Econometrica, Vol. 59, No. 3 (May, 1991), 667–686

A THEORY OF DISAPPOINTMENT AVERSION

BY FARUK GUL[1]

An axiomatic model of preferences over lotteries is developed. It is shown that this model is consistent with the Allais Paradox, includes expected utility theory as a special case, and is only one parameter (β) richer than the expected utility model. Allais Paradox type behavior is identified with positive values of β. Preferences with positive β are said to be disappointment averse. It is shown that risk aversion implies disappointment aversion and that the Arrow-Pratt measures of risk aversion can be generalized in a straight-forward manner, to the current framework.

KEYWORDS: Preferences over lotteries, expected utility theory, independence axiom, risk aversion, Arrow-Pratt measures of risk aversion.

INTRODUCTION

THE PURPOSE OF THIS PAPER is to develop an axiomatic model of decision making under uncertainty that (i) includes expected utility theory as a special case, (ii) is consistent with the Allais Paradox, and (iii) is the most restrictive possible model that satisfies (i) and (ii) above.

The difficulty is in providing a precise sense in which (iii) can be satisfied. We propose to do this as follows: We will present an intuitive explanation of the Allais Paradox. Then we will replace the independence axiom of expected utility theory with an alternative axiom which explicitly incorporates our intuitive explanation. An additional axiom which does not conflict with the intuitive explanation or with expected utility maximization will also be imposed. Analysis of the resulting model will reveal that it does indeed satisfy (i) and (ii) above and that no further qualitative restriction can be imposed without violating either (i) or (ii). Our aim is to show that the type of behavior exhibited by a large number of subjects in Allais' original experiment can be interpreted intuitively and justified within the framework of a reasonable model.

With this in mind, in what follows we characterize preferences that are described completely by a real-valued function u on the set of prizes and a real number $\beta > -1$ (Theorem 1). We show that u is unique up to an affine transformation and β is unique. Hence we isolate a class of preferences that is one parameter richer than von Neumann-Morgenstern preferences. We further show that $\beta = 0$ corresponds to the case of expected utility theory (where u is the von Neumann-Morgenstern utility function). We describe preferences with $\beta \geqslant 0$ as disappointment[2] averse and establish the equivalence of strict disap-

[1] I am indebted to Elchanan Ben-Porath, Eddie Dekel, Darrell Duffie, David Kreps, Mark Machina, Dilip Mookherjee, Ariel Rubinstein, Hugo Sonnenschein, Robert Wilson, two anonymous referees, and especially Outi Lantto for their help and criticism.

[2] The term *disappointment* was first used by Bell (1985) and Loomes and Sugden (1986). While we have borrowed the word from them, our motivation and the class of preferences that we consider are different.

pointment aversion (i.e. $\beta > 0$) and Allais Paradox type behavior (Theorem 2). Finally we show the relationship between risk aversion and disappointment aversion (Theorems 3–5). In particular we show that in this model risk aversion implies disappointment aversion.

THE ALLAIS PARADOX[3]

Consider an individual who is faced with the following two choice problems:

PROBLEM 1: Choose either p_1 or p_2 where p_1 is a degenerate lottery which yields 200 dollars for sure and p_2 is a lottery that yields 300 dollars with probability .8 and 0 dollars with probability .2.

PROBLEM 2: Choose either $\bar{p}_1$ or $\bar{p}_2$ where $\bar{p}_1$ is a lottery which yields 200 dollars with probability .5 and 0 dollars with probability .5 and $\bar{p}_2$ is a lottery which yields 300 dollars with probability .4 and 0 dollars with probability .6.

The propensity of decision makers to choose p_1 if confronted with the first problem and $\bar{p}_2$ if confronted with the second, is a phenomenon that is now widely known as the Allais Paradox. The term *paradox* is due to the fact that such preferences are not consistent with expected utility maximization.[4] In particular, this pair of choices is inconsistent with the independence axiom, which is a necessary condition for expected utility maximization. The independence axiom states that given any three lotteries p_1, p_2, and r and a number $\alpha \in (0, 1]$, p_1 is preferred to p_2 implies $\alpha p_1 + (1-\alpha)r$ is preferred to $\alpha p_2 + (1-\alpha)r$ (where $\alpha p_1 + (1-\alpha)r$ denotes the lottery which yields any prize x with probability $\alpha p_1(x) + (1-\alpha)r(x)$).

Letting r be the lottery which yields 0 dollars for sure, α equal $\frac{1}{2}$, and observing that $\bar{p}_1 = \alpha p_1 + (1-\alpha)r$, $\bar{p}_2 = \alpha p_2 + (1-\alpha)r$ establishes that the Allais Paradox above constitutes a violation of the independence axiom. Observe that in Problem 1, lottery p_1 has no chance of yielding a disappointing outcome whereas lottery p_2 has a .2 chance of yielding a disappointing outcome. One possible explanation of why the independence axiom fails in this particular example is that the lottery with a lower probability of disappointment suffers more when it is mixed with an inferior lottery (i.e., $r = 0$ dollars for sure); that is, if the lotteries were nearly indifferent initially, the lottery with the higher probability of disappointment becomes preferred after being mixed with the inferior lottery.

[3] This is not Allais' (1979) most famous example. This particular version is sometimes called the "Allais Ratio Paradox." It is also referred to as the common ratio effect or common consequence effect by Kahnemann and Tversky (1979). We use it here because the intuitive explanation of the type we wish to isolate is easier to express in terms of this slightly simpler example. However, the same intuitive argument applies to both this and the original version of the Allais Paradox and the notion of disappointment aversion resolves both versions.

[4] To check this note that $p_1 \succ p_2$ implies $u(200) > .8u(300) + .2u(0)$ and $\bar{p}_1 \prec \bar{p}_2$ implies $.5u(200) + .5u(0) < .4u(300) + .6u(0)$; i.e. $u(200) < .8u(300) + .2u(0)$, a contradiction.

In the words of Savage (1972, page 102), "Many people prefer Gamble 1 (p_1) to Gamble 2 (p_2) because, speaking qualitatively, they do not find the chance of winning a very large fortune in place of receiving a large fortune outright adequate compensation for even a small risk of being left in the status quo. Many of the same people prefer Gamble 4 ($\bar{p}_2$) to Gamble 3 ($\bar{p}_1$); because, speaking qualitatively, the chance of winning is nearly the same in both gambles, so the one with the much larger prize seems preferable." While 300 and 200 dollars hardly qualify as very large and large fortunes, it is clear that Savage's interpretation of the original version is closely related to our intuitive explanation here. What Savage calls the chance of winning is one minus what we have called the probability of disappointment which we will define formally in our model.

Before we begin our formal analysis two basic questions need to be addressed. First, why concentrate on the Allais Paradox as opposed to other systematic violations for the expected utility hypothesis? Second, what distinguishes our approach from other axiomatic models of choice under uncertainty that allow for Allais Paradox type behavior?

There is a large body of work on observed violations of the expected utility model. Historically the Allais Paradox has played a very significant role in the development of this literature. This is no doubt in part due to the intuitive appeal of the Allais Paradox choices. Hence it would appear that the Allais Paradox is a natural starting point for any attempt at reconciling the normative theory of choice under uncertainty with the existing empirical evidence.

One can identify at least three distinct ways that the non-expected utility literature has dealt with observed violations of the expected utility theory:

(a) By emphasizing the need for a purely descriptive theory. Such work has either attempted to describe the actual decision making process that is used by the subjects (see Kahneman and Tversky (1979) and Rubinstein (1988)) or to identify useful (i.e., consistent with the existing empirical evidence) functional forms. Regret theory (Bell (1982), Loomes and Sugden (1962)), the disappointment theory of Bell (1985) and Loomes and Sugden (1982), the subjective expected value models used in the psychology literature (see Edwards (1953) and Tversky (1967) among others) are some of the many examples that can be included under this category. What is common to this particular body of work is the emphasis on descriptive aspects and skepticism regarding relevance of a normative theory. Hence the models mentioned above often violate even the most basic desiderata of choice under uncertainty (transitivity, stochastic dominance, etc.).

(b) By rejecting the normative appeal of the independence axiom. Allais (1979) and Machina (1982) belong in this category. Allais argues for a cardinal measure of utility over sure prospects and postulates that individuals' utility for uncertain prospects will depend on the distribution of the cardinal measure, typically its first three moments. Machina (1982) considers preferences that can be represented by a "smooth" preference functional and develops the machinery for analyzing the local properties of a preference functional. He offers two empirical hypotheses which he states in terms of these local properties. The first is risk aversion. The second which is called Hypothesis II is shown to imply behavior consistent with the Allais Paradox and a number of other observed violations of the expected utility theory. A more detailed comparison between the model of this paper and Machina's generation of expected utility theory will be provided after the formal analysis of the next section.

(c) By modifying the independence axiom. This class of papers starts by offering similar (typically weaker) alternatives to the independence axiom. The resulting model is defended by pointing out that it is consistent with observed violations of expected utility theory and by arguing that the alternative assumption is more compelling than the independence axiom. Some examples of this type of work are: Chew and MacCrimmon (1979), Dekel (1986), Fishburn (1983), and Yaari (1987).

