MARK J. MACHINA

TWO ERRORS IN THE 'ALLAIS IMPOSSIBILITY THEOREM'

ABSTRACT. In his so-called 'Allais Impossibility Theorem', Allais (1988) asserts that the technique of 'generalized expected utility analysis' from Machina (1982) is invalid, on the grounds that its key tool, the 'local utility function', cannot be well-defined for preferences over finite-outcome lotteries. This paper presents a brief description of the local utility function and a summary of Allais' argument, and points out two separate errors in the latter. The first error consists of believing that two local utility functions should be affinely equivalent even when their outcome variables differ by a nonlinear transformation. The second error consists of an incorrect derivation of the local utility function, resulting from extending a moment representation function beyond its valid domain, and/or invoking the chain rule at a point where this function is not differentiable.

KEY WORDS: Generalized expected utility analysis, local utility functions, non-expected utility.

In Machina (1982, 1983) I presented an analytical approach for the study of 'smooth' (i.e., differentiable) non-expected utility preference functionals over bounded univariate probability distributions, which I termed 'generalized expected utility analysis'. Put simply, it consists of the observation that, as with *any* differentiable function or functional, a smooth functional over probability distributions will possess a 'local linear representation' at each point (i.e., probability distribution) in its domain. Since linearity in the probabilities is mathematically equivalent[1] to the expected utility property, this means that a smooth preference functional is 'locally expected utility' at each probability distribution in its domain. I termed the utility function that represents the local expected utility ranking at each probability distribution the 'local utility function' at that distribution.

The value of this approach is that, as with any application of finite or infinite-dimensional calculus, global hypotheses on the nature of a preference functional's derivatives (its local utility functions) are often equivalent to exact *global properties* of the preference functional itself. Thus, for example, a preference functional will be *globally* risk averse[2] if and only if its local utility functions are concave

Theory and Decision **38:** 231–250, 1995.

at each distribution in its domain (Machina, 1982, Thm. 2). This analytical approach has been extended to larger classes of preference functionals and distributions by Chew, Karni and Safra (1987), Karni (1987, 1989) and Wang (1993); it has been formally axiomatized by Allen (1987); and it has been applied to the analysis of choice under uncertainty by Chew, Epstein and Zilcha (1988), Chew and Nishimura (1992), Green and Jullien (1988), Machina (1984, 1987), Machina and Neilson (1987), Röell (1986) and others.

In Allais (1988), Maurice Allais presented an alleged proof of the following statement, which in Allais (1986) and elsewhere he refers to as the 'Allais Impossibility Theorem':

> Machina's 'local utility' function cannot be validly defined over the whole domain $(0, M)$ of variation of the x_i considered by Machina. (Allais, 1988, p. 380)

where the x_i's are the outcome values of a finite-support probability distribution. This had led Allais to conclude that

> the very foundation of Machina's theory does collapse. (Allais, 1988, p. 382)

The purpose of this paper is to identify two separate errors in Allais' argument. Section I defines the notion of the local utility function of a preference functional, and Section II gives a step-by-step presentation of Allais' argument. Sections III and IV identify the two mistakes, each of which renders the argument invalid. In each case, I present a counterexample to Allais' assertion, as well as describe the nature of his error. The first mistake is a misunderstanding of when two local utility functions (or for that matter, two von Neumann–Morgenstern utility functions) derived from the same preference ranking should be affinely equivalent. The second mistake consists of an incorrect derivation of the local utility function, resulting from extending a moment representation function beyond its valid domain, and/or invoking the chain rule at a point where this function is not differentiable. For the benefit of readers who do not wish to consult Allais (1988), the presentation in this paper is self-contained. For the benefit of readers who do wish to consult it, cross references by page and/or equation number are given throughout.

I. LOCAL UTILITY FUNCTIONS FOR PREFERENCES OVER FINITE-OUTCOME PROBABILITY DISTRIBUTIONS

In the analysis of Machina (1982), I considered a general preference function $V(F)$ over the set of all probability distributions – discrete, continuous or otherwise – over a real interval $[0, M]$, as represented by their cumulative distribution functions $F(\cdot)$ on $[0, M]$. Since it is Allais' contention that this approach cannot be applied to the subset of finite-outcome distributions over this interval, in the present development I restrict attention to the subset of all finite-outcome probability distributions $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)(1 \leqslant n < \infty)$ on $[0, M]$.

One way to express the idea of generalized expected utility analysis is that, once we reinterpret the standard results of expected utility analysis in terms of the probability derivatives of the expected utility preference functional, it turns out that many of these same results also apply to the probability derivatives of general *non-expected utility* preference functionals. The 'local utility function' is precisely this probability derivative.

Accordingly, recall that given any expected utility preference functional

$$(1) \qquad V(\mathbf{P}) \equiv V(x_1, \ldots, x_n, p_1, \ldots, p_n) \equiv \sum_{i=1}^{n} U(x_i) \cdot p_i$$

over finite-outcome probability distributions, we have the following classic results:

(a) $V(\cdot)$ exhibits *first order stochastic dominance preference* if and only if $U(\cdot)$ is increasing, i.e., if and only if

$$U(x) < U(x') \qquad \text{for all } x < x'$$

(b) $V(\cdot)$ exhibits *risk aversion* (i.e., aversion to mean preserving increases in risk) if and only if $U(\cdot)$ is concave, i.e., if and only if

$$\frac{U(x') - U(x)}{x' - x} > \frac{U(x'') - U(x')}{x'' - x'}$$

$$\text{for all } x < x' < x''$$

(c) some other expected utility preference functional $V^*(\mathbf{P}) \equiv \sum_{i=1}^{n} U^*(x_i) \cdot p_i$ is *at least as risk averse*[3] *as* $V(\cdot)$ if $U^*(\cdot)$ is a concave transformation of $U(\cdot)$, i.e., if[4]

$$\frac{U(x'') - U(x')}{U(x') - U(x)} \geqslant \frac{U^*(x'') - U^*(x')}{U^*(x') - U^*(x)}$$

$$\text{for all } x < x' < x''$$

In order to interpret these results in a manner that can be generalized beyond expected utility, observe that, for any $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$, the von Neumann–Morgenstern utility function $U(\cdot)$ at any outcome level x_i in the support of $\mathbf{P}$ can be expressed as the derivative of the formula $V(x_1, \ldots, x_n, p_1, \ldots, p_n) \equiv \sum_{i=1}^{n} U(x_i) \cdot p_i$ with respect to the probability of obtaining the outcome x_i:

$$(2) \qquad \begin{aligned} \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x_i)} &\equiv \frac{\partial V(x_1, \ldots, x_n, p_1, \ldots, p_n)}{\partial p_i} \\ &\equiv \frac{\partial [U(x_1) \cdot p_1 + \cdots + U(x_n) \cdot p_n]}{\partial p_i} \\ &\equiv U(x_i) \end{aligned}$$

