THE DUAL THEORY OF CHOICE UNDER RISK

By Menahem E. Yaari[1]

This paper investigates the consequences of the following modification of expected utility theory: Instead of requiring independence with respect to probability mixtures of risky prospects, require independence with respect to direct mixing of payments of risky prospects. A new theory of choice under risk—a so-called dual theory—is obtained. Within this new theory, the following questions are considered: (i) numerical representation of preferences; (ii) properties of the utility function; (iii) the possibility for resolving the "paradoxes" of expected utility theory; (iv) the characterization of risk aversion; (v) comparative statics. The paper ends with a discussion of other non-expected-utility theories proposed recently.

KEYWORDS: Risk, uncertainty, utility, duality.

1. INTRODUCTION

IN THIS ESSAY, a new theory of choice under risk is being proposed. It is a theory which, in a sense that will become clear, is *dual* to expected utility theory, hence the title "dual theory." Risky prospects are evaluated in this theory by a cardinal numerical scale which resembles an expected utility, except that the roles of payments and probabilities are reversed. This theme—the reversal of the roles of probabilities and payments—will recur throughout the paper. I should emphasize that playing games, with probabilities masquerading as payments and payments masquerading as probabilities, is not my object. Rather, I hope to convince the reader that the dual theory has intrinsic economic significance and that, in some areas, its predictions are superior to those of expected utility theory (while in other areas the reverse will be the case).

Two reasons have prompted me to look for an alternative to expected utility theory. The first reason is methodological: In expected utility theory, the agent's attitude towards risk and the agent's attitude towards wealth are forever bonded together. At the level of fundamental principles, risk aversion and diminishing marginal utility of wealth, which are synonymous under expected utility theory, are horses of different colors. The former expresses an attitute towards risk (increased uncertainty hurts) while the latter expresses an attitute towards wealth (the loss of a sheep hurts more when the agent is poor than when the agent is rich). A question arises, therefore, as to whether these two notions can be kept separate from each other in a full-fledged theory of cardinal utility. The dual theory will have this property.

The second reason that leads me to look for an alternative to expected utility theory is empirical: Behavior patterns which are systematic, yet inconsistent with expected utility theory, have often been observed. (Two prominent references, among many others, are Allais (1953) and Kahneman–Tversky (1979).) So deeply

rooted is our commitment to expected utility, that we tend to regard such behavior patterns as "paradoxical", perhaps even as "irrational." The dual theory, it turns out, rationalizes many of the "paradoxes" of expected utility theory. Obviously, the dual theory will have its own "paradoxes", many of which turn out to become rationalized under expected utility. Roughly speaking, we find each theory resolving "paradoxes" in the other theory.

The dual theory has the property that utility is linear in wealth, in the sense that applying an affine transformation to the payment levels of two gambles always leaves the direction of preference between them unchanged. (Under expected utility, this is true only when the agent is risk neutral.) In order to forestall needless arguments, let me come clean right away and say that I do not consider linearity in payments an empirically viable proposition. Behavior which is inconsistent with such linearity is probably often observed. However, such evidence should be viewed in proper perspective: Behavior which is inconsistent with linearity in *probabilities*—a vital component of expected utility theory—is also often observed. I shall return to this matter in Section 4, below.

In studying the behavior of *firms*, linearity in payments may in fact be an appealing feature. Under the dual theory, maximization of a linear function of profits can be entertained simultaneously with risk aversion. How often has the desire to retain profit maximization led to contrived arguments about firms' risk neutrality?

The most general way of using cardinal utility to treat choice under risk is one where preferences are represented by a measure which is defined on appropriate subsets of the payment-probability plane. Both expected utility and the dual theory are special cases of this approach, with the measure representing preferences being a *product* measure, factorizable into two marginal measures. In expected utility, the marginal measure along the probability axis is Lebesgue measure, and in the dual theory, the marginal measure along the payment axis is Lebesgue measure. Dropping the condition that one of the marginal measures be Lebesgue produces a theory which generalizes both expected utility and dual theory. A special version of this generalized theory has been proposed recently by Quiggin (1982), in a paper which studies the perception of risk from a cognitive point of view. The case of preferences being represented by a nonfactorizable measure has, to the best of my knowledge, not yet been studied.

An extension of the dual theory to the multivariate case exists, and is explored in a separate paper (Yaari (1986)). It is interesting to note that, in the multivariate version of the dual theory, linearity in payments ceases to be an issue.

2. A REPRESENTATION THEOREM

Let V be the set of all random variables defined on some given probability space, with values in the unit interval. I shall assume that the underlying probability space is "rich", in the sense that all distributions with supports contained in the unit interval can be generated from elements of V. For each $v \in V$, define the

decumulative distribution function (DDF for short) of v, to be denoted G_v, by

$$G_v(t) = \Pr\{v > t\}, \quad 0 \leqslant t \leqslant 1.$$

G_v is always nonincreasing, right-continuous, and satisfies $G_v(1) = 0$. For all $v \in V$, the following convenient relationship holds:

$$(1) \qquad \int_0^1 G_v(t)\, dt = Ev,$$

where Ev stands for the expected value of v.

The values of the random variables in V will be interpreted as payments, denominated in some monetary unit. This makes each $v \in V$ interpretable as a gamble or a lottery which a decision maker might consider holding. Restricting the values of random variables in V to the unit interval can be interpreted, via the choice of a suitable measurement scale, to mean that (i) no gambles can be considered which involve a possible loss exceeding the decision maker's total wealth, and (ii) no gambles exist which offer prizes exceeding some predetermined large number.

A preference relation $\gtrsim$ is assumed to be defined on V. Let the symbols $>$ and $\sim$ stand for strict preference and indifference, respectively. The following axiom suggests itself:

AXIOM A1—*Neutrality*: *Let u and v belong to V, with respective DDF's G_u and G_v. If $G_u = G_v$, then $u \sim v$.*

This axiom restricts attention to preferences which are not state-dependent. It implies, in particular, that preference among DDF's can be defined in an unambiguous manner. Specifically, we may construct a preference relation $(\gtrsim)$ among DDF's by writing $G(\gtrsim)H$ if, and only if, there exist two elements, u and v, of V such that $G_u = G$, $G_v = H$, and $u \gtrsim v$. Under Axiom A1, the assertions $u \gtrsim v$ and $G_u(\gtrsim)G_v$ are equivalent. Our assumptions on V imply that the domain of the relation $(\gtrsim)$ is the set of all DDF's with supports contained in the unit interval. More precisely, let a family of functions Γ be defined by

$$\Gamma = \{G:[0,1] \to [0,1] \mid G \text{ is nonincreasing, right-continuous and}$$
$$\text{satisfies } G(1) = 0\}.$$

Then, the assertion $G(\gtrsim)H$ is meaningful for every pair of functions, G and H, in Γ.

In order to reduce cumbersome notation, and with the reader's indulgence, I shall henceforth use the symbol $\gtrsim$ both for preference among random variables and for preference among DDF's. (In other words, the parentheses in $(\gtrsim)$ will henceforth be dropped.)