This paper belongs among the work cited under (c) above. What distinguishes the model of this paper is our emphasis on the Allais Paradox and the direct role it plays in our axiomatization. Hence we provide a narrow interpretation of the Allais Paradox and search for a generalization of expected utility theory which is consistent with this interpretation and yet allows us to retain as much of the insight offered by expected utility theory as possible.

The Model

For some b, w such that $b > w$, let $X = [w, b]$ be the set of all prizes. Let $\mathscr{L}$ be the set of all simple lotteries over these prizes. That is, $p \in \mathscr{L}$ implies that $\operatorname{supp}(p)$, the support of p is finite. For any $p, q \in \mathscr{L}$ and $\alpha \in [0,1]$, $\alpha p + (1-\alpha)q$ denotes the lottery $r \in \mathscr{L}$ such that for all $x \in X$, $r(x) = \alpha p(x) + (1-\alpha)q(x)$. When there is no risk of confusion we use $x \in X$ to denote the lottery p such that $p(x) = 1$. $\succsim$ is a binary relation on $\mathscr{L}$. We use $p \succ q$, "p is strictly preferred to q", to denote $p \succsim q$ and not $q \succsim p$. We use $p \sim q$, "p is indifferent to q", to denote $p \succsim q$ and $q \succsim p$.

Since, typically, we want to interpret $x \in X$ as a quantity of money, $x > y$ iff $x \succ y$, will be a maintained assumption throughout this paper.

DEFINITION 1: For any $\succsim$ and p, let

$$B(p, \succsim) = \{q \in \mathscr{L} \mid x \in \operatorname{supp}(q) \text{ implies } x \succsim p\};$$

$$W(p, \succsim) = \{q \in \mathscr{L} \mid x \in \operatorname{supp}(q) \text{ implies } p \succsim x\}.$$

We sometimes use $B(p), W(p)$ instead of $B(p, \succsim)$ and $W(p, \succsim)$.

Thus $B(p)$ and $W(p)$ denote the set of lotteries with supports consisting of prizes respectively, better than and worse than p.

DEFINITION 2: (α, q, r) is an *elation/disappointment decomposition* (EDD) of p iff $q \in B(p)$, $r \in W(p)$ and $\alpha q + (1-\alpha)r = p$.

Thus an EDD of p is constructed as follows: The lottery is divided into two parts, those prizes which are preferred to the certainty equivalent of p (called elation prizes) and those prizes which are less preferred to the certainty equivalent of p (called disappointment prizes). Then we normalize by dividing the probability of all elation prizes by α, the sum of all elation prize probabilities and obtain q. Similarly we divide all disappointment prize probability by $1-\alpha$ and obtain r. Hence $\alpha q + (1-\alpha)r = p$ (note that $\alpha q + (1-\alpha)r$ is p, not just indifferent to p). Obviously if the certainty equivalent of p is not in the support of p there is a unique EDD for p. Otherwise there will be an infinity of EDD's for p. To see this note that $(.2, x, x)$, $(.7, x, x)$ and $(0, b, x)$ are all EDD's of x.

Next we define elation, $e(p)$, and disappointment, $d(p)$, probabilities for a lottery p. Note that if p does not yield its certainty equivalent with positive probability, then $e(p) + d(p) = 1$ and $D(p) = \{(e(p), q, r)\}$ for some $q, r \in \mathscr{L}$.

DEFINITION 3: $e(p) \equiv \sum_{x \succ p} p(x)$ and $d(p) \equiv \sum_{p \succ x} p(x)$.

We use $D(p)$ to denote the set of all EDD's of p. Instead of $(\alpha, q, r) \in D(\alpha q + (1-\alpha)r)$ or equivalently $(\alpha, q, r) \in D((\alpha, q, r))$ we simply write $(\alpha, q, r) \in D$ where $D = \bigcup_{p \in \mathscr{L}} D(p)$. (Note that by definition if (α, q, r) is an EDD it must be an EDD of $\alpha q + (1-\alpha)r$.)

AXIOM 1—Preference Relation: $\succsim$ *is complete and transitive.*

AXIOM 2—Continuity: *For all* $p \in \mathscr{L}$ *the sets* $\{q \in \mathscr{L} \mid q \succsim p\}$ *and* $\{q \in \mathscr{L} \mid p \succsim q\}$ *are closed (under the topology generated by the* L^1 *metric).*[5]

Axiom 2 implies that the function CE: $\mathscr{L} \to [w, b]$ such that $CE(p) \sim p$ (i.e. CE is the certainty equivalent of p) is well-defined.

Next we will present a restriction of the independence axiom to the case in which the disappointment probabilities of the lotteries p_1 and p_2 are the same and no elation (disappointment) prize of p_i switches over to being a disappointment (elation) prize of $ap_i + (1-a)x$. The motivation for this is the intuitive explanation of the Allais Paradox that was presented earlier. Consider the lotteries $p_i^t = tp_i + (1-t)x$ as t decreases from 1 to a. If $p_1 \succsim p_2$ and no elation (disappointment) prize of p_i^t switches to being a disappointment (elation) prize, then the disappointment probabilities of p_1^t and p_2^t are always the same; hence our intuitive explanation of the Allais Paradox is not applicable so we would expect the conclusion of the independence axiom to be valid.

[5] To be more precise, let f: $\mathscr{L} \to \mathscr{L}^1$ where $f(p)$ is the cdf associate with p. Then, Axiom 2 requires that $f(\{q \in \mathscr{L} \mid q \succsim p\})$ and $f(\{q \in \mathscr{L} \mid p \succsim q\})$ are closed (for every $q \in \mathscr{L}$) in the relative topology on $f(\mathscr{L})$ generated by the L^1 metric.

Axiom 3—Weak Independence: $p_1 \succsim p_2$, $a \in [0,1]$, $z \in X$ *implies* $ap_1 + (1-a)z \succsim ap_2 + (1-a)z$ *whenever there exists* $(\lambda, q_i, r_i) \in D(p_i)$ *such that* $q_i \in B(ap_i + (1-a)z)$ *and* $r_i \in W(ap_i + (1-a)z)$ *for* $i = 1,2$.

It can be seen from the proof of Theorem 1, that Axioms 1–3 imply betweenness. That is if $\succsim$ satisfies Axioms 1–3 and $p \succ q (p \sim q)$, then $p \succ \alpha p + (1-\alpha) q \succ q (p \sim \alpha p + (1-\alpha)q)$ for all $\alpha \in (0,1)$. Hence preferences which satisfy Axioms 1–3 belong to the class studied by Dekel (1986).

Axiom 3 captures our intuitive explanation of the Allais Paradox by enabling the independence axiom to fail when disappointment effects are present. However, in order to verify the "minimality" criterion (iii) discussed in the introduction, we need to determine if the class of preferences which are characterized by Axioms 1–3 can be restricted further without excluding expected utility preferences of our intuitive explanation. To put it differently, are there situations in which the independence axiom would be applicable but our intuitive explanation would not? Consider the following example:

Let $\alpha x + (1-\alpha)w \succsim \alpha p + (1-\alpha)w$ and $p \in B(\alpha p + (1-\alpha)w)$. Thus the decision-maker prefers substituting x in place of p in $\alpha p + (1-\alpha)w$ when p consists of elation prizes of $\alpha p + (1-\alpha)w$. Now assume that $p \in W(\alpha b + (1-\alpha)p)$. Hence p consists of disappointment prizes of $\alpha b + (1-\alpha)p$. Note that the independence axiom would imply that $\alpha b + (1-\alpha)x \succsim \alpha b + (1-\alpha)p$ whenever $\alpha x + (1-\alpha)w \succsim \alpha p + (1-\alpha)w$. Furthermore observe that $\alpha b + (1-\alpha)x$, $\alpha b + (1-\alpha)p$, $\alpha x + (1-\alpha)w$, and $\alpha p + (1-\alpha)w$ all have the same disappointment probability $(1-\alpha)$. Thus substituting x in place of p does not result in the type of effect discussed in our intuitive explanation of the Allais Paradox. Hence we would again expect the independence axiom to hold (i.e., $\alpha b + (1-\alpha)x \succsim \alpha b + (1-\alpha)p$). This will be Axiom 4. To see why this particular application of the independence axiom is not covered by Axiom 3, note that by requiring that no elation prize switches to being a disappointment prize, Axiom 3 severs the connection between the individual's evaluation of elation prizes and his evaluation of disappointment prizes.

Axiom 4—Symmetry: *For* $i = 1,2$, $(\alpha, p_i, w), (\alpha, b, p_i) \in D$ *implies*

$$\alpha p_1 + (1-\alpha) w \succsim \alpha p_2 + (1-\alpha) w \qquad \textit{iff}$$

$$\alpha b + (1-\alpha) p_1 \succsim \alpha b + (1-\alpha) p_2 .$$

Theorem 1: $\succsim$ *satisfies Axioms* 1–4 *if and only if there exist functions* u: $X \to \Re$ *and* γ: $[0,1] \to [0,1]$ *such that*: (i) $(\alpha_i, q_i, r_i) \in D(p_i)$ *for* $i = 1,2$ *implies* $p_1 \succsim p_2$ *iff* $\gamma(\alpha_1)\Sigma_x u(x) q_1(x) + (1-\gamma(\alpha_1))\Sigma_x u(x) r_1(x) \geqslant \gamma(\alpha_2)\Sigma_x u(x) q_2(x) + (1-\gamma(\alpha_2))\Sigma_x u(x) r_2(x)$; (ii) γ', u' *satisfy* (i) *above implies* $u' = au + b$ *for some* $a > 0$, $b \in \Re$ *and* $\gamma' = \gamma$; (iii) u *is continuous and there exists* $\beta \in (-1, \infty)$ *such* *that*

$$\gamma(\alpha)=\frac{\alpha}{1+(1-\alpha)\beta} \qquad \textit{for all} \qquad \alpha\in[0,1].$$

PROOF: See Appendix.