Of course, taking some $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$ and increasing/decreasing a *single* p_i by some amount $\mathrm{d}p_i$ means that we no longer have a probability distribution, since the variables $\{p_1, \ldots, p_{i-1}, p_i + \mathrm{d}p_i, p_{i+1}, \ldots, p_n\}$ no longer sum to unity. In other words, while equation (2) is meaningful as the derivative of the *mathematical formula* $V(x_1, \ldots, x_n, p_1, \ldots, p_n) \equiv \sum_{i=1}^{n} U(x_i) \cdot p_i$ with respect to some p_i, it does not represent a change from $\mathbf{P}$ to some other actual probability distribution. However, as long as the *set* of changes $(\mathrm{d}p_1, \ldots, \mathrm{d}p_n)$ sum to zero, we can use these derivatives jointly to represent such a response. In other words, for any distribution $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$ and any vector of changes $(\mathrm{d}p_1, \ldots, \mathrm{d}p_n)$ summing to zero,[5] equation (2) implies[6]

$$(3) \qquad \mathrm{d}V(\mathbf{P}) \equiv \sum_{i=1}^{n} \frac{\partial V(\mathbf{P})}{\partial p_i} \cdot \mathrm{d}p_i \equiv \sum_{i=1}^{n} U(x_i) \cdot \mathrm{d}p_i$$

For purposes of pointing out one of the errors in Allais' argument, it is important to note that the interpretation of $U(x)$ as $\partial V(\mathbf{P})/\partial \text{prob}(x)$ is not restricted solely to outcome levels x_i in the support of $\mathbf{P}$. To see this, take any distribution $\mathbf{P} = (x_1, \ldots, x_n,$

and any outcome level x in $[0, M]$ but *not* in $\{x_1, \ldots, x_n\}$. Since formula (1) is defined for any positive integer n, we can write

$$
\begin{aligned}
V(x_1, \ldots, x_n, p_1, \ldots, p_n) &\equiv \sum_{i=1}^{n} U(x_i) \cdot p_i \\
&\equiv \left[\sum_{i=1}^{n} U(x_i) \cdot p_i + U(x) \cdot p_x \right] \Bigg|_{p_x=0} \\
&\equiv V(x_1, \ldots, x_n, x, p_1, \ldots, p_n, p_x)|_{p_x=0} \\
&\equiv V(x_1, \ldots, x_n, x, p_1, \ldots, p_n, 0)
\end{aligned}
\tag{4}
$$

so that we have

$$
\begin{aligned}
\frac{\partial V(\mathbf{P})}{\partial \mathrm{prob}(x)} &\equiv \left. \frac{\partial V(x_1, \ldots, x_n, x, p_1, \ldots, p_n, p_x)}{\partial p_x} \right|_{p_x=0} \\
&\equiv \left. \frac{\partial [\sum_{i=1}^{n} U(x_i) \cdot p_i \ + \ U(x) \cdot p_x]}{\partial p_x} \right|_{p_x=0} \\
&\equiv U(x)
\end{aligned}
\tag{5}
$$

Thus we have that for any probability distribution $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$, any outcome x in $[0, M] - \{x_1, \ldots, x_n\}$, and any vector of differential changes $(\mathrm{d}p_1, \ldots, \mathrm{d}p_n, \mathrm{d}p_x)$ summing to zero:

$$
\begin{aligned}
\mathrm{d}V(\mathbf{P}) &\equiv \sum_{i=1}^{n} \frac{\partial V(\mathbf{P})}{\partial p_i} \cdot \mathrm{d}p_i \ + \ \frac{\partial V(\mathbf{P})}{\partial p_x} \cdot \mathrm{d}p_x \\
&\equiv \sum_{i=1}^{n} U(x_i) \cdot \mathrm{d}p_i \ + \ U(x) \cdot \mathrm{d}p_x
\end{aligned}
\tag{6}
$$

Rewriting the above three results in terms of these ‘probability derivatives’, we obtain:

(a′) $V(\cdot)$ exhibits *first order stochastic dominance preference* if and only if at any distribution $\mathbf{P}$, $\partial V(\mathbf{P})/\partial \mathrm{prob}(x)$ is an increasing function of x, i.e., if and only if

$$\frac{\partial V(\mathbf{P})}{\partial \mathrm{prob}(x)} < \frac{\partial V(\mathbf{P})}{\partial \mathrm{prob}(x')} \qquad \text{for all } x < x'$$

(b′) $V(\cdot)$ exhibits *risk aversion* (aversion to mean preserving increases in risk) if and only if at any distribution $\mathbf{P}$, $\partial V(\mathbf{P})/\partial \mathrm{prob}(x)$

is a concave function of x, i.e., if and only if

$$\frac{\frac{\partial V(\mathbf{P})}{\partial \text{prob}(x')} - \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x)}}{x' - x} > \frac{\frac{\partial V(\mathbf{P})}{\partial \text{prob}(x'')} - \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x')}}{x'' - x'}$$

for all $x < x' < x''$

(c′) some other preference functional $V^*(\cdot)$ is *at least as risk averse as* $V(\cdot)$ if at any distribution $\mathbf{P}$, $\partial V^*(\mathbf{P})/\partial\text{prob}(x)$ is a concave transformation of $\partial V(\mathbf{P})/\partial\text{prob}(x)$, i.e., if