We can now proceed to the remaining axioms:

AXIOM A2—*Complete weak order*: $\gtrsim$ *is reflexive, transitive, and connected.*

AXIOM A3—*Continuity* (with respect to L_1-convergence): *Let G, G', H, H', belong to Γ; assume that $G > G'$. Then, there exists an $\varepsilon > 0$ such that $\|G - H\| < \varepsilon$ and $\|G' - H'\| < \varepsilon$ imply $H > H'$, where $\|\ \|$ is the L_1-norm, i.e., $\|m\| = \int |m(t)|\, dt$.*

It should be noted that the continuity assumed in A3 is stronger than that required for the development of standard expected utility theory. (The reason for this will become apparent shortly.)

AXIOM A4—*Monotonicity* (with respect to first-order stochastic dominance): *If $G_u(t) \geqslant G_v(t)$ for all t, $0 \leqslant t \leqslant 1$, then $G_u \gtrsim G_v$.*

With Axioms A1-A4 in hand, one can proceed to write down an appropriate independence axiom and obtain the result that preferences are representable by expected utility comparisons. Specifically consider:

AXIOM A5EU—*Independence*: *If G, G', and H belong to Γ and α is a real number satisfying $0 \leqslant \alpha \leqslant 1$, then $G \gtrsim G'$ implies $\alpha G + (1-\alpha)H \gtrsim \alpha G' + (1-\alpha)H$.*

For the record, I shall now state the expected utility theorem. Before doing so, let me introduce the following notation: If x and p both lie in the unit interval, then $[x; p]$ will stand for a random variable that takes the values x and 0 with probabilities p and $1-p$, respectively.

THEOREM 0: *A preference relation $\gtrsim$ satisfies Axioms A1–A4 and A5EU if, and only if, there exists a continuous and nondecreasing real function ϕ, defined on the unit interval, such that, for all u and v belonging to V,*

$$(2) \qquad u \gtrsim v \Leftrightarrow E\phi(u) \geqslant E\phi(v).$$

Moreover, the function ϕ, which is unique up to a positive affine transformation, can be selected in such a way that, for all t satisfying $0 \leqslant t \leqslant 1$, $\phi(t)$ solves the preference equation

$$(3) \qquad [1; \phi(t)] \sim [t; 1].$$

PROOF: See, e.g., Fishburn (1982, Theorem 3, p. 28). It follows readily from Axioms A2-A4 and A5EU that the premises of Fishburn's theorem hold, with the unit interval acting as the set of consequences and with distributions representing probability measures. The conclusion, therefore, is that a function ϕ satisfying (2) exists, uniquely up to a positive affine transformation and, moreover, that equation (3) provides the construction of ϕ. That ϕ is continuous and nondecreasing follows directly from A3 and A4, respectively, in conjunction with (3). Finally, the fact that the converse also holds is established by straightforward verification.
                                                                                        Q.E.D.

The dual theory of choice under risk is obtained when the independence axiom of expected utility theory (Axiom A5EU) is taken and, so to speak, "laid on its side." Instead of independence being postulated for convex combinations which are formed along the probability axis, it will now be postulated for convex combinations which are formed along the payment axis. The best way to do this is to consider appropriately defined *inverses* of distribution functions.

Let $G \in \Gamma$, so that G is the DDF of some $v \in V$. Now define a set-valued function, $\hat{G}$, by writing, for $0 \leq t \leq 1$,

$$\hat{G}(t) = \{x \mid G(t) \leq x \leq G(t-)\}$$

where $G(t-) = \lim_{s \to t, s < t} G(s)$ for $t > 0$, and $G(0-) = 1$. $\hat{G}$ is simply the set-valued function which "fills up" the range of G, to make it coincide with the unit interval. The values of $\hat{G}$ are closed and for each p, $0 \leq p \leq 1$, there exists some t such that $p \in \hat{G}(t)$. Using $\hat{G}$, we may now proceed to define the (generalized) *inverse* of G, to be denoted G^{-1}, by writing

(4) $G^{-1}(p) = \min \{t \mid p \in \hat{G}(t)\}.$

Note that G^{-1}, like G, belongs to Γ and that, for all $G \in \Gamma$, $(G^{-1})^{-1} = G$. Furthermore, if G and H belong to Γ and $\| \ \|$ stands for L_1-norm, then $\|G - H\| = \|G^{-1} - H^{-1}\|$. Of course, if G is invertible, then G^{-1} is just the usual inverse function of G.

A mixture operation for DDF's may now be defined as follows: If G and H belong to Γ and if $0 \leq \alpha \leq 1$, then $\alpha G \boxplus (1 - \alpha)H$ is the member of Γ given by

(5) $\alpha G \boxplus (1 - \alpha)H = (\alpha G^{-1} + (1 - \alpha)H^{-1})^{-1}.$

If $J = \alpha G \boxplus (1 - \alpha)H$, for some $0 \leq \alpha \leq 1$, then I shall say that J is a *harmonic convex combination* of G and H. With the operation $\boxplus$, the set Γ of all DDF's becomes a *mixture space*, in the sense of Herstein and Milnor (1953).

Returning to the preference relation $\succsim$, we are now in a position to state the axiom that gives rise to the dual theory of choice under risk:

AXIOM A5—*Dual Independence*: *If G, G' and H belong to Γ and α is a real number satisfying $0 \leq \alpha \leq 1$, then $G \succsim G'$ implies $\alpha G \boxplus (1 - \alpha)H \succsim \alpha G' \boxplus (1 - \alpha)H$.*

The economic significance of this axiom will be discussed in Section 3, below. The following representation theorem is now available:

THEOREM 1: *A preference relation $\succsim$ satisfies Axioms A1–A5 if, and only if, there exists a continuous and nondecreasing real function f, defined on the unit interval, such that, for all u and v belonging to V,*

(6) $u \succsim v \iff \int_0^1 f(G_u(t)) \, dt \geq \int_0^1 f(G_v(t)) \, dt.$

Moreover, the function f, which is unique up to a positive affine transformation, can be selected in such a way that, for all p satisfying $0 \leq p \leq 1$, $f(p)$ solves the preference equation

(7) $[1; p] \sim [f(p); 1].$

PROOF: Define a binary relation $\succsim^*$ on the family Γ of DDF's, as follows:

$$G \succsim^* H \quad \text{if, and only if,} \quad G^{-1} \succsim H^{-1},$$

for all G and H in Γ. Clearly, if u and v are random variables in V, then

$$u \succsim v \quad \Leftrightarrow \quad G_u^{-1} \succsim^* G_v^{-1}.$$

Checking Axioms A2–A4, we find that they hold for $\succsim$ if, and only if, they hold for $\succsim^*$. Furthermore, $\succsim$ satisfies A5 if, and only if, $\succsim^*$ satisfies A5EU. Hence, from Theorem 0, it follows that $\succsim$ satiefies A1–A5 if, and only if, $\succsim^*$ has the appropriate expected utility representation. In other words, $\succsim$ satisfies A1–A5 if, and only if, there exists a continuous and nondecreasing function f, defined on the unit interval, such that

$$u \succsim v \quad \Leftrightarrow \quad -\int_0^1 f(p)\, dG_u^{-1}(p) \geq -\int_0^1 f(p)\, dG_v^{-1}(p)$$

is true for all u and v in V. Let G be any member of Γ. Then, the equation

$$-\int_0^1 f(p)\, dG^{-1}(p) = \int_0^1 f(G(t))\, dt$$

holds, by introducing the change of variable $p = G(t)$, and this proves the first part of the theorem. Now, applying the second part of Theorem 0 to $\succsim^*$, we find that f can be selected so as to satisfy the preference equation

(8) $\qquad G_{[1; f(p)]} \sim^* G_{[p; 1]}$

for $0 \leq p \leq 1$. Note, however, that if G is the DDF of $[x; p]$ then G^{-1} is the DDF of $[p, x]$. Therefore, a rewriting of (8) in terms of the original preference relation, $\succsim$, produces (7). This completes the proof of the theorem.   Q.E.D.