Theorem 1 establishes that if Axioms 1–4 hold, then there exists a utility function V: $\mathscr{L}\to\Re$ which represents $\succsim$ and furthermore $V(p)$ can be calculated by taking an EDD (α, q, r) of p, computing the expected utilities of the elation and disappointment parts (q and r respectively) with respect to the utility index u, and taking a $\gamma(\alpha)$ weighted average of these utilities. Hence u and β are parameters of the individual's preferences and $V(p)$ is defined implicitly by the procedure above. To see that $V(p)$ is not explicitly defined note that the certainty equivalent of p needs to be known in order to determine an EDD of p.

However, a simple and finite algorithm (see Appendix) will enable us to construct all EDD's of p and commute $V(p)$ for arbitrary u and β.

The fact that $V(p)$ is well defined for any $\succsim$ which satisfies Axioms 1–4 is guaranteed by Theorem 1. However, this does not preclude the possibility that there might be no non-expected utility preference which satisfies Axioms 1–4. Defining $V(u,\beta,p)$ implicitly by $V(u,\beta,p)=\gamma(\alpha)Eu(q)+(1-\gamma(\alpha))Eu(r)$ for some α, q, r such that $\alpha q+(1-\alpha)r=p$ and $x\in\operatorname{supp}(q)$ implies $u(x)\geqslant V(u,\beta,p)$ and $x\in\operatorname{supp}(r)$ implies $u(x)\leqslant V(u,\beta,p)$ and showing that $V(u,\beta,\cdot)$ is a well defined function for arbitrary strictly increasing, continuous u and $\beta\in(-1,\infty)$ would establish that Axioms 1–4 characterize a rich class of preferences. This can be done using simple manipulations of the definition of $V(u,\beta,p)$.

Observe that expected utility theory corresponds to the special case $\gamma(\alpha)=\alpha$; that is, $\beta=0$. Furthermore, if $\beta>0$, then $\gamma(\alpha)<\alpha$ for all $\alpha\in(0,1)$ and $\gamma(\alpha)$ is convex. If $-1<\beta<0$ then $\gamma(\alpha)>\alpha$ for all $\alpha\in(0,1)$ and $\gamma(\alpha)$ is concave. We say that $\succsim$ is disappointment averse if $\beta\geqslant 0$ and $\succsim$ is elation loving if $\beta\in(-1,0]$. Note that unlike risk aversion, disappointment aversion is, by definition, a global property. Theorem 2 below (and its proof) reveals that β is a measure of the extent to which $\succsim$ is prone to Allais Paradox type behavior. Since the preferences which satisfy Axioms 1–4 are only one-parameter (β) richer than expected utility preferences, it would appear that no additional qualitative restrictions can be imposed without excluding either our intuitive explanation of the Allais Paradox or certain expected utility preferences.

THEOREM 2: *Let* $\succsim$ *satisfy Axioms* 1–4 *and* $p\sim q$. *Then if* $\beta>0$ ($\beta<0$) *there exists* $\bar{a}>0$ *such that* (i) $a<\bar{a}$, $e(p)>e(q)$ *implies* $ax+(1-a)p\succ(\prec)ax+(1-a)q$ *for* $x\succ p$; (ii) $a<\bar{a}$, $d(p)>d(q)$ *implies* $ax+(1-a)p\succ(\prec)ax+(1-a)q$ *for* $p\succ x$.

PROOF: See Appendix.

Let (u, β) denote the generic preference satisfying Axioms 1–4.

Define

$$\phi(x,v)=\begin{cases} u(x) & \text{for } x \text{ such that } u(x)\leqslant v,\\ \dfrac{u(x)+\beta v}{1+\beta} & \text{for } x \text{ such that } u(x)>v.\end{cases}$$

Observe that by using the definition of $v(p)$ provided in Theorem 1 we obtain that $\Sigma_x \phi(x,v)p(x)=v$ iff $V=v(p)$, hence ϕ is the local utility function for the preference (u,β) (see Dekel (1986) for the definition and analysis of local utility functions of this form).

Roughly speaking, given Axioms 1, 2, Axiom 3 guarantees that the local utility function has the following property: All elation prizes are evaluated with respect to one utility function and all disappointment prizes are evaluated with respect to another utility function. Symmetry (Axiom 4) guarantees that these utility functions represent the same preferences. That is, the utility function for elation prizes $(u(x)+\beta v)/(1+\beta)$ is a (positive) affine transformation of the utility function for disappointment prizes, $u(x)$.

Abandoning the symmetry (Axiom 4) assumption would lead to the following local utility function:[6]

$$\phi(x,v)=\begin{cases} u_d(x) & \text{for } x \text{ such that } u_d(x)\leqslant v,\\ u_e(x)-u_e\big(u_d^{-1}(v)\big)+v & \text{for } x \text{ such that } u_d(x)>v,\end{cases}$$

where u_d, u_e are two distinct functions from $[w,b]$ to $\Re$. Note that Axiom 4 implies $u_e - u_e(u_d^{-1}(v))+v$ and hence u_e, is an affine transformation of u_d and expected utility implies $u_e = u_d$.

It can be shown that for preferences which satisfy Axioms 1–4, $\beta \geqslant 0$ iff $d(\Sigma\phi(x,v)p(x))/dv$ is an increasing function of $d(p)$, the probability of disappointment (among p such that $V(p)=v$). This observation can be used to extend the notion of disappointment aversion to preferences which satisfy Axioms 1–3. For such preferences $d(\Sigma\phi(x,v)p(x))/dv$ is a decreasing function of $d(p)$ iff $(u_e'(y)/u_d'(y))<1$ for y such that $u_d(y)=v$. But now disappointment aversion has become a local property and global disappointment aversion can be imposed by requiring disappointment aversion at every point y. With this extended definition of disappointment aversion the results that Allais Paradox implies disappointment aversion (Theorem 2) and that risk aversion implies disappointment aversion (Theorem 3) can be generalized to preferences which satisfy Axioms 1–3.

We end this section by noting that imposing Axiom 4 is consistent with our objective of seeking a model in which any deviation from expected utility theory can be ascribed to disappointment aversion. Furthermore adding Axiom 4

[6] The existence of some local utility function is guaranteed by Dekel's (1986) Proposition 1 and the fact that Axioms 1–3 imply betweenness. Furthermore Axiom 3 implies that (local) preferences over disappointment outcomes and elation outcomes (but not combinations of disappointment and elation outcomes) are independent of v. Then a suitable normalization yields the representation provided above.

enables us to analyze a particular simple subclass of the preferences that satisfy Axioms 1–3 for which we can obtain a nearly closed form representation.

DISAPPOINTMENT AVERSION AND RISK AVERSION

In this section we analyze the relationship between disappointment aversion and risk aversion and develop measures of risk aversion for preferences satisfying Axioms 1–4. Hence, in what follows we will concentrate only on preferences satisfying Axioms 1–4 and sometimes use (u, β) to denote such preferences.

THEOREM 3: *(u, β) is risk averse (in the sense of weakly not preferring mean-preserving spreads) iff $\beta \geqslant 0$ and u is concave.*

PROOF: Dekel (1986) establishes that $\succsim$ is risk averse iff the local utility function is concave. Note that $\phi(x, v)$ is concave if $\beta \geqslant 0$ and u is concave. *Q.E.D.*

There are two main implications of Theorem 3. The first one is that risk aversion implies disappointment aversion. The second is that disappointment aversion and the concavity/convexity of u determine the individual's attitude towards risk.

The possibility of having concave u and $\beta < 0$ or convex u and $\beta > 0$ enables us to obtain preferences that display risk aversion with respect to certain types of gambles and risk loving with respect to others. For example, if $\beta = 4$ and $u(x) = x$ for $x \leqslant 0$ and $u(x) = 5x$ when $x > 0$ (hence x is convex), then the individual will be risk averse with respect to even chance gambles and gambles which yield a large loss with small probability but will be risk loving with respect to gambles that involve winning a large prize with small probability if his initial income is low. Hence, there are preferences consistent with Axioms 1–4 such that at all income levels, the individual would not accept fair even chance gambles, yet would still be willing to buy less than fair insurance. Furthermore such an individual would be willing to buy, at certain income levels, tickets to the state lottery.

It is possible to develop measures of absolute, relative, and comparative risk aversion for preferences satisfying Axioms 1–4, similar to those developed by Arrow and Pratt for expected utility theory. Of course, these measures coincide with the corresponding Arrow-Pratt measures when $\beta = 0$. However, it is interesting that essentially the same Arrow-Pratt measures are appropriate even when $\beta \neq 0$. More specifically, let $R_u^a(x) = -u''(x)/u'(x)$ and $R_u^r(x) = -xu''(x)/u'(x)$; then by noting that β does not depend on x and essentially replicating the corresponding proofs of expected utility theory we obtain the following theorem.

