1^*(s,F(s))ds} \geqslant \frac{\phi_1(x_3,F(x_3)}{\phi_1^*(x_3,F(x_3))}$$

so that $[\phi_1(x,F(x))/\phi_1^*(x,F(x))]$ is nonincreasing, which is equivalent to our statement. Suppose now that $[\phi_1(x,F(x))/\phi_1^*(x,F(x))]$ is nonincreasing. Then

$$\frac{\int_{x_1}^{x_2} \phi_1(s,F(s))ds}{\int_{x_1}^{x_2} \phi_1^*(s,F(s))ds} \geqslant \frac{\phi_1(x_2,F(x_2))}{\phi_1^*(x_2,F(x_2))} \geqslant \frac{\int_{x_2}^{x_3} \phi_1(s,F(s))ds}{\int_{x_2}^{x_3} \phi_1^*(s,F(s))ds}. \qquad \text{Q.E.D.}$$

Proof of Theorem 6:
Necessity:

$$\frac{dp_3}{dp_1} = \frac{\phi(x_2,p_1) - \phi(x_1,p_1)}{\phi(x_3,1-p_3) - \phi(x_2,1-p_3)}$$

must be nonincreasing with p_1, which implies $\phi_{12}(x,p) \leqslant 0$.

Sufficiency:

$$V(F_A) = V(F_B) \Leftrightarrow \int_X \int_{F_B(x)}^1 \phi_1(x,p)dp\ dx = \int_X \int_{F_A(x)}^1 \phi_1(x,p)dp\ dx.$$

$$V(F_C) \geqslant V(F_D) \Leftrightarrow \int_X \left\{ \int_{F_C(x)}^1 \phi_1(x,p)dp - \int_{F_D(x)}^1 \phi_1(x,p)dp \right\} dx \geqslant 0$$

$$\Leftrightarrow \int_{x<x^*} \left\{ \int_{F_C(x)}^{F_A(x)} \phi_1(x,p)dp - \int_{F_D(x)}^{F_B(x)} \phi_1(x,p)dp \right\} dx \geqslant 0$$

$$\Leftrightarrow \int_{x<x^*} \left\{ \int_{F_C(x)}^{F_A(x)} \phi_1(x,p)dp - \int_{F_C(x)+F_B(x)-F_A(x)}^{F_A(x)+F_B(x)-F_A(x)} \phi_1(x,p)dp \right\} dx \geqslant 0$$

$$\Leftrightarrow \int_{x<x^*} \int_{F_C(x)}^{F_A(x)} \{\phi_1(x,p) - \phi_1(x,p+F_B(x)-F_A(x))\} dp\ dx \geqslant 0.$$

As for $x \leqslant x^*$, $F_C(xd) \leqslant F_A(x)$ and $F_B(x) - F_A(x) \geqslant 0$, this is true if $\phi_1(x,p)$ is nonincreasing with p or $\phi_{12}(x,p) \leqslant 0$. Q.E.D.

Proof of Theorem 7:

We call F_p the distribution $(w - \varepsilon, w + (1 - p/p)\varepsilon, 1 - p, p)$.

$$V(G_w) = \int_0^1 \phi(w,s)ds,$$

$$V(F_p) = \int_0^{1-p} \phi(w-\varepsilon,s)ds + \int_{1-p}^1 \phi\left(w + \frac{1-p}{p}\varepsilon,s\right)ds,$$

$$\limsup_{p\to 0} V(F_p) = \int_0^1 \phi(w-\varepsilon,s)ds + \limsup_{p\to 0} \int_{1-p}^1 \phi\left(w + \frac{1-p}{p}\varepsilon,s\right)ds.$$

Suppose that $p_n \xrightarrow[n\to+\infty]{} 0$ and $V(G_w) \leqslant V(F_{p_n})$.
Then there exists $a > 0$, such that for all n,

$$\int_{1-p_n}^1 \phi\left(w + \frac{1-p_n}{p_n}\varepsilon,s\right)ds \geqslant 2a. \tag{A.5}$$

Now choose the sequence q_t by: $q_1 = p_1$, $q_{t+1} = p_n$ such that

$$\int_{1-q_t}^{1-q_{t+1}} \phi\left(w + \frac{1-q_t}{q_t}\varepsilon, s\right)ds > a.$$

Given that $q_t = p_n$ for some n, the inequality (A.5) insures that we can find q_{t+1}. Choose a distribution F as follows:

$$z(p) = w + \frac{1-q_t}{q_t}\varepsilon \qquad \text{if } p\varepsilon]1-q_t,\ 1-q_{t+1}[.$$

Then

$$V(F) \geqslant \sum_{t=1}^{\infty} \int_{1-q_t}^{1-q_{t+1}} \phi\left(w + \frac{1-q_t}{q_t}\varepsilon, s\right)ds = +\infty.$$

So the preferences are unbounded. Q.E.D.