Let v belong to V, with DDF G_v, and let $U(v)$ be defined by

(9) $\qquad U(v) = \int f(G_v(t))\, dt,$

with f defined in (7). Theorem 1 tells us that the function U is a *utility* on V, when preferences satisfy A1–A5. The hypothesis of the Dual Theory is that agents will choose among random variables so as to maximize U. This is in analogy (and in contrast) with the hypothesis of expected utility theory, which is that agents choose among random variables so as to maximize the function W, given by

(10) $\qquad W(v) = E\phi(v) = -\int_0^1 \phi(t)\, dG_v(t),$

with ϕ defined in (3). Note, incidentally, that (10) can be rewritten in a manner that makes the analogy with (9) stand out more clearly. Specifically we have

$$W(v) = \int_0^1 \phi(G_v^{-1}(p))\, dp.$$

Let $\succsim$ satisfy A1–A5, and let f be defined by (7). The phrase "f represents $\succsim$" will be used as convenient shorthand for the much longer phrase "the function U, derived from f in (9), is a utility representing $\succsim$."

The utility U of the dual theory has two noteworthy properties: First, U assigns to each random variable its *certainty equivalent*. In other words, if v belongs to V, then $U(v)$ is equal to that sum of money which, when received with certainty, is considered by the agent equally as good as v. The second important property of U is *linearity in payments*: When the values of a random variable are subjected to some fixed positive affine transformation, the corresponding value of U undergoes the same transformation. The following propositions provide a precise statement of these properties.

PROPOSITION 1: *Under Axioms A1–A5, the relationship*

(11) $\qquad v \sim [U(v); 1]$

holds for every $v \in V$.

PROOF: It follows from (9) that $U([x; 1]) = x$ for all x, $0 \leqslant x \leqslant 1$. In particular, $U([U(v); 1]) = U(v)$ and, by Theorem 1, $[U(v); 1] \sim v$, as was to be shown.
$$Q.E.D.$$

REMARK: In expected utility theory, the following *dual* to Proposition 1 exists: Let $\succsim$ satisfy A1–A4 and A5EU, and let ϕ and W be defined by (3) and (10), respectively. Then, $v \sim [1; W(v)]$ is true for every $v \in V$.

PROPOSITION 2: *Let v belong to V and let a and b be two real numbers, with $a > 0$. Define a function $av + b$ by writing $(av + b)(s) = av(s) + b$ for each state-of-nature s, and assume that $0 \leqslant av(s) + b \leqslant 1$ for all s. Then, $U(av + b) = aU(v) + b$.*

PROOF: Let G_v and G_{av+b} be the DDF's of v and $av + b$, respectively. Note that, for every t, $0 \leqslant t \leqslant 1$, we have

$$G_{av+b}(t) = \begin{cases} 1 & \text{for } 0 \leqslant t < av_0 + b, \\ G_v\left(\dfrac{t - b}{a}\right) & \text{for } t \geqslant av_0 + b, \end{cases}$$

where v_0 is the infimum of the range of v. Hence,

$$U(av + b) = av_0 + b + \int_{av_0 + b}^{1} f(G_{av+b})(t)) \, dt$$

$$= av_0 + b + \int_{av_0 + b}^{1} f\left(G_v\left(\frac{t - b}{a}\right)\right) dt.$$

Introducing the change of variable $s = (t - b)/a$, we get

$$U(av + b) = a\left[v_0 + \int_{v_0}^{1} f(G_v(s))\, ds \right] + b$$

$$= aU(v) + b,$$

as was to be shown.   Q.E.D.

COROLLARY: *If the preference relation $\gtrsim$ satisfies A1–A5, then, for all u and v belonging to V, we have*

$$u \gtrsim v \quad \Leftrightarrow \quad au + b \gtrsim av + b,$$

provided $a > 0$ and provided $au + b$ and $av + b$ both belong to V. In words, under A1–A5, agents always display constant absolute risk aversion as well as constant relative risk aversion.

PROOF: Apply Proposition 2.   Q.E.D.

Note that under expected utility theory, an agent with constant absolute risk aversion as well as constant relative risk aversion must be *risk-neutral*, i.e., this agent's preferences always rank random variables by comparing their means. Under the dual theory, we have linearity (in the sense of Proposition 2 and its Corollary) without risk neutrality being implied in any way. Indeed, let us see how risk neutrality is characterized under the dual theory. It follows from (6), in conjunction with (1), that under Axioms A1–A5, the agent's preference relation $\gtrsim$ ranks random variables by comparing their means if, and only if, the function f representing $\gtrsim$ coincides with the identity, i.e., $f(p) = p$ for $0 \leq p \leq 1$. In other words, risk neutrality is characterized in the dual theory by the function f in (7) being the identity. But there is nothing in Theorem 1 to *force* f to coincide with the identity: Any continuous and nondecreasing function f, satisfying $f(0) = 0$ and $f(1) = 1$ can be obtained in (7), for some preference relation $\gtrsim$ satisfying A1–A5. In the dual theory, the agent's attitude towards wealth—restricted as it is—does not prejudice the agent's attitude towards risk.

It is interesting to compare the construction of the function f in the dual theory with the construction of the von Neumann–Morgenstern utility ϕ in expected utility theory. Consider the preference equation

(11) $[1; p] \sim [t; 1]$.

We know, from (7) and (3), that $f(p)$ is the value of t that solves (11), while $\phi(t)$ is the value of p that solves (11). It follows, therefore, that $f = \phi^{-1}$. Of course, when writing $f = \phi^{-1}$, we should not lose sight of the fact that only *one* of the two functions, ϕ and f, can be relevant to the characterization of the agent's overall behavior in risky situations.

3. THE MEANING OF DUAL INDEPENDENCE

In the foregoing section, dual independence (Axiom A5) appeared without an economic interpretation. My aim now is to re-state A5 in a way that will make its economic content clear.

Consider once again the set V of random variables, on which preferences are defined, and let (S, Σ, P) be the underlying probability space. (V, then, is the set of all Σ-measurable functions on S, with values in the unit interval.)

DEFINITION: Let u and v belong to V. We say that u and v are *comonotonic* if, and only if, for every s and s' in S, the inequality

$$(u(s) - u(s'))(v(s) - v(s')) \geq 0$$

is true.