THEOREM 4: *(u, β) is increasingly (decreasingly, constant) absolute (relative) risk averse iff $R^a(R^r)$ is increasing (decreasing, constant).*

DEFINITION 4: $\succsim_1$ is more risk averse than $\succsim_2$ iff $p \succsim_1 x$ implies $p \succsim_2 x$ for all $p \in \mathscr{L}$, $x \in X$.

THEOREM 5: *(u_1, β_1) is more risk averse than (u_2, β_2) if $\beta_1 \geqslant \beta_2$ and $R^a_{u_1}(x) \geqslant R^a_{u_2}(x)$ for all $x \in (w, b)$. Furthermore if (u_1, β_1) is more risk averse than (u_2, β_2), then $\beta_1 \geqslant \beta_2$.*

PROOF: See Appendix.

Machina (1984) provides the following stronger notion of comparative risk aversion: $\succsim_1$ is more risk averse than $\succsim_2$ iff for all $\lambda \in [0,1]$, $p, p' \in \mathscr{L}$; $\lambda p' + (1-\lambda)p \sim_2 \lambda p' + (1-\lambda)\bar{x}$ and $\lambda p' + (1-\lambda)p \sim_1 \lambda p' + (1-\lambda)x$ implies $x \leqslant \bar{x}$.

Given Theorem 3 and the observation that β and the curvature of u determines the curvature of the local utility function, we would expect that being more risk averse corresponds to having a higher β and a more concave u. This is almost correct in the sense that (u_1, β_1) is more risk averse (in the stronger sense) than (u_2, β_2) iff $R^a_{u_1}(x) \geqslant R^a_{u_2}(x)$ for all $x \in (w, b)$ and $\beta_1 \geqslant 0 \geqslant \beta_2$. To see why $\beta_1 \geqslant 0 \geqslant \beta_2$ cannot be replaced by $\beta_1 \geqslant \beta_2$, note that if $\beta_1 \geqslant \beta_2 > 0$,[7] then both (u_1, β_1) and (u_2, β_2) are risk averse (Theorem 3). Furthermore the certainty equivalent y_2, of $\lambda p + (1-\lambda)p'$ for (u_2, β_2) is greater than its certainty equivalent y_1, for (u_1, β_1) (Theorem 5). Then, at y_2, the local utility function ϕ_1 is smooth whereas the local utility function ϕ_2 has a concave kink. This means that ϕ_1 is less concave than ϕ_2 at y_2. Machina (1982), however, shows that more risk aversion is equivalent to the greater concavity of the local utility function.[8] Thus the comparison fails. Intuitively the fact that $y_2 > y_1$ makes it possible for (u_2, β_2) to view p' as increasing the probability of disappointment while (u_1, β_1) views p' as decreasing the probability of disappointment and therefore is less reluctant to accept it.

OTHER MODELS ON DECISION MAKING UNDER UNCERTAINTY AND EXPERIMENTAL EVIDENCE

It should be noted that although we have so far concentrated on the case where X is a compact interval, it is clear that the type of preference that we have been considering can just as well be defined on $\mathscr{L}$ when X is either finite or an unbounded interval and $\mathscr{L}$ includes nonsimple lotteries. All of the qualitative conclusions of this paper (Theorems 2–5) would still hold. If X is finite, however, Axioms 1–4 are not sufficient to characterize (u, β) preferences. This is due to the fact that our proof necessitates that certainty equivalents be well defined. The case of unbounded X and/or nonsimple lotteries can be dealt with as is done in expected utility theory. For nonsimple lotteries a finite algorithm for explicitly computing $V(p)$ would no longer exist. Instead an algorithm which converges to $V(p)$ can be constructed.

The most striking feature of these preferences is that by adding only one new variable, β, to the expected utility model, we are able to construct a class of

[7] A symmetric argument applies to the case $0 > \beta_1 \geqslant \beta_2$.

[8] Theorem 4 of Machina (1982) assumes the global differentiability of local utility functions. However, similar arguments can be used for the preferences considered here.

preferences which include the *EU* model as a special case, are compatible with many of the observed violations of this model, and enable us to explain these violations in terms of the notion of disappointment aversion. Observe that when we are comparing binary lotteries with each other or binary lotteries with sure things, then the functional form we have considered can be expressed as

$$V(ax+(1-a)z)=\pi_1(a)u(x)+\pi_2(1-a)u(z)$$
$$=\frac{a}{1+(1-a)\beta}u(x)+\frac{(1-a)(1+\beta)}{1+(1-a)\beta}u(z)$$

where $x \geqslant z$ and obviously $\pi_2(1-a)=1-\pi_1(a)$.

But this is very similar to the subjective expected utility models that have been used extensively in the psychology and economics literature (see, for example, Kahneman and Tversky (1979)) and is a special case (for binary lotteries) of Quiggin (1982).

Note also that for the case of $\beta > 0$, the weight function of the good prize is convex. Hence, a small increase in the probability of the good prize increases utility much more, when the chance of getting the good prize is already high. This is very suggestive of what Kahneman and Tversky (1979) refer to as the tendency for "people [to] overweight outcomes that are considered certain, relative to outcomes that are merely probable," i.e., the so-called certainty effect. The class of (u,β) preferences also have the following feature: If X is finite and $|X|=n$, then the preferences of any individual can be determined uniquely by asking him $n-1$ simple questions and solving a quadratic equation. This is only one more than the number of questions one would need to ask under the assumptions of expected utility theory.

Consider a person, with preferences (u,β), who will receive income x in state 1 and income y in state 2. Let α and $(1-\alpha)$ denote the probabilities of states 1 and 2 respectively (see Figure 1). An indifference curve for such a person will be described by the equation

$$\frac{\alpha}{1+(1-\alpha)\beta}u(x)+\frac{(1-\alpha)(1+\beta)}{1+(1-\alpha)\beta}u(y)=v \qquad \text{for } x \geqslant y$$

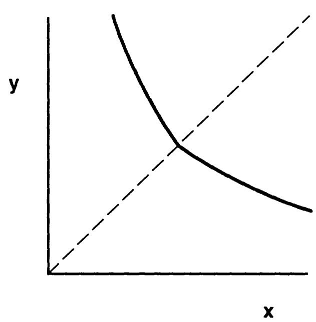

FIGURE 1

and

$$\frac{1-\alpha}{1+\alpha\beta}u(y)+\frac{\alpha(1+\beta)}{1+\alpha\beta}u(x)=v \qquad \text{for } y>x.$$

Hence

$$\frac{dy}{dx}=-\frac{\alpha u'(x)}{(1-\alpha)u'(y)(1+\beta)} \qquad \text{if } y<x,$$

$$\frac{dy}{dx}=-\frac{\alpha u'(x)(1+\beta)}{(1-\alpha)u'(y)} \qquad \text{if } y>x.$$

Thus if u is concave and $\beta>0$, then these preferences would look as in Figure 1 above. Note that the shape of the indifference curve on either side of the 45° line is determined by the curvature of u and the nature of the kink is determined by β. In particular there will be a kink so long as $\beta \neq 0$. Figure 1 shows why $\beta>0$ is necessary for risk aversion (see Theorem 3). This particular shape of the indifference curve has two important implications. The first is that risk averse individuals will purchase full insurance at less than fair odds; the second is that these preferences are not "differentiable" when $\beta \neq 0$ and hence do not belong to the class of preferences considered by Machina (1982).

Figure 2 illustrates the indifference map of (u,β) for lotteries over three prizes x, y, z where $x<y<z$. Hence any point (p_x, p_z) in Figure 2 corresponds to the lottery p such that $p(x)=p_x$, $p(y)=1-p_x-p_z$, and $p(z)=p_z$. After normalizing u so that $u(z)=1$ and $u(x)=0$, the indifference curve through any lottery p such that $p \succ y$ is defined by equation (1) and the indifference curve through any p such that $p \precsim y$ is defined by equation (2):

$$(1) \qquad p_z=\frac{(1+\beta)[u(y)p_x+v-u(y)]}{1+\beta v-(1+\beta)u(y)} \qquad \text{for } v>u(y),$$

$$(2) \qquad p_z=\frac{(u(y)+\beta v)p_x+v-u(y)}{1-u(y)} \qquad \text{for } v\leqslant u(y).$$

Both 1 and 2 can be derived from the definition of $V(p)$ provided in Theorem 1.

Equations 1 and 2 imply that the indifference curves on the top half of Figure 2 all intersect at the point $((1-u(y))/u(y)\beta, (1+\beta)/\beta)$ and indifference curves on the bottom half of Figure 2 all intersect at the point $(-1/\beta, -(1+\beta)u(y)/\beta(1-u(y)))$. Hence for $\beta>0$ all indifference curves in the bottom half are "fanning out" and all indifference curves in the top half are coming together from left to right (as depicted in Figure 2). For $\beta<0$ the opposite is true. Hence for the case of lotteries over three prizes, taking the preference (u,β) and "flipping" half of its indifference map, we obtain the type of preference considered by Chew and MacCrimmon (1979) and Fishburn

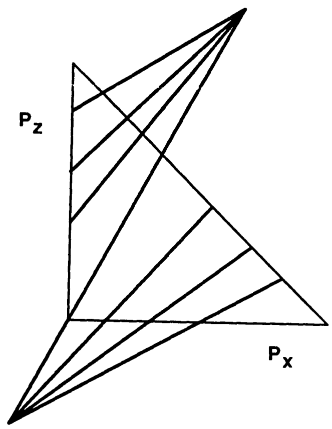

FIGURE 2

(1985) which are defined by the property that all indifference curves originate from the same point.