$$\frac{\frac{\partial V(\mathbf{P})}{\partial \text{prob}(x'')} - \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x')}}{\frac{\partial V(\mathbf{P})}{\partial \text{prob}(x')} - \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x)}} \geqslant \frac{\frac{\partial V^*(\mathbf{P})}{\partial \text{prob}(x'')} - \frac{\partial V^*(\mathbf{P})}{\partial \text{prob}(x')}}{\frac{\partial V^*(\mathbf{P})}{\partial \text{prob}(x')} - \frac{\partial V^*(\mathbf{P})}{\partial \text{prob}(x)}}$$

for all $x < x' < x''$

To see that expressing the above results in this manner allows for their direct generalization to the non-expected utility case, consider result (a/a′) on first order stochastic dominance preference. It is well known that any first order stochastically dominating change from one discrete probability distribution to another one consists of one or more shifts of probability mass from lower outcome values up to higher outcome values. It is also clear from the inequality in (a′) that any such increase in the probability of a higher outcome x' at the expense of a matching decrease in the probability of a lower outcome x will lead to a net increase in the value of the preference functional $V(\cdot)$, so that all such first order stochastically dominating changes – be they 'small' (differential) or 'large' – will be strictly preferred. It is crucial to note that nowhere in this argument do we need the assumption that $V(\cdot)$ is necessarily linear in the probabilities (i.e., that it is necessarily expected utility), so that result (a′) is true for *all* smooth preference functionals $V(\cdot)$. Similar arguments demonstrate that results (b′) and (c′) also hold for general smooth probability distributions.[7]

To highlight the correspondence between the probability derivative $\partial V(\mathbf{P})/\partial\text{prob}(x)$ of a general smooth preference functional $V(\cdot)$ and the von Neumann–Morgenstern utility function $U(\cdot)$ of an expected utility preference functional, I have adopted the following notation for the probability derivative:

$$(7) \qquad U(x;\mathbf{P}) \overset{\text{def}}{\equiv} \frac{\partial V(\mathbf{P})}{\partial \text{prob}(x)} \qquad \text{for all } x \in [0, M]$$

and refer to the function $U(\cdot;\mathbf{P})$ as the *local utility function* of $V(\cdot)$ at the distribution $\mathbf{P}$. Using this notation, we can write the above 'generalized expected utility' results (a′), (b′) and (c′) as:

(a″) $V(\cdot)$ exhibits *first order stochastic dominance preference* if and only if at any distribution $\mathbf{P}$, $U(x;\mathbf{P})$ is an increasing function of x, i.e., if and only if

$$U(x;\mathbf{P}) < U(x';\mathbf{P}) \qquad \text{for all } x < x'$$

(b″) $V(\cdot)$ exhibits *risk aversion* (aversion to mean preserving increases in risk) if and only if at any distribution $\mathbf{P}$, $U(x;\mathbf{P})$ is a concave function of x, i.e., if and only if

$$\frac{U(x';\mathbf{P}) - U(x;\mathbf{P})}{x' - x} > \frac{U(x'';\mathbf{P}) - U(x';\mathbf{P})}{x'' - x'}$$

$$\text{for all } x < x' < x''$$

(c″) Some other preference functional $V^*(\cdot)$ is *at least as risk averse as* $V(\cdot)$ if at any distribution $\mathbf{P}$, $U^*(x;\mathbf{P})$ is a concave transformation of $U(x;\mathbf{P})$, i.e., if

$$\frac{U(x'';\mathbf{P}) - U(x';\mathbf{P})}{U(x';\mathbf{P}) - U(x;\mathbf{P})} \geqslant \frac{U^*(x'';\mathbf{P}) - U^*(x';\mathbf{P})}{U^*(x';\mathbf{P}) - U^*(x;\mathbf{P})}$$

$$\text{for all } x < x' < x''$$

For additional applications of generalized expected utility analysis to the study of smooth non-expected utility preference functionals, the reader is referred to the articles cited in the introduction to this paper, especially the insightful piece by Chew, Epstein and Zilcha (1988).

Finally, since I shall make important use of the following fact, it is worth possibly belaboring the obvious by stating it explicitly: Equations (2), (5) and (7) together imply that the local utility function $U(\cdot;\mathbf{P})$ for any *expected utility* preference functional $V(\mathbf{P}) \equiv \sum_{i=1}^{n} U(x_i) \cdot p_i$ at any probability distribution $\mathbf{P}$ is exactly its von Neumann–Morgenstern utility function $U(\cdot)$.

II. ALLAIS' 'IMPOSSIBILITY THEOREM'

Professor Allais' argument, as sketched out in Section 2 (especially pp. 358–361) of Allais (1988) and presented more formally in its Appendix A (pp. 377–386), proceeds as follows:[8]

Allais begins with a general non-expected utility preference functional

$$(8) \qquad \Phi = H(x_1, \ldots, x_n, p_1, \ldots, p_n) \qquad (1)$$

over finite-outcome probability distributions. Defining the kth absolute moment of the probability distribution $\mathbf{P} = (x_1, \ldots, x_n, p_1, \ldots, p_n)$ by

$$(9) \qquad M_k = \sum_{i=1}^{n} p_i \cdot x_i^k \qquad (4)$$

and stating that "a discrete distribution of order n is fully determined if its $2n-1$ first moments are given" (p. 357), he obtains the following representation of the preference functional $H(\cdot)$:

$$(10) \quad H(x_1, \ldots, x_n, p_1, \ldots, p_n) \equiv G(M_1, \ldots, M_{2n-1}) \qquad (5)$$

Allais now considers a general increasing (though not necessarily linear) transformation

$$(11) \quad y = f(x) \qquad (7)$$

of the outcome variable x. Since $f(\cdot)$ is a one-to-one function, we can equivalently represent the original preference functional in equation (8) as

$$(12) \quad H^*(y_1, \ldots, y_n, p_1, \ldots, p_n) \equiv H(x_1, \ldots, x_n, p_1, \ldots, p_n) \, (12)$$

where of course

$$(13) \quad y_i = f(x_i) \qquad (13)$$

Defining the moments of the random variable y by[9]

$$(14) \quad \mathcal{M}_k = \sum_{i=1}^{n} p_i \cdot y_i^k = \sum_{i=1}^{n} p_i \cdot f^k(x_i) \qquad (14,17)$$

he obtains a corresponding moment representation of the preference functional $H^*(\cdot)$:

$$(15)\qquad H^*(y_1,\dots,y_n,p_1,\dots,p_n) \equiv G^*(\mathcal{M}_1,\dots,\mathcal{M}_{2n-1}) \qquad (15)$$