This definition makes it possible to state the following axiom, directly on preference among random variables (without going to distributions):

AXIOM A5*—*Direct Dual Independence*: *Let u, v, and w belong to V and assume that u, v, and w are pairwise comonotonic. Then, for every real number α satisfying $0 \leq \alpha \leq 1$, $u \gtrsim v$ implies $\alpha u + (1 - \alpha)w \gtrsim \alpha v + (1 - \alpha)w$.*

Note that here we are dealing with ordinary convex combinations of real functions and that $\alpha u + (1 - \alpha)w$ is *not* a probability mixture of u and w.

It turns out that A5 and A5* are, in fact, equivalent:

PROPOSITION 3: *Let $\gtrsim$ be a preference relation on V, satisfying Axiom A1. Then, $\gtrsim$ satisfies Axiom A5* if, and only if, the corresponding preference relation among DDF's (also denoted $\gtrsim$) satisfies Axiom A5.*

PROOF: Under Axiom A1, the underlying probability space can be chosen to suit our convenience, as long as all DDF's in Γ can be generated. Accordingly, let (S, Σ, P) consist of the unit interval, the Borel sets, and Lebesgue measure. Now let u, v, and w be pairwise comonotonic and suppose that A5 holds. We must show that $u \gtrsim v$ implies $\alpha u + (1 - \alpha)w \gtrsim \alpha v + (1 - \alpha)w$, where $0 \leq \alpha \leq 1$. By comonotonicity, there exists a measure-preserving transformation, mapping the unit interval onto itself which, when composed with any of the random variables u, v, and w, rearranges it in nonincreasing order, without affecting its distribution. Thus, without loss of generality, we may assume not only that u, v, and w are pairwise comonotonic, but that each one of them is a nonincreasing function on the unit interval. Moreover, having selected Lebesgue measure for the underlying probability measure, we find that the right-continuous *inverse* of u, u^{-1}, is precisely the DDF G_u of u, and similarly for v and w. Therefore, the assertion that $G_u \gtrsim G_v$ implies $\alpha G_v \boxplus (1 - \alpha)G_w \gtrsim \alpha G_v \boxplus (1 - \alpha)G_w$ in A5 reduces precisely to $u \gtrsim v$

implying $\alpha u + (1-\alpha)w \gtrsim \alpha v + (1-\alpha)w$. Conversely, let G, G', and H belong to Γ and assume that A5* holds. We must show that

$$G \gtrsim G' \quad \text{implies} \quad \alpha G \boxplus (1-\alpha)H \gtrsim \alpha G' \boxplus (1-\alpha)H, \quad \text{for} \quad 0 \le \alpha \le 1.$$

Defining u, v, and w to be the inverses of G, G', and H, respectively, we find that u, v, and w are pairwise comonotonic so, by A5*, $u \gtrsim v$ implies $\alpha u + (1-\alpha)w \gtrsim \alpha v + (1-\alpha)w$. This assertion, when written in terms of preference among DDF's, gives the desired result, and the proof is complete.   Q.E.D.

The foregoing proposition makes it clear that the economic interpretation of dual independence lies in the intuitive meaning of comonotonicity. Recall that comonotonicity is a distribution-free property, in the sense that it is invariant under changes in the underlying probability measure. It is, in fact, an analogue of perfect correlation for this distribution-free setting. When two random variables are comonotonic, then it can be said that neither of them is a *hedge* against the other. The variability of one is never tempered by counter-variability of the other. (A discussion of this no-hedge condition appeared in Yaari (1969), where comonotonic random variables were referred to as "bets on the same event.") Suppose, for example, that u and v are random variables such that $u \gtrsim v$. Would this preference be retained when both u and v are mixed, half and half, with some third random variable, say w? (Recall that we are not dealing here with a probability mixture, but rather with a pointwise averaging of the values of the two random variables.) If the agent whose preferences are being discussed is risk averse, and w is a hedge against v but not against u, then this agent might well have reason to reverse the direction of preference: i.e., the assertions $u \gtrsim v$ and $\frac{1}{2}v + \frac{1}{2}w > \frac{1}{2}u + \frac{1}{2}w$ will both be true. Similarly, if the agent for whom $u \gtrsim v$ is true is risk seeking, and w is a hedge against u but not against v, then, once again, there will be reason for the agent to reverse the direction of preference as above. Thus, the demand that $u \gtrsim v$ should imply $\alpha u + (1-\alpha)w \gtrsim \alpha v + (1-\alpha)w$ seems to be justified only in the case where w is neither a hedge against u nor a hedge against v. This is precisely what dual independence says. Actually, dual independence is *weaker*, in that the conclusion is only required to hold when u and v themselves are not a hedge against each other. This further weakening becomes important when the agent's initial wealth is allowed to vary. In this paper, however, variations in initial wealth will not be considered.

We see, in summary, that dual independence requires the direction of preference to be retained under mixing of payments, provided hedging is not involved. Two comments are in order at this point.

(a) Comonotonicity, i.e. the no-hedge condition, is sensitive to random variables being changed on sets of probability zero. In a recent paper, Röell (1985) has adopted a weaker notion of comonotonicity, defined with joint distributions, which is invariant under changes occurring on sets of probability zero. Röell then uses this alternative definition in an axiom like A5*.

(b) Axiom A5* is, of course, quite strong, and one could think of weakening it in the following way: Suppose that u, v, and w are pairwise comonotonic and

that $u \succsim v$. Then, $\alpha u + (1-\alpha)w \succsim \alpha v + (1-\alpha)w$ should be required to hold only if w is *relatively* a better hedge against u than against v. (Presumably, one could try to define the relation "relatively a better hedge..." using correlation coefficients.) This condition would weaken the notion of independence, in comparison with A5*, while simultaneously restricting the analysis to the case of a risk averse agent. Exploring the resulting theory would be, it seems to me, an interesting task. Here, A5* will be maintained, with risk aversion to be treated separately (see Section 5, below).

4. PARADOXES AND DUAL PARADOXES

Behavior which is inconsistent with expected utility theory has been observed systematically, and often such behavior has been branded "paradoxical." As it turns out, behavior which is "paradoxical" under expected utility theory is, in many cases, entirely consistent with the dual theory. This does not mean, however, that the dual theory is "paradox-free." We find, on the contrary, that for each "paradox" of expected utility theory, one can usually construct a "dual paradox" of the dual theory, by interchanging the roles of payments and probabilities. Under these "dual paradoxes," reasonable behavior—and probably easily observable behavior—is found to be inconsistent with the dual theory and to be entirely in keeping with expected utility theory. I would like to illustrate this, using a couple of prominent examples.