Machina (1982, 1987) considers a large class of systematic violations of the expected utility model and shows that these violations, which are frequently encountered in experiments, all imply that indifference curves in Figure 2 are fanning out. Since any individuals with preferences (u, β) will have indifference curves that fan out either in the top half or the bottom half (but not both) of Figure 2, no such individual can exhibit all of the violations considered by Machina.

It is undeniably true that fanning out over some range is necessary for Allias Paradox type behavior. The preference (u, β) will display the fanning out properly over the range concerning the Allais Paradox if and only if $\beta > 0$. Hence there is no conflict between Machina's observation regarding fanning out and Theorem 2. There are however substantial differences in the two interpretations of the Allais Paradox and the underlying approaches to violations of expected utility theory. Whereas Machina views the Allais Paradox as a special case of (global) fanning out, we have emphasized our narrower intuitive explanation and sought to provide a model which incorporates this intuitive explanation while retaining many of the features of expected utility theory. In the case of lotteries over three outcomes no two distinct lotteries can be indifferent and have the same disappointment probability. Thus Figure 2 understates the similarity between (u, β) preferences and expected utility theory. To see this, consider the following example: Suppose $\succsim$ satisfies

$$(3) \qquad \frac{1}{2} \times \left(\frac{1}{2} \times 1000 + \frac{1}{2} \times 800 \right) + \frac{1}{2} \times 100 \sim \frac{1}{2} \times 880 + \frac{1}{2} \times 100 \sim 400;$$

then Axioms 1–4 imply that $\succsim$ satisfies

$$(4) \qquad \frac{1}{2} \times \left(\frac{1}{2} + 1000 + \frac{1}{2} \times 800 \right) + \frac{1}{2} \times 50 \sim \frac{1}{2} \times 880 + \frac{1}{2} \times 50.$$

Under Axioms 1–4 the necessity of satisfying (4) whenever (3) is satisfied does not conflict in the displaying Allais Paradox type behavior at every income level. Under Machina's interpretation any individual who displays Allais type behavior at every income level and satisfies (3) must violate (4). This follows immediately from fanning out.

Not surprisingly there is some empirical and experimental evidence conflicting with Axioms 1–4 and Hypothesis II. In particular for $\beta > 0$ (which we consider to be the more important case), (u, β) will be consistent with the common ratio effect, partly consistent with the common sequence effect (including the Allais Paradox), and inconsistent with the common ratio effect with negative numbers (see Machina (1987)). Conversely for $\beta < 0$, (u, β) will be consistent with the common ratio effect with negative numbers, partly consistent with the common consequence effect (excluding the Allais Paradox), and inconsistent with the common ratio effect.

Neilson (1989) considered lotteries over three prizes and concludes that existing empirical evidence suggests the need for preferences which fan in on the top part and fan out on the bottom part of the probability triangle (i.e., exactly the situation depicted by Figure 2). The model of this paper (for $\beta > 0$) always has this property. To put it another way, we have identified Allais Paradox with precisely this mixed-fanning property.

CONCLUSION

We have taken what is considered to be the most compelling argument against the independence axiom and attempted to find an alternative to expected utility theory which is immune to this particular argument and yet retains as much of the expected utility theory as possible. The notion of disappointment aversion offers good intuition as to why the independence axiom is so often violated. Axioms 1–4 aim to capture the notion of disappointment aversion that lead to a rather restricted class of preferences with acceptable normative properties capable of accommodating many of the experimental results. The simple characterization of these preferences suggest that they might constitute a useful step in better understanding the failure of the independence axiom.

Graduate School of Business, Stanford University, Stanford, CA 94305, U.S.A.

Manuscript received April, 1988; final revision received April, 1990.

APPENDIX

PROOF OF THEOREM 1: The proof will make use of the following two lemmas.

LEMMA 1: (I) $x > y$, $\lambda \in (0,1)$ *implies* $x \succ \lambda x + (1-\lambda)y \succ y$.

(II) $\lambda, a \in (0,1)$, $y > z$, $p = ay + (1-a)z$ *imply* (1) $\lambda x + (1-\lambda)p \succ p$ *whenever* $x \succ p$; (2) $p \succ \lambda x + (1-\lambda)p$ *whenever* $p \succ x$; (3) $x \sim \lambda x + (1-\lambda)p$ *whenever* $x \sim p$; (4) $x \succ \lambda x + (1-\lambda)p$ *whenever* $x \succ p$; (5) $\lambda x + (1-\lambda)p \succ x$ *whenever* $p \succ x$.

(III) II *above holds for arbitrary* p.

PROOF: (I) Assume the contrary; then there exists λ such that $\lambda x + (1-\lambda)y \succsim x$ or $y \succsim \lambda x + (1-\lambda)y$. Take the first case and let $\alpha = \inf\{\lambda \in [0,1] | \lambda x + (1-\lambda)y \succsim x\}$. By Axiom 2 $\alpha \in (0,1)$ and $\alpha x + (1-\alpha)y \sim x$. Then Axiom 3 implies that $\alpha^2 x + (1-\alpha^2)y \sim \alpha x + (1-\alpha)y$ if $x \succsim \alpha^2 x + (1-\alpha^2)y \succsim y$. But this follows immediately from Axiom 3. Hence $\alpha^2 x + (1-\alpha^2)y \succsim \alpha x + (1-\alpha)y \succsim x$. But $\alpha^2 < \alpha$ which contradicts the fact that α is the infimum. A similar argument establishes a contradiction for the $y \succsim \alpha x + (1-\alpha)y$ case.

(II) (1) Assume the contrary; hence there exists $\lambda, a \in (0,1)$, $p = ay + (1-a)z$, and $x \in X$ such that $x \succ p$ and $p \succsim \lambda x + (1-\lambda)p$. Then by Axiom 2 there exists $\lambda^* \in (0,1)$ such that $\lambda^* x + (1-\lambda^*)p \sim p$. Let $T = \{(y,z) | ay + (1-a)z \sim p\}$ and $\Delta = \{y^* - z^* > 0 | y^* \succsim \lambda^* x + (1-\lambda^*)(ay^* + (1-a)z^*) \succsim z^*$ for $(y^*, z^*) \in T\}$. By assumption Δ is nonempty. Let $\delta = \inf \Delta$. Then by Axiom 2 there exist y^* such that $ay^* + (1-a)(y^* - \delta) \sim p$ and then, by Axiom 3, $\lambda^* x + (1-\lambda^*)(ay^* + (1-a)(y^* - \delta)) \sim p$. Therefore, by Axiom 2, $\delta = 0$; otherwise we can find $\bar{y}$ such that $\bar{y} \succsim \lambda^* x + (1-\lambda^*)(a\bar{y} + (1-a)(\bar{y} - \bar{\delta})) \succsim \bar{y} - \delta$, $a\bar{y} + (1-a)(\bar{y} - \bar{\delta}) \sim p$, and $0 < \bar{\delta} < \delta$, which contradicts the fact that $\delta = \inf \Delta$. But if $\delta = 0$, by Axiom 2, $\lambda^* x + (1-\lambda^*)y^* \sim p$ and $y^* \sim p$ which contradicts (I) above.

(2) follows from a symmetric argument.

(3) follows from (1) above and Axiom 2.

(4) If $x > y$, then set

$$\hat{p} = \frac{\lambda}{\lambda + (1-\lambda)a} x + \frac{(1-\lambda)a}{\lambda + (1-\lambda)a} y.$$

Then by (I) and (II) (2) above we have $\hat{p} \succ tz + (1-t)\hat{p}$ whenever $t \in (0,1)$. Set $t = (1-\lambda)(1-a)$ to obtain $\hat{p} \succ tz + (1-t)\hat{p} = \lambda x + (1-\lambda)p$. But $x \succ \hat{p}$ by (I), so $x \succ \lambda x + (1-\lambda)p$.

If $y \geqslant x$ then let $\alpha^* x + (1-\alpha^*)p \sim x$ and $T = \{(y^*, z^*) | y^* \succsim \alpha^* x + (1-\alpha^*)(ay^* + (1-a)z^*) \succsim z^*$ and $ay^* + (1-a)z^* \sim p\}$. Define $\bar{y} = \inf\{y | (y,z) \in T$ for some $z \in X\}$. Observe that $\bar{y} = x$. Otherwise by Axiom 3 (and Axiom 2) $\alpha^* x + (1-\alpha^*)(a\bar{y} + (1-a)\bar{z}) \sim p$ for some $\bar{z}$. This would imply that there exists $y' < \bar{y}$ and z' such that $(y', z') \in T$, a contradiction to the fact that $\bar{y}$ is the infimum. But if $\bar{y} = x$ we have $\alpha^* x + (1-\alpha^*)(ax + (1-a)\bar{z}) \sim x$ and $\bar{z} < x$ which contradicts (I) above.

(5) follows from a symmetric argument to the one provided in (4) above.