The remainder of Allais argument consists of three steps:[10] First, he claims (p. 381) that, “According to Machina” the local utility functions of the preference functionals $H(\cdot)$ and $H^*(\cdot)$ *“should be identical up to within a linear* [i.e., increasing affine] *transformation for any value of* $x(0 \leqslant x \leqslant M)$”. Second, he uses the moment representations (10) and (15) of $H(\cdot)$ and $H^*(\cdot)$ to derive what he claims are their respective local utility functions. Finally, he shows that the functions he obtains cannot be increasing affine transformations of each other.

To calculate the local utility function $U(\cdot;\mathbf{P})$[11] of the preference functional $H(\cdot)$, Allais takes the differential expansion of its moment representation function $G(M_1,\dots,M_{2n-1})$, to obtain

$$(16)\qquad \begin{aligned} \mathrm{d}\Phi &= \sum_{i=1}^{n}\left[\sum_{k=1}^{2n-1}\frac{\partial G(M_1,\dots,M_{2n-1})}{\partial M_k}\cdot\frac{\partial M_k}{\partial p_i}\right]\cdot \mathrm{d}p_i \\ &= \sum_{i=1}^{n}\left[\sum_{k=1}^{2n-1}\frac{\partial G(M_1,\dots,M_{2n-1})}{\partial M_k}\cdot x_i^k\right]\cdot \mathrm{d}p_i \end{aligned} \qquad (19)$$

so that (according to Allais) the local utility function of $H(\cdot)$ for outcome values $x \in [0, M]$ must be

$$(17)\qquad \begin{aligned} U(x;\mathbf{P}) &= \frac{\partial H(\mathbf{P})}{\partial \mathrm{prob}(x)} = \frac{\partial G(M_1,\dots,M_{2n-1})}{\partial \mathrm{prob}(x)} \\ &= \sum_{k=1}^{2n-1}\frac{\partial G(M_1,\dots,M_{2n-1})}{\partial M_k}\cdot x^k \end{aligned} \qquad (20)$$

From equations (14) and (15), a similar differential expansion of $G^*(\mathcal{M}_1,\dots,\mathcal{M}_{2n-1})$ yields

$$(18)\qquad \mathrm{d}\Phi = \sum_{i=1}^{n}\left[\sum_{k=1}^{2n-1}\frac{\partial G^*(\mathcal{M}_1,\dots,\mathcal{M}_{2n-1})}{\partial \mathcal{M}_k}\cdot f^k(x_i)\right]\cdot \mathrm{d}p_i \qquad (21)$$

so that (according to Allais) the local utility function of $H^*(\cdot)$ for outcomes $x \in [0, M]$ must be

$$U^*(x;\mathbf{P}) = \frac{\partial H^*(\mathbf{P})}{\partial \mathrm{prob}(x)} = \frac{\partial G^*(\mathcal{M}_1,\dots,\mathcal{M}_{2n-1})}{\partial \mathrm{prob}(x)}$$

(19) (22)

$$= \sum_{k=1}^{2n-1} \frac{\partial G^*(\mathcal{M}_1,\ldots,\mathcal{M}_{2n-1})}{\partial \mathcal{M}_k} \cdot f^k(x)$$

The final step in Allais' argument, showing that the formulas he obtains for the local utility functions $U(\cdot,\mathbf{P})$ and $U^*(\cdot;\mathbf{P})$ cannot be increasing affine transformations of each other, follows from the structure of the formulas (17) and (19). Note that since the sets of moments $\{M_1,\ldots,M_{2n-1}\}$ and $\{\mathcal{M}_1,\ldots,\mathcal{M}_{2n-1}\}$, and hence the sets of partial derivatives $\{\partial G/\partial M_1,\ldots,\partial G/\partial M_{2n-1}\}$ and $\{\partial G^*/\partial \mathcal{M}_1,\ldots,\partial G^*/\partial \mathcal{M}_{2n-1}\}$, are functions of the probability distribution alone and *not* the value of x, it would follow from (17) that the local utility function $U(x;\mathbf{P})$ is a $2n-1$ degree polynomial in x over the interval $[0, M]$, while it would follow from (19) that the local utility function $U^*(x;\mathbf{P})$ is a $2n-1$ degree polynomial in $f(x)$. Since the function $f(\cdot)$ was arbitrary, argues Allais, these two formulas cannot be identical up to an increasing affine transformation over any real interval.

III. WHEN SHOULD TWO LOCAL UTILITY FUNCTIONS (OR TWO VON NEUMANN–MORGENSTERN UTILITY FUNCTIONS) BE AFFINELY EQUIVALENT?

The essence of Allais' argument is to present two local utility functions derived from the same underlying preference ranking that are not affinely equivalent, even though "according to Machina *they should be identical up to within a linear* [i.e., increasing affine] *transformation*" (Allais, 1988, p. 381). Allais' first error is in believing that the property of affine equivalence should be expected to hold, or that I claim it should hold, *under nonlinear transformations of the outcome variable.* The following counterexample shows that this is not true, even in the simple special case of expected utility preferences:

COUNTEREXAMPLE 1. Define the expected utility preference functional $H(x_1,\ldots,x_n,p_1,\ldots,p_n) \equiv \sum_{i=1}^{n} U(x_i)\cdot p_i$ where $U(x) \equiv x^2$ over the domain $x \in [2,3] \subset R^1$. Since the local utility function of any expected utility preference functional is just its von Neumann–Morgenstern utility function, it is clear that at any distribution $\mathbf{P}$, the local utility function of $H(\cdot)$ is given by the formula $U(x;\mathbf{P}) \equiv U(x) \equiv x^2$, or any increasing affine transformation of this formula, which will clearly be a *convex* function over its domain $[2,3]$.