A famous "paradox" of expected utility theory is the so-called *common ratio effect*: Dividing all the probabilities by some common divisor reverses the direction of preference. Kahnemen and Tversky (1979), for example, have found that a great majority of subjects prefer [0.3; 1] over [0.4; 0.8] but that an equally large majority prefer [0.4; 0.2] over [0.3; 0.25]. (The symbol [x; p], it will be recalled, stands for a random variable which takes the values x and 0 with probabilities p and $1-p$, respectively. Here, payments are measured in units of $10,000, so that [0.3; 1] is the gamble that yields $3000 with certainty, etc.) This pattern, which is obviously inconsistent with expected utility theory, is entirely in keeping with the dual theory. Specifically, with the utility U defined in (9), we find that $U([0.3; 1]) = 0.3$, $U([0.4; 0.8]) = (0.4)f(0.8)$, $U([0.3; 0.25]) = (0.3)f(0.25)$ and $U([0.4; 0.2]) = (0.4)f(0.2)$, and these numbers will support the preference pattern $[0.3; 1] > [0.4; 0.8]$ and $[0.4; 0.2] > [0.3; 0.25]$ if

$$f(0.8) < \frac{3}{4} < \frac{f(0.2)}{f(0.25)}.$$

This inequality is satisfied, for example, when f is of the form $f(p) = p/(2-p)$, for $0 \leq p \leq 1$. (This f is in fact *risk averse*, as we shall see in Section 5.)

Now, to get a "dual paradox" for the common ratio effect, we must look for a case where dividing all the *payments* by some common factor would lead to preference reversal. In order to obtain such behavior, which would clearly be inimical to the dual theory, we would have to gather a group of subjects, pay each one of them $5 per hour for "Participating in an Interesting Experiment on

Decision Making" and proceed to elicit from these subjects a pattern of responses which is inconsistent with constant relative risk aversion. Alas, I cannot claim to have done this. But happily I join the critics of the dual theory in saying that such "deviant" behavior is, no doubt, quite common.

A similar state of affairs exists with Allais' celebrated paradox (Allais (1953)). On the one hand, the non-expected-utility preference pattern, which Allais had found prevalent, turns out to be consistent with the dual theory. On the other hand, examples can be found which resemble Allais' gambles—with the roles of payments and probabilities reversed—where one would expect to observe behavior which is inconsistent with the dual theory while being consistent with expected utility theory. I shall omit the details.

Proceeding now to the theory of *income distribution*, we find yet another "paradox": Newbery (1979) has shown that there does *not* exist a von Neumann–Morgenstern utility whose expected value ranks distributions (with a fixed mean) in the same order as their Gini coefficients of equality. (The Gini coefficient of equality is defined as twice the area under the Lorenz Curve.) Under expected utility theory, it is "irrational" to evaluate income distributions according to the Gini coefficient. Given the frequency with which the Gini has actually been used for comparing income distributions, Newbery's finding is surely as much a paradox of expected utility theory as the common ratio effect or Allais' gambles. Under the dual theory, the paradox disappears. In fact, if we let the function f of Theorem 1 be given by $f(p) = p^2$ for $0 \leq p \leq 1$, we find that, for DDF's with a fixed integral, the ordering induced by the integral $\int f(G(t)) \, dt$ is precisely the Gini equality ordering. Indeed, for mean-normalized distributions, the quantity $\int (G(t))^2 \, dt$ is precisely the Gini equality coefficient for G. This result is due to Dorfman (1979). Now, as might be expected, it is easy to think of a "paradox" that would be the *dual* of the foregoing: Just as Gini-type measures of equality (or of inequality) are not rationalizable under expected utility theory, so Atkinson's (1970) measures of equality (or of inequality) are not rationalizable under the dual theory.

5. RISK AVERSION

How would risk aversion be characterized under the dual theory? The following heuristic argument is meant to sound suggestive: Under expected utility theory, preferences are represented by a von Neumann–Morgenstern utility, ϕ. Under the dual theory, preferences are represented by a function f, as per Theorem 1. The construction of ϕ and f in the two theories (equations (3) and (7)) implies that $f = \phi^{-1}$. Since the concavity of ϕ is equivalent to the convexity of ϕ^{-1} and since, under expected utility, the concavity of ϕ characterizes risk aversion, we should expect the *convexity* of f to characterize risk aversion under the dual theory. Showing that this conclusion is indeed correct—even though f and ϕ belong to different theories—is my task in the present section.

Letting $\succeq$ be a preference relation on V, as before, we say that $\succeq$ is risk averse if $v = u + \text{noise}$ implies $u \succeq v$: Adding noise can never be $\succeq$-improving. Drawing

on the work of Blackwell (1950) and Rothschild-Stiglitz (1970), we obtain the following definition:

DEFINITION: Let u and v belong to V, with DDF's G_u and G_v, respectively, and consider the inequality

(12) $$\int_0^T G_u(t)\, dt \geq \int_0^T G_v(t)\, dt.$$

A preference relation $\gtrsim$ on V is said to be *risk averse* if $u \gtrsim v$ whenever (12) holds for all T satisfying $0 \leq T \leq 1$, with equality for $T = 1$.

The following theorem is now available:

THEOREM 2: *Consider the class of preference relations on V satisfying Axioms A1–A5. A preference relation $\gtrsim$ in this class is risk averse if, and only if, the function f representing $\gtrsim$ (see Theorem 1) is convex.*

PROOF: Let $\gtrsim$ satisfy A1–A5, and assume that $\gtrsim$ is risk averse. Take five real numbers, x, y, p, q, r, such that $0 \leq y \leq x \leq 1$ and $0 \leq q \leq p \leq r \leq 1$, and construct two random variables, u and v, in the following manner: u takes the values x, y, and 0 with probabilities q, $r - q$, and $1 - r$, respectively, and v takes the values x and 0 with probabilities p and $1 - p$, respectively. Assume that $(p - q)x = (r - q)y$. Then, by direct calculation, (12) holds for $0 \leq T \leq 1$, with equality for $T = 1$. Hence, $u \gtrsim v$. By Theorem 1, $u \gtrsim v \Leftrightarrow U(u) \geq U(v)$, where U is defined in (9). Computing, one finds that $U(u) = yf(r) + (x - y)f(q)$ and $U(v) = xf(p)$. The following implication

(13) $$(p - q)x = (r - q)y \Rightarrow yf(r) + (x - y)f(q) \geq xf(p)$$

has therefore been derived, for any five numbers, x, y, p, q, r satisfying $0 \leq y \leq x \leq 1$ and $0 \leq q \leq p \leq r \leq 1$. Note that (13) is trivial when $r = q$, so assume $r > q$. Define λ, $0 \leq \lambda \leq 1$, by writing $\lambda = (p - q)/(r - q)$ and note that $p = \lambda r + (1 + \lambda)q$. Now (13) reduces to the condition that

$$y = \lambda x \Rightarrow yf(r) + (x - y)f(q) \geq xf(\lambda r + (1 - \lambda)q)$$

must hold for all λ and x in the unit interval. For $x > 0$, this is precisely the statement that f is convex. Conversely, let $\gtrsim$ satisfy A1–A5 and suppose that u and v satisfy (12) for $0 \leq T \leq 1$, with equality for $T = 1$. Then, by a theorem of Hardy, Littlewood, and Polya (1929, Theorem 10), the inequality

$$\int_0^1 f(G_u(t))\, dt \geq \int_0^1 f(G_v(t))\, dt$$

holds for every convex and continuous f. Checking (9), we conclude that $U(u) \geq U(v)$—or $u \gtrsim v$—holds whenever f is continuous and convex. Thus, if the function f of Theorem 1 is convex, then $\gtrsim$ is risk averse, as was to be shown.   *Q.E.D.*