(III) The results of (II) can be generalized to p with arbitrary supports as follows: assume that (1)–(5) hold for all p such that $|\text{supp}(p)| \leqslant n$. Then for p such that $|\text{supp}(p)| = n+1$ we can conclude by Axiom 2 that there exists $v \in (w, b)$ such that $v \sim p$. Let $\Delta = \{\lambda \in (0,1) | p \succsim \lambda x + (1-\lambda)p\}$ for some $x \succ p$. If (1) is false, then Δ is nonempty. If $\inf \Delta = 0$, then choose $(a, q, r) \in D(p)$ such that $x' \sim p$ implies $x' \in \text{supp}(r)$. Next let $(a, y, z) \sim p$ for some $y > z$. By Axiom 2 such y and z exist. Finally choose $\lambda^* \in \Delta \cap \{\lambda > 0 | y \succsim \lambda x + (1-\lambda)(ay + (1-a)z) \succsim z,\ r \in W(\lambda x + (1-\lambda)p)\}$. By Axiom 2 such λ^* exist. Then Axiom 3 implies $\lambda^* x + (1-\lambda^*)p \sim \lambda^* x + (1-\lambda^*)(ay + (1-a)z)$. But by (II) $\lambda^* x + (1-\lambda^*)(ay + (1-a)z) \succ p$, hence a contradiction.

If $\inf \Delta \neq 0$ there exists some $\hat{v} \in (w, x)$ such that $\hat{v} = CE(\alpha x + (1-\alpha)p)$ for some $\alpha \in (0, \lambda)$ and for all $\varepsilon > 0$ there exists $\alpha' \in (\alpha, \alpha + \varepsilon)$ such that $CE(\alpha' x + (1-\alpha')p) < CE(\alpha x + (1-\alpha)p)$.

Set $\hat{p} = \alpha x + (1-\alpha)p$ and observe that $\inf \hat{\Delta} = 0$ where $\hat{\Delta} = \{\hat{\lambda} \in (0,1) | \hat{p} \succsim \hat{\lambda} x + (1-\hat{\lambda})p\}$. Hence we can use the argument above to establish the desired conclusion.

(2) follows from a symmetric argument.

(3) follows from Axiom 2 and (1) above.

(4) Assume the contrary; then choose α such that $0 < \alpha < \lambda^*$, $a \in (0,1)$, $q, r \in \mathscr{L}$ and $z < x$ such that $(a, q, r) \in D(\alpha x + (1-\alpha)p)$, $\lambda^* = \inf\{\lambda | \lambda x + (1-\lambda)p \succsim x\}$ and $x' \in \text{supp}(p)$ and $x' < x$ implies $x' < V$. By Axiom 2 all of this is possible. Then set $c = (\lambda^* - \alpha)/(1-\alpha)$ and observe that

$$cx + (1-c)(\alpha x + (1-\alpha)p) = \lambda^* x + (1-\lambda^*)p \sim x, \qquad \text{by Axiom 2,}$$

$$cx + (1-c)(\alpha x + (1-\alpha)p) \sim cx + (1-c)(ax + (1-a)z), \qquad \text{by Axiom 3.}$$

Hence $x \sim (cx + (1-c)ax + (1-c)(1-a)z)$, contradicting (I) above.

(5) follows from a symmetric argument. *Q.E.D.*

Next we will define a binary relation R on $\mathscr{L}^0 = \{p \in \mathscr{L} | w \in \text{supp}(p) \text{ implies } b \notin \text{supp}(p)\}$.

sRs' iff $(\alpha, s, r) \in D, (\alpha, s', r) \in D$ implies $\alpha s + (1-\alpha)r \succsim \alpha s' + (1-\alpha)r$. We write sIs' to denote sRs' and $s'Rs$.

It is easy to see that Axiom 2, Axiom 3, and Lemma 1 (III) imply the following, stronger version of A3:

AXIOM 3: For $i = 1,2$ $(\alpha, q_i, r_i) \in D(p_i)$, $q_i \in B(\lambda x + (1-\lambda)p_i)$ and $r_i \in W(\lambda x + (1-\lambda)p_i)$, $\lambda \in [0,1]$ implies $p_1 \succsim p_2$ iff $\lambda x + (1-\lambda)p_1 \succsim \lambda x + (1-\lambda)p_2$.

But Axiom 1, Axiom 3*, and Lemma 1 (III) imply that R is preference relation on $\mathscr{L}^0$.

LEMMA 2: *sIy implies for all* $\lambda \in [0,1]$, $\lambda s + (1-\lambda)p \sim \lambda y + (1-\lambda)p$ *whenever*

$$s \in B(\lambda s + (1-\lambda)p) \cup W(\lambda s + (1-\lambda)p).$$

PROOF: It follows from Axiom 3* and Lemma 1 (II) that $s \in B(\lambda s + (1-\lambda)p)$ implies sIy iff $\lambda s + (1-\lambda)p \sim \lambda y + (1-\lambda)p$. Assume that $s \in W(\lambda s + (1-\lambda)p)$. Then by Axiom 3* and Lemma 1 (III) we have $\lambda s + (1-\lambda)p \sim \lambda y + (1-\lambda)p$ iff $\alpha s + (1-\alpha)b \sim \alpha y + (1-\alpha)b$ for all α sufficiently small. But then Axiom 4 implies that $\lambda s + (1-\lambda)p \sim \lambda y + (1-\lambda)p$ iff $\alpha s + (1-\alpha)w \sim \alpha y + (1-\alpha)w$ for all α such that $s \in B(\alpha s + (1-\alpha)w)$; that is, $\lambda s = (1-\lambda)p \sim \lambda y + (1-\lambda)p$ iff sIy. *Q.E.D.*

PROOF OF THEOREM 1: Define w_0 such that $w < w_0 < b$ and a function α_0: $[w_0, b] \to (0,1]$ such that $\alpha_0(x)x + (1-\alpha_0(x))w \sim w_0$ for all $x \in [w_0, b]$. Axiom 2 and Lemma 1 establish that α_0 is well-defined and continuous. It is easy to show, using Lemma 1, that α_0 is strictly decreasing. Next define u_0: $[w_0, b] \to [0,1]$ by

$$u_0(x) = \frac{\alpha_0(b)(1-\alpha_0(x))}{(1-\alpha_0(b))\alpha_0(x)}.$$

If follows from the continuity and strict decreasingness of α_0 that u_0 is continuous and strictly increasing. We will now show that (*) sRr iff $\Sigma_x u_0(x)s(x) \geqslant \Sigma_x u_0(x)r(x)$ for all $s, r \in \mathscr{L}(w_0)$ where $\mathscr{L}(w_0) = \{p \in \mathscr{L} | x \in \operatorname{supp}(p)$ implies $x \geqslant w_0\}$. To do this first we will show that $\alpha_0(x)[u(x)b + (1-u(x))w_0] + (1-\alpha_0(x))w \sim \alpha_0(x)x + (1-\alpha_0(x))w$. By Lemma 1 (III), part (3), $t\alpha_0(b)b + (1-t)w_0 + t(1-\alpha_0(b))w \sim w_0$. Set

$$t = \frac{1-\alpha_0(x)}{1-\alpha_0(b)}.$$

So,

$$\frac{\alpha_0(b)(1-\alpha_0(x))}{1-\alpha_0(b)}b + \frac{\alpha_0(x)-\alpha_0(b)}{1-\alpha_0(b)}w_0 + (1-\alpha_0(x))w \sim w_0;$$

hence,

$$\alpha_0(x)\left[\frac{\alpha_0(b)(1-\alpha_0(x))}{(1-\alpha_0(b))\alpha_0(x)}b + \left(1 - \frac{\alpha_0(b)(1-\alpha_0(x))}{(1-\alpha_0(b))\alpha_0(x)}\right)w_0\right] + (1-\alpha_0(x))w \sim w_0;$$

i.e.,

$$\alpha_0(x)\left[u_0(b)b + (1-u_0(x))w_0\right] + (1-\alpha_0(x))w \sim w_0.$$

Therefore

$$\alpha_0(x)\left[u_0(x)b + (1-u_0(x))w_0\right] + (1-\alpha_0(x))w \sim \alpha_0(x)x + (1-\alpha_0(x))w.$$

Then, applying Lemma 2 yields that for all s such that $\operatorname{supp}(s) = 2$,

$$(\substack{*\\*}) \qquad s \in B(\alpha s + (1-\alpha)p) \cup W(\alpha s + (1-\alpha)p)$$

implies $\alpha s + (1-\alpha)p \sim \alpha y + (1-\alpha)p$ iff $\Sigma_x u(x)s(x) = u(y)$. Hence, applying $(\substack{*\\*})$ repeatedly establishes that $(\substack{*\\*})$ holds for arbitrary s. A3 and Lemma 1 imply that $\alpha y + (1-\alpha)p \succsim \alpha y' + (1-\alpha)p$. Therefore $(\substack{*\\*})$ establishes the desired conclusion.