Now define the nonlinear transformation $y = f(x) = x^4$, so that y ranges over the interval $[16,81] \subset R^1$. It is clear that we can represent the same individual's risk preferences in terms of the outcome variable y by means of the preference functional $H^*(y_1,\ldots,y_n, p_1,\ldots,p_n) \equiv \sum_{i=1}^{n} U^*(y_i)\cdot p_i$, where $U^*(y) \equiv U(f^{-1}(y)) = \sqrt{y}$ over the domain $y \in [16,81]$. Once again, $H^*(\cdot)$ is an expected utility functional, so that at any $\mathbf{P}$, its local utility function is given by $U^*(y;\mathbf{P}) \equiv U^*(y) \equiv \sqrt{y}$, or any increasing affine transformation of this formula, which will clearly be a *concave* function over its domain [16, 81]. Thus, the local utility functions $U(\cdot;\mathbf{P})$ and $U^*(\cdot;\mathbf{P})$ cannot be identical up to an affine transformation. ∎

To my knowledge, *no* researcher (including myself) has ever claimed that two von Neumann–Morgenstern utility functions representing the same underlying preference ranking should be affinely equivalent *when their outcome variables are related by a nonlinear transformation*[12] and I have certainly never claimed this for *local* utility functions either.[13] In fact, when the preference functionals $H(\cdot)$ and $H^*(\cdot)$ are related by a transformation of the outcome variable, the *very issue* of whether their local utility functions must satisfy an identity of the form

$$(20) \quad U^*(z;\mathbf{P}) \underset{z}{\equiv} a\cdot U(z;\mathbf{P}) + b$$

is a meaningless question, since as Counterexample 1 illustrates, the functions $U^*(\cdot;\mathbf{P})$ and $U(\cdot;\mathbf{P})$ will generally be defined over different domains (and in the case of our counterexample, the *completely disjoint* intervals [2, 3] and [16, 81]).

To derive the *correct* relationship between the local utility functions $U(\cdot;\mathbf{P})$ and $U^*(\cdot;\mathbf{P})$, we begin with the identity linking the preference functionals $H(\cdot)$ and $H^*(\cdot)$, namely

$$(21) \quad \begin{aligned} &H^*(y_1,\ldots,y_n,p_1,\ldots,p_n) \\ &\quad \equiv H(f^{-1}(y_1),\ldots,f^{-1}(y_n),p_1,\ldots,p_n) \end{aligned} \qquad (12,13)$$

At any distribution $\mathbf{P} = (y_1, \dots, y_n, p_1, \dots, p_n)$ and for any outcome level y_i in its support $\{y_1, \dots, y_n\}$, we accordingly have:

$$\begin{aligned} U^*(y_i; \mathbf{P}) &= \frac{\partial H^*(y_1, \dots, y_n, p_1, \dots, p_n)}{\partial p_i} \\ &= \frac{\partial H(f^{-1}(y_1), \dots, f^{-1}(y_n), p_1, \dots, p_n)}{\partial p_i} \qquad (22) \\ &= \frac{\partial H(\hat{\mathbf{P}})}{\partial \text{prob}(f^{-1}(y_i))} = U(f^{-1}(y_i); \hat{\mathbf{P}}) \end{aligned}$$

where $\hat{\mathbf{P}} = (f^{-1}(y_1), \dots, f^{-1}(y_n), p_1, \dots, p_n)$ is the probability distribution of the variable $x = f^{-1}(y)$ when the variable y has the distribution $\mathbf{P} = (y_1, \dots, y_n, p_1, \dots, p_n)$. For any outcome level y *not* in $\{y_1, \dots, y_n\}$, we similarly have

$$\begin{aligned} U^*(y; \mathbf{P}) &= \left. \frac{\partial H^*(y_1, \dots, y_n, y, p_1, \dots, p_n, p_y)}{\partial p_y} \right|_{p_y=0} \\ (23) \quad &= \left. \frac{\partial H(f^{-1}(y_1), \dots, f^{-1}(y_n), f^{-1}(y), p_1, \dots, p_n, p_y)}{\partial p_y} \right|_{p_y=0} \\ &= \left. \frac{\partial H(\hat{\mathbf{P}})}{\partial \text{prob}(f^{-1}(y))} \right|_{\text{prob}(f^{-1}(y))=0} = U(f^{-1}(y); \hat{\mathbf{P}}) \end{aligned}$$

Thus, at any distribution $\mathbf{P}$, we have that the local utility functions of $H(\cdot)$ and $H^*(\cdot)$ are linked by the relationship

$$U^*(y; \mathbf{P}) \underset{y}{\equiv} U(f^{-1}(y); \hat{\mathbf{P}}) \quad \text{or more generally}$$

$$(24) \quad U^*(y; \mathbf{P}) \underset{y}{\equiv} a \cdot U(f^{-1}(y); \hat{\mathbf{P}}) + b \quad (a > 0)$$

for all values of y, whether or not they lie in the support of $\mathbf{P}$.[14] This – as it must be – is a generalization of the relationship $U^*(y) \equiv U(f^{-1}(y))$ exhibited by the von Neumann–Morgenstern utility functions $U(\cdot)$ and $U^*(\cdot)$ in Counterexample 1.

IV. INVALIDITY OF DERIVING THE LOCAL UTILITY FUNCTION FROM THE MOMENT REPRESENTATION FUNCTION

Allais' second (and separate) error consists in asserting that, under the moment representation

$$(8,10) \quad \Phi = H(x_1, \dots, x_n, p_1, \dots, p_n) \equiv G(M_1, \dots, M_{2n-1}) (1,5)$$

the differential expansion

$$(16)\qquad \begin{aligned} \mathrm{d}\Phi &= \sum_{i=1}^{n}\left[\sum_{k=1}^{2n-1}\frac{\partial G(M_1,\ldots,M_{2n-1})}{\partial M_k}\cdot\frac{\partial M_k}{\partial p_i}\right]\cdot \mathrm{d}p_i \\ &= \sum_{i=1}^{n}\left[\sum_{k=1}^{2n-1}\frac{\partial G(M_1,\ldots,M_{2n-1})}{\partial M_k}\cdot x_i^k\right]\cdot \mathrm{d}p_i \end{aligned}\qquad (19)$$