The fact that risk aversion is characterized in the dual theory by the convexity of f has a useful interpretation when f happens to be differentiable. Let v belong to V, with DDF G_v, and let $U(v)$ be the utility number assigned to v under the dual theory, i.e., $U(v) = \int f(G_v(t)) \, dt$. If f is differentiable, then the expression for $U(v)$ can be integrated by parts to obtain

$$U(v) = \int_0^1 tf'(G_v(t)) \, dF_v(t),$$

where F_v is the *cumulative* distribution of v. Note that $\int f'(G_v(t)) \, dF_v(t) = 1$, i.e., $\{f'(G_v(t))\}$ is a system of nonnegative weights summing to 1, and recall that $\int t \, dF_v(t)$ is the mean of v. In $U(v)$, a similar integral is being calculated, but each t is given a weight $f'(G_v(t))$. In other words, $U(v)$ is a corrected mean of v, in which the payment level t receives a weight of size $f'(G_v(t))$. If f is convex, then f' is nondecreasing; i.e., those values of t for which $G_v(t)$ is small receive relatively low weights and those values of t for which $G_v(t)$ is large receive relatively high weights. Thus, $U(v)$ is a corrected mean of v, in which low payments (bad outcomes) receive relatively high weights while high payments (good outcomes) receive relatively low weights. The agent behaves pessimistically, as though bad outcomes are more likely than they really are and good outcomes are less likely than they really are. It should be emphasized however, that this is *not* a case where probabilities are being distorted in the agent's perception. For the analysis undertaken in this essay deals with how perceived risk is processed into choice, and not with how actual risk is processed into perceived risk. This is necessarily true in any theory that subscribes to the neutrality axiom, A1. Let (S, Σ, P) be the probability space underlying the set V, over which preferences are defined. Then, the measure P must be interpreted as the agent's *perceived* probability measure, whether it coincides with some "objective" probability measure or not. If P were a measure that was liable to be modified (or "distorted") before entering the agent's choice process, then assuming neutrality with respect to P would have been completely unwarranted.

Having seen how risk aversion is characterized under the dual theory, one is led to ask about how the *degree* of risk aversion might be assessed, and to seek tools for carrying out *comparisons* of risk aversion. These topics are taken up in a separate paper (Yaari (1986)).

6. LIQUIDITY PREFERENCE AND COMPARATIVE STATICS

One of the hallmarks of expected utility theory is its treatment of portfolio selection. It is therefore interesting to see how the dual theory would cope with this classical topic. We begin by considering Tobin's (1958) basic liquidity preference problem.

There are two assets: A safe asset (cash) and a risky security. The rate of return on cash is 0 and the rate of return on the risky security is θ, where θ is a random variable distributed on the interval $[-1, a]$, for some $a > 0$. One must assume, of course, that $E\theta > 0$. A decision maker wishes to invest a fixed amount K,

satisfying $0 \leqslant K \leqslant 1/(1+a)$, and faces the problem of dividing this amount between cash and the risky security. Let x be the amount invested in the risky security, $0 \leqslant x \leqslant K$. Then, the decision maker's gross return from his/her portfolio is given by the random variable $K + \theta x$, which belongs to the class V of the previous sections.

Let $\succeq$ be the decision maker's preference order on V, and assume that $\succeq$ satisfies Axioms A1–A5 of Section 2. Then, by Theorem 1, there exists a continuous and nondecreasing real function f, satisfying the preference equation (7), such that picking the best portfolio is equivalent to selecting an x in the interval $[0, K]$ so as to maximize the quantity

$$(14) \qquad \Psi(x) = \int_0^1 f(G_{K+\theta x}(t)) \, dt,$$

where $G_{K+\theta x}$ is the DDF of $K + \theta x$.

PROPOSITION 4: *The function $\Psi(\cdot)$, defined in (14), is of the form*

$$\Psi(x) = K + cx, \quad 0 \leqslant x \leqslant K,$$

where the constant c, is given by

$$(15) \qquad c = \int_{-1}^{a} f(G_\theta(t)) \, dt - 1,$$

with G_θ being the DDF of θ.

PROOF: Essentially the same as the proof of Proposition 2, in Section 2.

Q.E.D.

With $\Psi(x)$ being linear in x, we find the dual theory predicting plunging, rather than diversification. Specifically, letting x^* be the maximizer of $\Psi(x)$ under $0 \leqslant x \leqslant K$, we find, from Proposition 4, that

$$x^* = \begin{cases} 0 & \text{if } \int_{-1}^{\infty} f(G_\theta(t)) \, dt < 1, \\[2ex] \text{any value in } [0, K] & \text{if } \int_{-1}^{\infty} f(G_\theta(t)) \, dt = 1, \\[2ex] K & \text{if } \int_{-1}^{\infty} f(G_\theta(t)) \, dt > 1. \end{cases}$$

The term "plunging" must not be confused with risk seeking. Indeed, consider a risk averse investor. Under the dual theory, the behavior of such an agent can be described, so to speak, as waiting in the wings until the rate of return is high enough, and then going whole hog. Under expected utility theory, on the other hand, diversification is universal, in the sense that the amount invested in the risky security is always positive, sometimes reaching the total available for

investment. This point deserves to be emphasized: Under expected utility, a risk averse investor will always put *some* resources into a risky security, provided its expected rate of return is positive. I am prepared to argue that both positions— "never stay put" in expected utility theory and "stay put until plunging becomes justified" in the dual theory—are extreme. Real investment behavior probably lies somewhere in between.

The dual theory, because of its linearity property, tends to produce corner solutions in optimization problems. This is why we get plunging behavior in the foregoing liquidity preference problem. However, it is easy to think of more complex portfolio problems, where diversification and corner solutions can coexist. Let us consider, for example, a three asset portfolio selection problem, with a safe asset (cash) earning no return and two risky securities whose rates of return are independent, identically distributed random variables. Under the dual theory, a risk averse investor facing this situation will either hold his/her assets in cash or in a diversified portfolio consisting of the two risky securities in equal amounts. Letting θ be the random variable describing the rate of return on this mixed asset and letting x be the amount invested in it, we find that Proposition 4 is applicable as it stands for the analysis of the investor's decision in this situation. An analysis of the general portfolio selection problem, in a dual theory setting, appears in Röell (1985).

We come now to the question of comparative statics. I shall claim that, despite the awkwardness brought about by corner solutions, the dual theory possesses desirable comparative statics properties. The framework, once again, will be that of the basic, two asset, liquidity preference problem. Recall that, under the dual theory, optimal behavior in this setting is determined by the constant c, given in (15). Plunging is optimal if $c > 0$, and holding back is optimal if $c < 0$. The constant c, therefore, acts like a measure of the agent's propensity to invest (i.e., to plunge), with environmental changes that reduce c tending to inhibit plunging and environmental changes that raise c tending to encourage plunging. Thus, it would be of interest to see how changes in various parameters affect this constant. Looking at equation (15), we note that c depends, on the one hand, on the function f representing the preference relation $\succsim$ and, on the other hand, on the DDF G_θ describing the rate of return on the risky security. To study the effect of a change in f, consider two functions, f_1 and f_2, representing two preference relations, $\succsim_1$ and $\succsim_2$, respectively. Intuitively, if $\succsim_1$ is more risk averse than $\succsim_2$, then f_1 will lie uniformly below f_2. (For a more rigorous discussion, see Yaari (1986).) Thus, if $\succsim_1$ is more risk averse than $\succsim_2$, then the corresponding respective values of c in (15), call them c_1 and c_2, will satisfy $c_1 \leqslant c_2$: Increased risk aversion inhibits plunging, and, as we might expect, the more risk averse the population the fewer the plungers.