Proof of Lemma (8.2):

Define $F_p(x) = (1-p)F(x + [p/1-p]\varepsilon) + pF(x-\varepsilon)$.

$$V(F_p) - V(F) = \int_X \int_{F_p(x)}^{F(x)} \phi_1(x,s)ds$$

$$= \int_X \left\{ \int_{(1-p)F(x)+pF(x-\varepsilon)}^{F(x)} \phi_1(x,s)ds + \int_{F[x+(p/1-p)\varepsilon]}^{F(x)} \phi_1(x,s)ds \right.$$

$$\left. + \int_{F(x)}^{(1-p)F(x)+pF(x-\varepsilon)} \phi_1(x,s)ds + \int_{(1-p)F[x+(p/1-p)\varepsilon]+pF(x-\varepsilon)}^{F[x+(p/1-p)\varepsilon]} \phi_1(x,s)ds \right\} dx.$$

a) The second line has derivative 0 at $p = 0$.

$$\forall \alpha > 0,\ \exists \beta > 0 / |s - F(x)| < \beta \Rightarrow |\phi_1(x,s) - \phi_1(x,F(x))| < \alpha.$$

Define

$$G(p) = \int_X \left\{ \int_{F(x)}^{(1-p)F(x)+pF(x-\varepsilon)} \phi_1(x,s)ds + \int_{(1-p)F[x+(p/1-p)\varepsilon]+pF(x-\varepsilon)}^{F[x+(p/1-p)\varepsilon]} \phi_1(x,s)ds \right\} dx.$$

$$\int_X \phi_1(x,F(x))\left(F\left(x + \frac{p}{1-p}\varepsilon\right) - F(x)\right)dx + 2\alpha \int_X (F(x) - F(x-\varepsilon))dx$$

$$\geqslant \frac{G(p)}{p} \geqslant \int_X \phi_1(x,F(x))\Big(F\Big(x + \frac{p}{1-p}\,\varepsilon\Big) - F(x)\Big)dx - 2\alpha \int_X (F(x) - F(x-\varepsilon))dx.$$

Using the right continuity of F, we find

$$\forall\alpha \qquad 2\alpha \int_X (F(x) - F(x-\varepsilon))dx \geqslant \lim_{p\to 0} \Big|\frac{G(p)}{p}\Big| \text{ or } \lim_{p\to 0} \frac{G(p)}{p} = 0.$$

b) The first term of the first line is $V((1-p)F + pF_\varepsilon) - V(F)$, where F_ε is the distribution $F_\varepsilon(x) = F(x-\varepsilon)$. By definition of U_F, its derivative is

$$\int_X U_F(x)\{dF(x-\varepsilon) - dF\} = E_F\{U_F(x+\varepsilon) - U_F(x)\}$$

The second term can be written

$$\int_X \int_{F(x)+(p/1-p)\varepsilon}^{F(x)} \phi_1(x,s)ds\ dx = \int_s \Big\{\phi\Big(z(s) - \frac{p}{1-p}\,\varepsilon,s\Big) - \phi(z(s),s)\Big\}ds.$$

Its derivative at $p = 0$ is $-\varepsilon\int_s\phi_1(z(s),s)ds$.

So the overall derivative is

$$E_F\{U_F(x+\varepsilon) - U_F(x)\} - \varepsilon \int_s \phi_1(z(s),s)ds.$$

Lemma (8.2) follows directly.

Proof of Lemma (8.3):

Define $F_p(x) = F(x + p\varepsilon)$ the distribution of the agent if he insures the risk.

$$V(F_p) - V(G) = \int_X \int_{F(x+p\varepsilon)}^{pF(x+\varepsilon)+(1-p)F(x)} \phi_1(x,s)ds$$

$$= \int_X \int_{F(x)}^{pF(x+\varepsilon)+(1-p)F(x)} \phi_1(x,s)ds + \int_X \int_{F(x+p\varepsilon)}^{F(x)} \phi_1(x,s)ds.$$

The first term is just $V(F) - V((1-p)F + pF_\varepsilon)$ where $F_\varepsilon(x) = F(x+\varepsilon)$; its derivative is

$$\int U_F(x)\{dF(x) - dF(x+\varepsilon)\} = E_F\{U_F(x) - U_F(x-\varepsilon)\}.$$

The second term is $\int_s\{-\phi(z(s),s) + \phi(z(s) - p\varepsilon,s)\}\,ds$; its derivative is $-\varepsilon\int_s\phi_1(z(s),s)ds$. So the total derivative is

$$E_F(U_F(x) - U_F(x - \varepsilon)) - \varepsilon \int_s \phi_1(z(s),s)ds.$$

Proof of Theorem 9:

For $F = G_w$ the relation (4) reduces to

$$\frac{\phi(w + \varepsilon,1) - \phi(w,1)}{\varepsilon} \geqslant \int_0^1 \phi_1(w,p)dp. \tag{A.6}$$

Since $\phi_{12}(w,p) < 0$, $\int_0^1 \phi_1(w,p)dp > \phi_1(w,1)$. Choose w arbitrarily; relation (A.6) implies that we can find $\varepsilon_1 > 0$, such that

$$\phi(w + \varepsilon_1,1) - \phi(w,1) \geqslant \varepsilon_1 \int_0^1 \phi_1(w,p)dp > \varepsilon_1\phi_1(w,1),$$

$$\phi_1(w + \varepsilon_1,1) \geqslant \phi_1(w,1).$$

Recursively, define ε_n by

$$\phi(w + \varepsilon_1 \ldots + \varepsilon_n,1) - \phi(w + \varepsilon_1 \ldots + \varepsilon_{n-1},1) \geqslant \varepsilon_n \int_0^1 \phi_1(w + \ldots + \varepsilon_{n-1})dp,$$

$$\phi_1(1 + \varepsilon_1 \ldots + \varepsilon_n,1) \geqslant \phi_1(w + \ldots + \varepsilon_{n-1},1). \tag{A.7}$$

Then

$$\phi(w + \ldots + \varepsilon_n,1) - \phi(w,1) > \phi_1(w,1)[\varepsilon_1 \ldots + \varepsilon_n].$$

Suppose that $\phi(x,1)$ is bounded; then $w + \varepsilon_1 \ldots + \varepsilon_n$ must converge to some $\bar{w}$. But then

$$\int_0^1 \phi_1(w + \varepsilon_1 \ldots + \varepsilon_{n-1},p)dp \xrightarrow[n\to+\infty]{} \int_0^1 \phi_1(\bar{w},p)dp,$$

while

$$\frac{\phi(w + \ldots \varepsilon_n,1) - \phi(w + \varepsilon_1 \ldots + \varepsilon_{n-1},1)}{\varepsilon_n} \xrightarrow[n\to+\infty]{} \phi_1(\bar{w},1).$$

So we find $\phi_1(\bar{w},1) \geqslant \int_0^1 \phi_1(\bar{w},p)dp$ which contradicts $\phi_{12}(x,p) < 0$. Therefore $\phi(x,1)$ is unbounded. Since $\phi(x,p) = \int_0^x \phi_1(s,p)ds > \int_0^x \phi_1(s,1)ds = \phi(x,1)$, $\phi(x,p)$ is unbounded for all p. Q.E.D.

References

Chew, S.H., and L.G. Epstein, (1987), "A Unifying Approach to Axiomatic Non-Expected Utility Theories," mimeo, University of Toronto.

Chew, S.H., E. Karni and Z. Safra, (1987), "Risk Aversion in the Theory of Expected Utility with Rank-Dependent Probabilities," *Journal of Economic Theory* 42, 370–381.

Chew, S.H. and K.R. MacCrimmon (1979), "Alpha Utility Theory: A Generalization of Expected Utility Theory," Faculty of Commerce and Business Administration Working Paper #669, University of British Columbia.

Dekel, E. (1986), "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory* 40, 304–318.

Gorman, W.M. (1968), "The Structure of Utility Functions," *Review of Economic Studies* 35, 367–390.

Gul, F. (1988), "Neo-Bernoullian Extensions and the Expected Utility Hypotheses and a Theory of Disappointment", mimeo., Stanford Graduate School of Business.

Friedman, M. and L.J. Savage (1948), "The Utility Analysis of Choices Involving Risk", *Journal of Political Economy* 56, 279–304.

Kahneman, D. and A. Tversky (1984), "Choices, Values and Frames", *American Psychologist* 39, 341–350.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47, 263–291.

Loomes, G. and R. Sugden (1982), "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty", *Economic Journal* 92, 805–824.

Machina, M. (1982), "'Expected Utility' Analysis without the Independence Axiom", *Econometrica* 50, 277–323.

MacCrimmon, K. and S. Larsen (1979), "Utility Theory: Axioms versus 'Paradoxes'", in *Expected Utility Hypotheses and the Allais Paradox.* M. Allais and O. Hager (eds.) Holland: D. Reidel.

Quiggin, J. (1982), "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization* 3, 323–343.

Royden, H. (1963), *Real Analysis*. Macmillan, New York.

Segal, U. (1987). "Two-Stage Lotteries without the Reduction Axiom". University of Toledo working paper no. 8708.

Segal, U. (1984), "Nonlinear Decision Weights With the Independence Axiom," UCLA Working Paper #353.

Yaari, M. (1987), "The Dual Theory of Choice Under Risk," *Econometrica* 55, 96–115.