which properly pertains only to outcomes $\{x_1,\ldots,x_n\}$ in the support of $\mathbf{P} = (x_1,\ldots,x_n,p_1,\ldots,p_n)$ and their associated probabilities $\{p_1,\ldots,p_n\}$, *also* applies for general outcome values $x \in [0,M]$ outside the support of $\mathbf{P}$, so that (according to Allais) we would have the local utility function formula

$$(17)\qquad \begin{aligned} U(x;\mathbf{P}) &= \frac{\partial H(\mathbf{P})}{\partial \mathrm{prob}(x)} = \frac{\partial G(M_1,\ldots,M_{2n-1})}{\partial \mathrm{prob}(x)} \\ &= \sum_{k=1}^{2n-1}\frac{\partial G(M_1,\ldots,M_{2n-1})}{\partial M_k}\cdot x^k \end{aligned}\qquad (20)$$

for *all* outcome levels $x \in [0,M]$. This is not correct, as shown by the following counterexample:

COUNTEREXAMPLE 2. To see that *something* must be wrong with formula (17), recall that since the $2n-1$ partial derivatives $\{\partial G(M_1,\ldots,M_{2n-1})/\partial M_1,\ldots,\partial G(M_1,\ldots,M_{2n-1})/\partial M_{2n-1}\}$ are each independent of x, it implies that, regardless of the nature of the preference functional $H(\cdot)$, its local utility function $U(x;\mathbf{P})$ must be a $2n-1$ degree polynomial in x. Thus, for example, equation (17) states that the local utility function of the *expected utility* preference functional

$$(25)\qquad H(x_1,\ldots,x_n,p_1,\ldots,p_n) \equiv \sum_{i=1}^{n} U(x_i)\cdot p_i \equiv \sum_{i=1}^{n} e^{x_i}\cdot p_i$$

which we know to be precisely the exponential function

$$(26)\qquad U(x;\mathbf{P}) \equiv U(x) \equiv e^{x}$$

must be a $2n-1$ degree polynomial! ■

Allais' erroneous derivation of equation (17) stems from the following oversight: Recall that the general formula for the local utility function at an outcome level x outside the support of $\mathbf{P}$ is given by the probability derivative

$$(27)\quad U(x;\mathbf{P}) = \left.\frac{\partial H(x_1,\ldots,x_n,x,p_1,\ldots,p_n,p_x)}{\partial p_x}\right|_{p_x=0}$$

Note that if the distribution $\mathbf{P} = (x_1,\ldots,x_n,p_1,\ldots,p_n)$ is an n-point distribution and x is *not* in the support of $\mathbf{P}$, then the evaluation of the derivative in (27) requires knowledge of the function $H(\cdot)$ *over the set of* $(n+1)$*-point distributions with support* $\{x_1,\ldots,x_n,x\}$.[15] But this means that the moment representation function $G(M_1,\ldots,M_{2n-1})$ of equation (10), which is only valid up to n-point distributions, *cannot be used to obtain the local utility representation* (17) *for general outcome values* $x \in [0,M]$ *outside the support of any n-point distribution* $\mathbf{P}$.

Although this oversight renders Allais' formula for $U(x;\mathbf{P})$ invalid for general outcome values $x \in [0,M]$ for any n-point distribution $\mathbf{P}$, one might argue that formula (17), involving the function $G(M_1,\ldots,M_{2n-1})$, could be valid for general outcome values $x \in [0,M]$ at any distribution $\mathbf{P}$ whose support contains *fewer* than n points.[16] However, this would involve a second oversight, namely the fact that *the moment representation function* $G(M_1,\ldots,M_{2n-1})$ *is not generally differentiable at any set of moments* $\{M_1,\ldots,M_{2n-1}\}$ *that correspond to a distribution* $\mathbf{P}$ *with a less than n-point support.* Thus, an application of the chain rule on equation (10) to obtain equation (17) would not be valid in this case. The following counterexample demonstrates this:

COUNTEREXAMPLE 3. To see that the moment representation function $G(M_1,\ldots,M_{2n-1})$ is not generally differentiable at the moments of any probability distribution with a support of fewer than n points, it is necessary to go no farther than *Allais' own* algebraic illustration for the case $n = 2$ (Allais, 1988, pp. 383–385). In this case, the moment representation of any probability distribution $\mathbf{P} = (x_1,x_2,p_1,p_2)$ involves the first $2n-1 = 3$ moments $\{M_1,M_2,M_3\}$. Allais shows that if the preference functional is given by

$$(28)\quad \Phi = H(x_1,x_2,p_1,p_2) = M_4 = x_1^4\cdot p_1 + x_2^4\cdot p_2 \qquad (1^*, 18^*, 32)$$

then the function $G(M_1,M_2,M_3)$ that satisfies the identity

$$(29)\quad H(x_1,x_2,p_1,p_2) \equiv G(M_1,M_2,M_3)$$

is given by the formula

$$(30)\quad G(M_1, M_2, M_3) \equiv \frac{M_2^3 + M_3^2 - 2M_1M_2M_3}{M_2 - M_1^2} \qquad (6^*)$$

Allais neglects to define the value of the function $G(M_1, M_2, M_3)$ at any set of moments of the form $(M_1, M_2, M_3) = (z, z^2, z^3)$, which correspond to any probability distribution $\mathbf{P} = (z, 1)$ with one-point support (note that for such a distribution, both the numerator and the denominator of his formula for $G(M_1, M_2, M_3)$ will be zero). However, from equations (28) and (29) it follows that $G(z, z^2, z^3)$ must take the value

$$(31)\quad G(z, z^2, z^3) \equiv z^4 \qquad \text{for all } z \in [0, M]$$

To demonstrate that $G(M_1, M_2, M_3)$ is not differentiable at points of the form $(M_1, M_2, M_3) = (z, z^2, z^3)$, it suffices to show that its directional derivatives are not linear in the direction vector. Consider the following three direction vectors out of a point $\vec{\mathrm{M}} = (M_1, M_2, M_3)$:

$$(32)\quad \begin{array}{l} \text{Direction } A : \vec{\Delta}^A \equiv (\Delta M_1, \Delta M_2, \Delta M_3) = (1, 0, 1) \\ \text{Direction } B : \vec{\Delta}^B \equiv (\Delta M_1, \Delta M_2, \Delta M_3) = (0, 1, 0) \\ \text{Direction } C : \vec{\Delta}^C \equiv (\Delta M_1, \Delta M_2, \Delta M_3) = (1, 1, 1) \end{array}$$