Also of interest is the effect of a change in the distribution of returns on optimal behavior. In particular, one would like to know what the effect would be of an increase in the *riskiness* of the rate of return on the constant c. Consider two random variables, θ_1 and θ_2, taking values in the interval $[-1, a]$ and satisfying $E\theta_1 = E\theta_2 > 0$. Suppose that θ_1 is riskier than θ_2 (i.e., $\theta_1 = \theta_2 + \text{noise}$) and let c_1

and c_2 be the values of the constant c in (15), for $\theta = \theta_1$ and $\theta = \theta_2$, respectively. Assume that the investor is risk averse. Then, it is easy to see that the inequality $c_1 \leqslant c_2$ must hold. (This can be seen either directly in (15), or by looking at the equation $c = (\Psi(K) - K)/K$ and noting that $\Psi(K)$ is the utility which is assigned, under the dual theory, to the random variable $(1 + \theta)K$. For a risk averse agent, this utility decreases as riskiness increases.) Thus, we find that increased riskiness inhibits plunging, when investors are risk averse. With risk averse investors, the more risky the rate of return, the fewer the plungers. Compare this observation with the corresponding result in expected utility theory: Under expected utility, a risk averse investor may actually *increase* his/her security holding, in response to a rise in the riskiness of the security. Increased riskiness only inhibits investment under suitable assumptions on the *third* derivative of the utility function, assumptions that govern the relationship between the degree of risk aversion and the level of wealth. Under the dual theory, the only property needed is risk aversion itself.

Comparative statics without third derivative conditions is a general feature of the dual theory. This feature comes into its own in the multivariate version of the theory, where corner solutions no longer prevail. (See Yaari (1986) for details.)

7. MACHINA, QUIGGIN, SCHMEIDLER

The dual theory of choice under risk needs to be viewed in the light of other non-expected-utility theories that have been proposed recently. For want of space, I shall restrict my attention to three prominent and representative contributions, namely those of Mark Machina, John Quiggin, and David Schmeidler. I wish to emphasize that restricting myself in this way should by no means be construed as belittling the various other contributions to non-expected-utility theory that have appeared recently. I apologize also for imposing my own notation upon the work that I am about to cite.

In a paper well on its way to becoming a milestone, Machina (1982) studies preference among random variables in a spirit not unlike that of Section 2, above. In particular, conditions closely resembling Axioms A1–A4 are imposed. Like expected utility theory and the dual theory, Machina also needs a fifth axiom, but he rejects independence, whether "primal" or "dual". Instead, Machina's fifth condition is one that ensures existence of a Frechet-differentiable functional, call it "the Machina functional" and let it be denoted M, such that $u \succsim v \Leftrightarrow M(G_u) \geqslant M(G_v)$ holds for all random variables u and v, with respective DDF's G_u and G_v. Machina's fifth axiom is not stated directly on the preference relation and, to the best of my knowledge, a general axiom on preferences guaranteeing the existence of a suitable Machina functional has not yet been discovered. (However, see Allen (1986).) Of course, Axiom A5EU will do it, because of the linearity of M under expected utility. Axiom A5, on the other hand, fails to produce a suitable Machina functional. An example of this is easily obtained, by taking the function f that represents preferences under the dual theory to be nondifferentiable. Recently, Chew, Karni, and Safra (1985) have shown, in fact,

that even when f is differentiable, the functional M given by $M(G) = \int f(G(t))\, dt$ need not be Frechet differentiable. Strictly speaking, the dual theory falls outside Machina's framework. However, there does exist an extension of Machina's work, in which Frechet differentiability is replaced by the weaker Gateaux differentiability. (See Chew, Karni, and Safra, op. cit.) The dual theory does fall into this extended Machina framework, in those cases where the function f representing preferences happens to be differentiable. Indeed, the extended Machina framework can be used to prove a special case of our Theorem 2, namely that if preferences satisfy A1–A5 and the resulting function f is differentiable, then risk aversion is equivalent to f being convex. The main difference between Machina's work and the work being presented here is, in my opinion, a difference of *intent*. Machina's aim is to construct a general tool for analyzing *all* non-expected-utility theories. The aim of the dual theory, on the other hand, is to concentrate on a specific alternative.

Now let me try, as best I can, briefly to summarize Quiggin's proposal (Quiggin (1982)) in a way that will facilitate comparison with the dual theory. The basic approach is perceptional: Probabilities are liable to be adjusted (or distorted) in the decision maker's perception, before becoming an input in the decision process. We could think of a real function h such that, if p is a probability, then $h(p)$ stands for how p is perceived. A modified expected utility theory can now be constructed using what the agent perceives when facing a random variable. Specifically, if the agent faces a random variable taking the values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$, respectively, then, under such a modified theory, the utility number assigned to the random variable would be of the form $\sum h(p_i)\phi(x_i)$, with ϕ being a von Neumann–Morgernstern utility. Quiggin looks at this representation and notes, as several others have, that under continuity and mononicity (Axioms A3 and A4) the function h in the foregoing representation must coincide with the identity. He therefore offers the following more general representation: The utility number to be assigned to a random variable taking the values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$, respectively, shall be of the form $\sum h_i(p_1, \ldots, p_n)\phi(x_i)$, where $h = (h_1, \ldots, h_n)$ is now an n-component vector function, defined on the $(n-1)$-dimensional unit simplex. (Such a vector function is required to exist for every positive integer n.) The decision weight being applied to the ith value of the random variable now depends not only on p_i but on the entire vector $(p_1, \ldots, p_n)$. Quiggin now finds that, in order for this new representation to be consistent with A3 and A4, there must exist a real function f, defined on the unit interval, such that

$$h_i(p_1, \ldots, p_n) = f\left(\sum_{j=i}^{n} p_j\right) - f\left(\sum_{j=i+1}^{n} p_j\right).$$

He is therefore led to seek axioms which imply that preferences can be represented by a utility of the form

$$(16) \qquad Q(x_1, \ldots, x_n, p_1, \ldots, p_n) = \sum_{i=1}^{n} \phi(x_i)\left[f\left(\sum_{j=i}^{n} p_j\right) - f\left(\sum_{j=i+1}^{n} p_j\right)\right],$$

and he finds that a suitably weakened version of A5EU, together with suitable versions of A2, A3, and A4 will do the trick. (The neutrality axiom, A1, whose appropriateness in a theory of risk perception is questionable, is assumed implicitly.) Rewriting (16) for any random variable v belonging to V, we find preferences being represented by a utility Q of the form

$$(17) \qquad Q(v) = \int_0^1 \phi(t) \, d(f \circ G_v)(t) = \int_0^1 f(G_v(t)) \, d\phi(t),$$

where G_v is the DDF of v and ϕ is a von Neumann–Morgenstern utility. Now, when ϕ is the identity, (17) reduces to (9) and when f is the identity, (17) reduces to (10). Quiggin's representation theorem generalizes the representation theorems of both expected utility theory and the dual theory.