Choose some sequence $\{w_i\}_{i=1,2\ldots}$ such that $w < w_{i+1} < w_i$ for $i = 1,2\ldots$ and $\lim_{i\to\infty}\{w_i\} = w$. Define α_i: $[w_i, b] \to [0,1]$ by substituting w_i in place of w_0 in the definition of α_0. Furthermore let

$$u_i(x) = \frac{\alpha_i(b)}{1-\alpha_i(b)} \frac{(1-\alpha(x))}{\alpha(x)}.$$

By the argument above, u_i satisfies (*). Note that on the interval $[w_i, b]$ both u_i and u_{i+1} satisfy (*); hence $u_{i+1} = au_i + c$ by a familiar argument from expected utility theory. $u_{i+1}(b) = u_i(b) = 1$, so $a = 1 - c$. Furthermore $u_i(w_i) = 0$; hence $0 \leqslant c$ and therefore $0 < c < 1$ and $u_{i+1} = (1-c)u_i + c \geqslant u_i$. Hence $u_i(x)$ is an increasing sequence. Let $j^*(x) = \inf\{j | w_j \leqslant x\}$ for all $x \in (w, b]$. Define $u(x)$ by $u(x) = \lim_{i \geqslant j^*(x)} u_i(x)$. Clearly $u(x)$ is well-defined for all $x \in (w, b]$; furthermore for all i, u_i satisfies strict monotonicity, (*), and continuity, and hence u satisfies those properties also. Note that since $\lim_{x \to w} u(x) = 0$, u can be extended to $[w, b]$ with all of those properties by setting $u(w) = 0$.

Next define γ: $[0,1] \to [0,1]$ by $\gamma(\alpha) = u(x)$ where $\alpha b + (1-\alpha)w \sim x$. It follows from Axiom 2 that γ is well-defined and continuous. From Lemma 1 it follows that γ is strictly increasing. Furthermore $\gamma(1) = 1$, $\gamma(0) = 0$. We will show that $\gamma(\alpha) = \alpha/(1 + (1-\alpha)\beta)$ for some $\beta \in (-1, \infty)$. Let $\alpha > \hat{\alpha}$ and $\gamma(\alpha) = u(x)$, $\gamma(\hat{\alpha}) = u(y)$. By definition $\alpha b + (1-\alpha)w \sim x$. By Lemma 1 (II) part (3) $\lambda\alpha b + (1-\lambda)x + \lambda(1-\alpha)w \sim x$. Choose λ such that $(1-\lambda)u(x) = u(y)$. Then $\alpha b + (1-\alpha)w \sim \alpha[\lambda b + (1-\lambda)x] + (1-\alpha)[(1-\lambda)x + \lambda w]$. From (*) and Axiom 3 we have

$$\alpha b + (1-\alpha)w \sim \alpha[\lambda b + (1-\lambda)x] + (1-\alpha)y.$$

By taking a $\hat{\alpha}/\alpha$ convex combination of both sides with w and applying Axiom 3 we get

$$\hat{\alpha} b + (1-\hat{\alpha})w \sim \hat{\alpha}[\lambda b + (1-\lambda)x] + \hat{\alpha}\frac{(1-\alpha)}{\alpha}y + \frac{\alpha - \hat{\alpha}}{\alpha}w.$$

Let a satisfy $a + (1-a)u(y) = u(x)$; then Axiom 3 and (*) yields

$$\hat{\alpha} b + (1-\hat{\alpha})w \sim \hat{\alpha}(\lambda + (1-\lambda)a)b + \left[\hat{\alpha}(1-\lambda)(1-a) + \frac{\hat{\alpha}(1-\alpha)}{\alpha}\right]y + \frac{\alpha - \hat{\alpha}}{\alpha}w.$$

Then by Axiom 3

$$\hat{\alpha} b + (1-\hat{\alpha})w \sim \frac{\hat{\alpha}(\lambda + (1-\lambda)a)}{D}b + \left[1 - \frac{\hat{\alpha}(\lambda + (1-\lambda)a)}{D}\right]w,$$

where $D = 1 - \hat{\alpha}(1-\lambda)(1-a) - (\hat{\alpha}(1-\alpha)/\alpha)$. But this implies that

$$\hat{\alpha} = \frac{\hat{\alpha}(\lambda + (1-\lambda)a)}{D}.$$

Substituting the value for λ and a, and some simplifying, yields:

$$u(y) = \frac{u(x)\hat{\alpha}(1-\alpha)}{\alpha(1-u(x)) - \hat{\alpha}(\alpha - u(x))}.$$

Substituting $\gamma(\alpha)$ for $u(x)$ and $\gamma(\hat{\alpha})$ for $u(y)$, we obtain

$$\gamma(\hat{\alpha}) = \frac{\gamma(\alpha)\hat{\alpha}(1-\alpha)}{\alpha(1-\gamma(\alpha)) - \hat{\alpha}(\alpha - \gamma(\alpha))};$$

that is,

$$\gamma(\hat{\alpha}) = \frac{\hat{\alpha}}{\dfrac{\gamma(\alpha)(1-\alpha) + \alpha - \gamma(\alpha)}{\gamma(\alpha)(1-\alpha)} - \hat{\alpha}\dfrac{\alpha - \gamma(\alpha)}{\gamma(\alpha)(1-\alpha)}}.$$

Define

$$\beta(\alpha) = \frac{\alpha - \gamma(\alpha)}{\gamma(\alpha)(1-\alpha)};$$

then,

$$\gamma(\hat{\alpha}) = \frac{\hat{\alpha}}{1 + (1-\hat{\alpha})\beta(\alpha)}.$$

If we can show that $\beta(\alpha)$ is a constant, we are done. For $0 < \hat{\alpha} < \alpha < 1$,

$$\beta(\hat{\alpha}) = \frac{\hat{\alpha} - \gamma(\hat{\alpha})}{\gamma(\hat{\alpha})(1-\hat{\alpha})}.$$

Substituting

$$\gamma(\hat{\alpha}) = \frac{\hat{\alpha}}{1 + (1-\hat{\alpha})\beta(\alpha)}$$

into the above equation yields $\beta(\hat{\alpha}) = \beta(\alpha)$. Hence β is constant. $\beta > -1$ follows from the fact that $b \succ \alpha b + (1-\alpha)w$ for all $\alpha \in (0,1)$ (by Lemma 1). Next we will show that the function

$$V(p) = \gamma(\alpha)\sum_x u(x)q(x) + (1-\gamma(\alpha))\sum_x u(x)r(x)$$

represents $\succsim$ for $(\alpha, q, r) \in D(p)$. To show this we will prove that $(\alpha, x, z) \in D$ implies $(\alpha, x, z) \sim y$ iff $\gamma(\alpha)u(x) + (1-\gamma(\alpha))u(z) \sim u(y)$. Then, the fact that V represents $\succsim$ will follow from Lemma 2 and the observation that $ax^* + (1-a)y^* \succ ax^* + (1-a)z^*$ iff $y^* > z^*$.

$$(\alpha, x, z) \sim y \quad \text{iff} \quad \alpha(cb + (1-c)y) + (1-\alpha)(dy + (1-d)w) \sim y$$

for $c = (u(x) - u(y))/(1-u(y))$ and $d = u(z)/u(y)$ (by Lemma 2).

Hence $(\alpha, x, z) \sim y$ iff $\alpha cb + [\alpha(1-c) + (1-\alpha)d]y + (1-\alpha)(1-d)w \sim y$. Hence $(\alpha, x, z) \sim y$ iff $tb + (1-t)w \sim y$ where

$$t = \frac{\alpha c}{1 - \alpha(1-c) - (1-\alpha)d}.$$

But by construction

$$tb + (1-t)w \sim y \quad \text{iff}$$
$$\gamma(t) = u(y), \quad \text{i.e. iff}$$
$$\frac{t}{1+(1-t)\beta} = u(y), \quad \text{i.e. iff}$$
$$t = \frac{u(y)(1+\beta)}{1+u(y)\beta}.$$

Substituting for t, c, and d yields

$$(\alpha, x, z) \sim y \quad \text{iff}$$
$$\alpha = \frac{(1+\beta)(u(y)-u(z))}{u(x)-u(z)+\beta[u(y)-u(z)]} \quad \text{which holds iff}$$
$$\gamma(\alpha)u(x) + (1-\gamma(\alpha))u(z) = u(y).$$

Hence we have proven part (i) and (iii) of the theorem. That u is unique up to an affine transformation follows from the familiar argument of expected utility theory. The uniqueness of γ is obvious. *Q.E.D.*

An Algorithm for Computing V(p)