Direction C is clearly the vector sum of Directions A and B. Thus, *if* the function $G(M_1, M_2, M_3)$ were differentiable at the point $\vec{\mathrm{M}} = (M_1, M_2, M_3)$ – in other words, *if* Allais' use of the chain rule to derive equation (16) from equation (10) were valid – we ought to have the following relationship among the three directional derivatives:

$$(33)\quad \left.\frac{\mathrm{d}G(\vec{\mathrm{M}} + \vec{\Delta}^C \cdot t)}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}G(\vec{\mathrm{M}} + \vec{\Delta}^A \cdot t)}{\mathrm{d}t}\right|_{t=0} + \left.\frac{\mathrm{d}G(\vec{\mathrm{M}} + \vec{\Delta}^B \cdot t)}{\mathrm{d}t}\right|_{t=0}$$

In order to evaluate these three directional derivatives at the point $\vec{\mathrm{M}} = (M_1, M_2, M_3) = (1, 1, 1)$ (the moments of the one-point distribution $\mathbf{P} = (1, 1)$), we evaluate the functions:

$$G(\vec{\mathrm{M}} + \vec{\Delta}^A \cdot t) = G(1 + t, 1, 1 + t)$$

$$
\begin{aligned}
&= \frac{1+(1+t^2)-2(1+t)^2}{1-(1+t)^2} \\
&= 1 \\
G(\vec{\mathbf{M}}+\vec{\Delta}^B\cdot t) &= G(1,1+t,1) \\
&= \frac{(1+t)^3+1-2(1+t)}{(1+t)-1} \qquad (34)\\
&= 1+3t+t^2 \\
G(\vec{\mathbf{M}}+\vec{\Delta}^C\cdot t) &= G(1+t,1+t,1+t) \\
&= \frac{(1+t)^3+(1+t)^2-2(1+t)^3}{(1+t)-(1+t)^2} \\
&= 1+t
\end{aligned}
$$

which imply directional derivative values of

$$
\begin{aligned}
&\left.\frac{\mathrm{d}G(\vec{\mathbf{M}}+\vec{\Delta}^A\cdot t)}{\mathrm{d}t}\right|_{t=0}=0 \qquad \left.\frac{\mathrm{d}G(\vec{\mathbf{M}}+\vec{\Delta}^B\cdot t)}{\mathrm{d}t}\right|_{t=0}=3 \qquad (35)\\
&\text{and} \quad \left.\frac{\mathrm{d}G(\vec{\mathbf{M}}+\vec{\Delta}^C\cdot t)}{\mathrm{d}t}\right|_{t=0}=1
\end{aligned}
$$

which do not satisfy the adding-up condition (33). ∎

In fact, the problem is even more serious than this counterexample suggests. Recall that Allais' 'theorem' depends crucially upon the argument that, at any probability distribution P, each of the partial derivatives $\partial G(M_1,\ldots,M_{2n-1})/\partial M_k$ in equation (17) is a well-defined value. In his algebraic illustration for $n=2$ (so $2n-1=3$), equation (17) takes the form

$$
\begin{aligned}
U(x;\mathbf{P}) = {} & \frac{\partial G(M_1,M_2,M_3)}{\partial M_1}\cdot x+\frac{\partial G(M_1,M_2,M_3)}{\partial M_2}\cdot x^2 \\
& +\frac{\partial G(M_1,M_2,M_3)}{\partial M_3}\cdot x^3
\end{aligned} \qquad (36) \qquad (20^*)
$$

At any set of moments (M_1,M_2,M_3), the partial derivative $\partial G(M_1, M_2,M_3)/\partial M_3$, *if it exists*, is given by the formula

$$
\frac{\partial G(M_1,M_2,M_3)}{\partial M_3} \overset{\text{def}}{=} \left.\frac{\mathrm{d}G(M_1,M_2,M_3+t)}{\mathrm{d}t}\right|_{t=0} \qquad (37)
$$

However, for the Allais moment representation function

$$(30) \quad G(M_1, M_2, M_3) \equiv \frac{M_2^3 + M_3^2 - 2M_1M_2M_3}{M_2 - M_1^2} \qquad (6^*)$$

we have

$$(38) \quad \begin{aligned} G(z, z^2, z^3 + t) &= \frac{(z^2)^3 + (z^3 + t)^2 - 2(z)(z^2)(z^3 + t)}{(z^2) - (z)^2} \\ &= \frac{z^6 + z^6 + 2z^3t + t^2 - 2z^6 - 2z^3t}{z^2 - z^2} = \frac{t^2}{0} \end{aligned}$$

But since $G(z, z^2, z^3 + t)$ equals positive infinity for any nonzero value of t, the partial derivative $\partial G(M_1, M_2, M_3)/\partial M_3$ in formula (37) (and hence in formula (36)) is *not well-defined* at the moments $(M_1, M_2, M_3) = (z, z^2, z^3)$ of any one-point distribution $\mathbf{P} = (z, 1)$.

To summarize the nature of Allais' second error: The only way to use the moment representation function $G(M_1, \ldots, M_{2n-1})$ to obtain the local utility function $U(x; \mathbf{P})$ at a general outcome level x *not* in the support of $\mathbf{P}$ would be if the distribution $\mathbf{P}$ has a less than n-point support. However, at such distributions, the function $G(M_1, \ldots, M_{2n-1})$ is not necessarily differentiable, in that its directional derivatives are not necessarily linear in the direction vectors, and some partial derivatives may not even exist. Accordingly, Allais' use of the chain rule on this function in order to derive his local utility formula (17) is invalid,[17] and leads to incorrect derivations of the local utility function in even the simplest of cases, as can be seen for the simple finite-outcome expected utility example presented in Counterexample 2.