Since Quiggin's approach is perceptional, there is an empirical observation that he can use. This is the observation, often noted by students of risk perception, that a 50–50 proposition is in fact perceived by decision makers as a 50–50 proposition. The implication of this, under Quiggin's theory, is that, when facing 50–50 propositions, agents always act like expected utility maximizers, and this property is relied upon heavily in Quiggin's arguments. It follows, however, from this property that the function f in (17) must satisfy $f(\frac{1}{2}) = \frac{1}{2}$. This fact, in conjunction with the recent work of Chew, Karni, and Safra (1985), implies that all risk averse agents in Quiggin's framework must be expected utility maximizers, because the only convex f satisfying $f(0) = 0$, $f(\frac{1}{2}) = \frac{1}{2}$, and $f(1) = 1$ is the identity. (Perhaps I should mention also that Quiggin's representation theorem is incorrect as it stands. Utility representations of the form $\sum x_i w(p_i)$, with w continuous and satisfying $w(p) + w(1 - p) \leq 1$ for all p, satisfy Quiggin's axioms but they do not agree with (17), unless $w(p) = p$. To fix things up, Quiggin's Dominance Axiom must certainly be modified, and possibly also his Independence Axiom.)

It is interesting to note that the dual theory can serve as a building block in an alternative axiomatization of (17). The idea is related to a recent paper by Shubik (1985). Suppose that an agent who faces a random variable v, belonging to V, acts in the following way: First, the agent considers the payment levels $v(s)$, for all states-of-nature s. Each payment level, $v(s)$, is processed by the agent into a utility level, $\phi(v(s))$, where ϕ is a cardinal utility generated from some *riskless* intensity-of-preference framework. (See, e.g., Shapley (1975).) Now the agent faces the random variable $\phi(v)$ which belongs to V, under a suitable normalization of a bounded ϕ. The axioms of the dual theory (i.e., Axioms A1–A5 above) may now be postulated for preferences among these utility-valued random variables. The result is a theory in which preferences among the original, money-valued, random variables are represented by a utility of the form $U(\phi \circ v)$, where U is defined in (9). But now we find that $U(\phi \circ v) = \int f(G_v(t)) \, d\phi(t)$ for some appropriate real function f, so (17) is obtained as a utility representation for preferences among the money-valued random variables.

Finally, an interesting relationship between utility representations of the form (17) and the notion of comonotonicity (see Section 3) can be seen in a recent paper by Schmeidler (1984). Schmeidler's concern is to show that preference

among *acts* (not among random variables) can be represented, under suitable assumptions, by an expected utility, in which expectation is taken with respect to some *nonadditive* measure. (An act is a measurable real function on some measurable space without a probability.) Following Anscombe and Aumann (1963), Schmeidler assumes that agents can always toss coins, if they wish, thereby obtaining "objective" probability mixtures of acts. What Anscombe and Aumann had done was to write down an independence axiom for such mixtures which, together with other suitable axioms, implies that preference among acts has an expected utility representation. Schmeidler's idea is to require this kind of independence only for pairwise comonotonic acts. From this, together with other standard axioms, he obtains the result that there exists a von Neumann-Morgenstern utility ϕ and a nonadditive measure μ such that $u \gtrsim v \Leftrightarrow \int \phi(u) \, d\mu \geq \int \phi(v) \, d\mu$ holds for every pair of acts, u and v. (This result has recently been extended by Gilboa (1985) to the case where uncertainty is totally subjective; i.e., "objectively" mixed acts are not necessarily available.) Now let us return to (17) and recall that the preference relation being treated there is over *random variables*, with some underlying probability measure, P. Defining a nonadditive measure μ by $\mu = f \circ P$, we find preferences being represented precisely by Schmeidler's utility, $\int \phi(v) \, d\mu$. Note that Schmeidler's independence axiom deals with *probability* mixtures of comonotonic functions, whereas the independence axiom of the dual theory (Axiom A5*) deals with pointwise mixtures of the *values* of comonotonic functions.

Institute for Advanced Studies, Hebrew University, Givat Ram, Jerusalem, Israel

REFERENCES

ALLAIS, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque," *Econometrica*, 21, 503–546.

ALLEN, B. (1986): "Smooth Preferences and the Local Expected Utility Hypothesis," *Journal of Economic Theory*, forthcoming.

ANSCOMBE, F. J., AND R. J. AUMANN (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

ATKINSON, A. B. (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263.

BLACKWELL, D. (1950): "Comparison of Experiments," *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics*. Berkeley: University of California Press, 93–102.

CHEW, S. H., E. KARNI, AND Z. SAFRA (1985): "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities," Working Paper No. 16-85, Foerder Institute, Tel-Aviv University.

DORFMAN, R. (1979): "A Formula for the Gini Coefficient," *Review of Economics and Statistics*, 61, 146–149.

FISHBURN, P. C. (1982): *The Foundations of Expected Utility*. Dordrecht: Reidel Publishing Co.

GILBOA, I. (1985): "Expected Utility with Purely Subjective Non-Additive Probabilities," Working Paper No. 6-85, Foerder Institute, Tel-Aviv University.

HARDY, G. H., J. E. LITTLEWOOD, AND G. POLYA (1929): "Some Simple Inequalities Satisfied by Convex Functions," *Messenger of Mathematics*, 58, 145–152.

HERSTEIN, I. N., AND J. MILNOR (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21, 291–297.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.

MACHINA, M. J. (1982): "Expected Utility Analysis without the Independence Axiom," *Econometrica*, 50, 277-323.

NEWBERY, D. (1979): "A Theorem on the Measurement of Inequality," *Journal of Economic Theory*, 2, 264-266.

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

RÖELL, A. (1985): "Risk Aversion in Yaari's Rank-Order Model of Choice under Uncertainty," mimeo., London School of Economics.

ROTHSCHILD, M., AND J. E. STIGLITZ (1970): "Increasing Risk: A Definition," *Journal of Economic Theory*, 2, 225-243.

SCHMEIDLER, D. (1984): "Subjective Probability and Expected Utility without Additivity," IMA Preprint Series, University of Minnesota.

SHAPLEY, L. S. (1975): "Cardinal Utility from Intensity Comparisons," RAND Publication R-1683-PR.

SHUBIK, M. (1985): "A Note on Cardinal Utility with Probability Mixtures and Intensity Comparisons," Cowles Foundation Preliminary Paper No. 850423, Yale University.

TOBIN, J. (1958): "Liquidity Preference as Behavior Toward Risk," *Review of Economic Studies*, 25, 65-86.

YAARI, M. E. (1969): "Some Remarks on Measures of Risk Aversion and on Their Uses," *Journal of Economic Theory*, 1, 315-329.

——— (1986): "Univariate and Multivariate Comparisons of Risk Aversion: A New Approach," forthcoming in *Essays in Honor of Kenneth J. Arrow*, ed. by W. P. Heller, R. Starr, and D. Starrett. Cambridge: Cambridge University Press.