Let $\{x_1, x_2, \ldots x_k\}$ be the support for some lottery p. Without loss of generality assume $x_j > x_{j-1}$ for all $j = 2, 3, \ldots k$. If $CE(p) \notin \{x_1, x_2, \ldots x_k\}$, then there exists a unique j^* such that $\{x_1, x_2, \ldots x_{j^*}\}$ constitute all the disappointment prizes of p and $\{x_{j+1}, x_{j+2} \ldots x_k\}$ constitute all the elation prizes of p. Hence there are $k-1$ candidates for an EDD of p. These are (α_j, q_j, r_j) (for $j = 1, 2, \ldots, k-1$) where $\alpha_j = \sum_{i=j+1}^{k} p(x_i)$, $q_j(x) = 0$ for $x \notin \{x_{j+1}, x_{j+1}, \ldots x_k\}$, $q_j(x) = p_j(x)/\alpha_j$ for $x \in \{x_{j+1}, x_{j+2} \ldots x_k\}$, $r_j(x) = 0$ for $x \notin \{x_1, x_2 \ldots x_j\}$, and $r_j(x) = p(x)/(1-\alpha_j)$ for $x \in \{x_1, x_2 \ldots x_j\}$. Let $\hat{V}(\alpha_j, q_j, r_j) = \gamma(\alpha_j)Eu(q_j) + (1-\gamma(\alpha_j))Eu(r_j)$ where $\gamma(\alpha) = \alpha/(1+(1-\alpha)\beta)$. The condition $u(x_j) < \hat{V}(\alpha_j, q_j, r_j) < u(x_{j+1})$ will be satisfied if and only if $j = j^*$. For j^*, $\hat{V}(\alpha_{j^*}, q_{j^*}, r_{j^*}) = V(p)$ and $\{(\alpha_{j^*}, q_{j^*}, r_{j^*})\} = D(p)$. Hence in at most $k-1$ steps we can isolate the unique EDD of p and determine $V(p)$. If $CE(p) \in \text{supp}(p)$ then there will exist a unique j^* such that $\hat{V}(\alpha_{j^*-1}, q_{j^*-1}, r_{j^*-1}) = \hat{V}(\alpha_{j^*}, q_{j^*}, r_{j^*}) = u(x_{j^*})$. For $j < j^* - 1$, $\hat{V}(\alpha_j, q_j, r_j)$ and for $j > j^*$, $\hat{V}(\alpha_j, q_j, r_j) < u(x_j)$. Thus $D(p) = \{(\alpha, q, r) | \alpha q + (1-\alpha) r = p$ and $q(x) = 0$ if $x \notin \{x_{j^*}, x_{j^*-1}, \ldots x_k\}$, $r(x) = 0$ if $x \notin \{x_1, x_2, \ldots x_{j^*}\}\}$. Furthermore for all $(\alpha, q, r) \in D(p)$, $\hat{V}(\alpha, q, r) = u(x_{j^*}) = V(p)$. So again in at most $k-1$ steps, $V(p)$ and $D(p)$ can be determined.

Proof of Theorem 2: Let $e(p) > e(q)$, $x \succ p$, and $\beta > 0$. Then there exists $(\alpha, s, r) \in D(p)$ such that $\alpha = e(p)$. Furthermore, by Axiom 2, there exists $\bar{a} > 0$ such that $s \in B(ax + (1-a)p)$ and $r \in W(ax + (1-a)p)$ for all $a < \bar{a}$. Then for $a < \bar{a}$,

$$V(ax + (1-a)p) = \gamma(a + (1-a)\alpha) \cdot \frac{au(x) + (1-a)\alpha c}{a + (1-a)\alpha} + (1 - \gamma(a + (1-a)\alpha))d$$

where $c = \Sigma_y u(y) s(y)$ and $d = \Sigma_y u(y) r(y)$. Hence,

$$\left.\frac{dV(ax + (1-a)p)}{da}\right|_{a=0} = \gamma'(\alpha)(1-\alpha)c + \gamma(\alpha)\frac{u(x) - c}{\alpha} - \gamma'(\alpha)(1-\alpha)d.$$

Substituting $\gamma'(\alpha) = 1 + \beta/(1 + (1-\alpha)\beta)^2$ and $\gamma(\alpha) = \alpha/1 + (1-\alpha)\beta$ and rearranging terms yields

$$\begin{aligned}\left.\frac{dV(ax + (1-a)p)}{da}\right|_{a=0} &= \frac{u(x)}{1 + (1-\alpha)\beta} - \frac{\alpha c + (1-\alpha)(1+\beta)d}{(1 + (1-\alpha)\beta)^2} \\ &= \frac{u(x) - V(p)}{1 + (1-\alpha)\beta} = \frac{u(x) - V(p)}{1 + (1 - e(p))\beta}.\end{aligned}$$

Repeating the same argument for q yields

$$\left.\frac{dV(ax + (1-a)q)}{da}\right|_{a=0} = \frac{u(x) - V(q)}{1 + (1 - e(q))\beta}.$$

By assumption $V(p) = V(q)$ and $u(x) - V(p) > 0$. Hence $e(p) > e(q)$ implies

$$\left.\frac{dV(ax + (1-a)p)}{da}\right|_{a=0} > \left.\frac{dV(ax + (1-a)q)}{da}\right|_{a=0},$$

which establishes that $ax + (1-a)p \succ ax + (1-a)q$ for sufficiently small a. All remaining cases follow from symmetric arguments. *Q.E.D.*

Proof of Theorem 5: First we will prove that $\beta_2 > \beta_1$ implies $\succsim_1$ is not more risk averse than $\succsim_2$. Choose $x \in (w, b)$ and $\varepsilon > 0$. Let α solve

$$u_1(x) = \frac{\alpha u_1(x+\varepsilon) + (1-\alpha)(1+\beta_1)u_1(x-\varepsilon)}{1 + (1-\alpha)\beta_1}.$$

Then obviously $x \sim p$ where $p(x+\varepsilon) = \alpha$ and $p(x-\varepsilon) = 1-\alpha$. Using Taylor series expansion of u_1 around x we obtain

$$\begin{aligned}(1 + (1-\alpha)\beta_1)u_1(x) = \alpha&[u_1(x) + \varepsilon u_1'(x)] \\ &+ (1-\alpha)(1+\beta_1)[u_1(x) - \varepsilon u_1'(x)] + o(\varepsilon);\end{aligned}$$

hence

$$\alpha = \frac{1+\beta_1}{2+\beta_1} + o(\varepsilon).$$

Similar argument establishes that if $a'(x+\varepsilon) + (1-\alpha')(x-\varepsilon) \sim_2 x$, then

$$\alpha' = \frac{1+\beta_2}{2+\beta_2} + o(\varepsilon).$$

Hence for ε sufficiently small $\alpha' > \alpha$. That is $x \succ_2 \alpha(x+\varepsilon) + (1-\alpha)(x-\varepsilon) = p$.

Next assume that $R_1^a(x) \geqslant R_2^a(x)$ for all $x \in (w, b)$, $\beta_1 \geqslant \beta_2$ and $p \succsim_1 \bar{x}$. Then, in particular $p \sim_1 x$ for some $x > \bar{x}$. Since u_1 and u_2 are unique only up to affine transformations we can, without loss of generality, assume that $u_1(x) = u_2(x)$ and $u_1'(x) = u_2'(x)$. Then note that $u_1(y) \leqslant u_2(y)$ for all $y \in [w, b]$. Let $(\alpha, q, r) \in D(p, \succsim_1)$. Then we have

$$\gamma_1(\alpha) \sum_y u_1(y) q(y) + (1 - \gamma_1(\alpha)) \sum_y u_1(y) r(y) = u_1(x).$$

But $u_1(x) = u_2(x)$, $\gamma_1(\alpha) \leqslant \gamma_2(\alpha)$ and $u_1(y) \leqslant u_2(y)$; hence

$$\gamma_2(\alpha) \sum_y u_2(y) q(y) + (1 - \gamma_2(\alpha)) \sum_y u_2(y) r(y) \geqslant u_2(x).$$

Therefore

$$u_2(x) \leqslant \sum_y u_2(y) p(y) + \beta_2 \sum_{y \leqslant x} (u_2(y) - u_2(x)) p(y).$$

Since $V_2(p) = \Sigma_y u_2(y) p(y) + \beta_2 \Sigma_{y \leqslant p} (u_2(y) - V_2(p)) p(y)$, $V_2(p) \geqslant u_2(x)$, which establishes that $p \succsim_2 x \succsim_2 \bar{x}$ and proves that $\succsim_1$ is more averse than $\succsim_2$. *Q.E.D.*

REFERENCES

Allais, M. (1979): "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School," in *Expected Utility Hypothesis and the Allais Paradox*, ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Co.

Bell, D. (1982): "Regret in Decision Making Under Uncertainty," *Operations Research*, 30, 961–981.

——— (1985): "Disappointment in Decision Making Under Uncertainty," *Operations Research*, 33, 1–27.

Chew, S. H., and K. MacCrimmon (1979): "Alpha-Nu Choice Theory: A Generalization of Expected Utility Theory," University of British Columbia, Faculty of Commerce and Business Administration, Working Paper No. 686.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences Under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 40, 309–318.

Edwards, W. (1953): "Probability-Preferences in Gambling," *American Journal of Psychology*, 66, 56–67.

Fishburn, P. C. (1983): "Transitive Measurable Utility," *Journal of Economic Theory*, 31, 293–317.

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 Rational Choice Under Uncertainty," *Economic Journal*, 92, 805–824.

——— (1986): "Disappointment and Dynamic Consistency in Choice Under Uncertainty," *Review of Economic Studies*, 53, 271–282.

Machina, M. J. (1982): "Expected Utility Without the Independence Axiom," *Econometrica*, 50, 273–323.

——— (1984): "Temporal Risk and the Nature of Induced Preferences," *Journal of Economic Theory*, 33, 199–231.

——— (1987): "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives*, 1, 121–154.

Neilson, W. S. (1989): "Behavior in the Probability Triangle," mimeo.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 225–243.

Rubinstein, A. (1988): "Similarity and Decision-Making Under Risk," *Journal of Economic Theory*, 46, 145–153.

Savage, L. J. (1972): *Foundations of Statistics*. New York: Dover Publications.

Tversky, A. (1967): "Utility Theory and Additivity Analysis of Risky Choices," *Journal of Experimental Psychology*, 75, 27–36.

Yaari, M. E. (1987): "The Dual Theory of Choice Under Risk," *Econometrica*, 55, 95–115.