ACKNOWLEDGMENTS

I am grateful to Joel Sobel and Max Stinchcombe for helpful discussions on this material, and to the National Science Foundation (Grant SES 92-09012) for financial support.

APPENDIX

Reader's Note on some typos in Allais (1988): In the manuscript version of Allais (1988), dated June 5, 1986, his preference functions '$H(\cdot)$' and '$H^*(\cdot)$' originally appeared as '$F(\cdot)$' and '$F^*(\cdot)$' throughout. In converting this notation for the published version, some changes were apparently missed, so that in Allais (1988): the symbols 'F' on page 378 (his equation (12)) and on page 383 (his equation (1*)) should both properly read 'H', and the symbol 'F^*' on page 378 (following his equation (11)) should properly read 'H^*'. In citing the two equations, I have corrected the notation.

NOTES

[1] This paper is about mathematical issues only. The question of the 'true nature of utility' is not addressed here.

[2] That is, averse to all mean preserving increases in risk, in the sense of Rothschild and Stiglitz (1970).

[3] See Machina (1982, pp. 298–300; 1984, pp. 203–205) for the formal list of properties that this implies.

[4] The equivalence of this inequality to a concave transformation of $U(\cdot)$ is shown in Pratt (1964, Thm. 1).

[5] Actually, we must impose one more condition on the vector $(\mathrm{d}p_1, \ldots, \mathrm{d}p_n)$ to ensure that $(p_i + \mathrm{d}p_1, \ldots, p_n + \mathrm{d}p_n)$ is a set of actual probabilities, namely, if $p_i = 0$ then $\mathrm{d}p_i \geqslant 0$. This condition will be understood to hold throughout.

[6] Throughout both Allais' formulation (see Allais, 1988, p. 359 top) and our own, all differential expansions $\mathrm{d}V(\mathbf{P})$ (and later, $\mathrm{d}\Phi$) are with respect to the probabilities alone, not the outcome levels.

[7] More formal (and more general) versions of the results (a′/a″), (b′/b″) and (c′/c″), along with their proofs, appear as Theorems 1, 2 and 4 of Machina (1982).

[8] Equation numbers at the right margin refer to Allais (1988, Appendix A, pp. 377–386). Unless otherwise stated, all references to equations are to my own equation numbers, which will continue to appear at the left margin.

[9] In keeping with Allais' notation, I use $f^k(x_i)$ to denote the kth power of the value $f(x_i)$.

[10] In Sections III and IV, I show that each of these first two steps is invalid, so that the third step is moot.

[11] Allais denotes local utility functions by '$\P(x;p)$' in the text of his paper and by '$\pi(x;p)$' in his equations. For consistency with the original notation used in Machina (1982) and in subsequent articles by myself and others, I replace his terms $\P(x;p)$ and $\pi(x;p)$ by the standard term $U(x;\mathbf{P})$ in the following cited equations and text.

[12] An even simpler example: Typically, *after-tax income* y is a nonlinear (say concave) function of *before-tax income* x. It is obvious that if I am risk neutral in after-tax income, then I *cannot* be risk neutral in before-tax income, or vice versa. Even if I'm not risk neutral, my utility function over y should not be expected to be an affine transformation of my utility function over x. Nor should any of my local utility functions.

[13] My discussion of the cardinal properties of local utility functions is given in Machina (1988).

[14] Although the terms a and b in (24) will be independent of the outcome value y (or $f^{-1}(y)$), they can depend upon the probability distribution $\mathbf{P}$ (or $\hat{\mathbf{P}}$), in a manner made precise in Machina (1988).

[15] This is true even though we are *evaluating* this derivative at the n-point distribution $\mathbf{P}$, where $p_x = 0$. An analogy is that for any univariate function $h(\cdot)$, evaluating the derivative $h'(0)$ requires knowledge of $h(z)$ *over non-zero values of* z in a neighborhood of 0.

[16] On this point: In a 'Postscript' to Allais (1988) (pp. 398–403), Allais reminds us that the proof of his Impossibility Theorem is based on the representation of an n-point distribution by its first $2n - 1$ moments, and states that this "assumes that all the x_i are different from each other and that no p_i is nil. If one of these conditions is not met, there is degenerescence and the distribution considered is lower order than n" (pp. 400–401, P.1.3, a.,c.). It is not clear to me what this statement means. If it means that the function $G(M_1, \ldots, M_{2n-1})$ (and hence formula (17)) *cannot be used* to represent preferences over distributions with less than n-point support, the invalidity of formula (17) is then fully established by the argument of the previous paragraph, and the argument of the rest of this Section is not necessary. Otherwise, the argument of the rest of the Section establishes this invalidity.

[17] Of course, it follows that Allais' derivation of the local utility function $U^*(\cdot; \mathbf{P})$ of the preference functional $H^*(\cdot)$ (equation (19)) is similarly invalid.

REFERENCES

Allais, M.: 1986, 'Machina's Local Utility and the Allais Impossibility Theorem – Additional Comments', C.A.E. Manuscript C-4384.

Allais, M.: 1988, 'A New Neo-Bernoullian Theory: The Machina Theory. A Critical Analysis', in *Risk, Decision and Rationality*, edited by B. Munier, D. Reidel Publishing Co., Dordrecht.

Allen, B.: 1987, 'Smooth Preferences and the Local Expected Utility Hypothesis', *Journal of Economic Theory* **41**(2), 340–355.

Chew, S., Epstein, L., and Zilcha, I.: 1988, 'A Correspondence Theorem between Expected Utility and Smooth Utility', *Journal of Economic Theory* **46**(1), 186–193.

Chew, S., E. Karni and Z. Safra: 1987, 'Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities', *Journal of Economic Theory* **42**(2), 370–381.

Chew, S. and Nishimura, N.: 1992, 'Differentiability, Comparative Statics, and Non-Expected Utility Preferences', *Journal of Economic Theory* **56**(2), 294–312.

Green, J. and Jullien, B.: 1988, 'Ordinal Independence in Non-Linear Utility Theory', *Journal of Risk and Uncertainty* **1**(4), 355–387.

Karni, E.: 1987, 'Generalized Expected Utility Analysis with State-Dependent Preferences', *International Economic Review* **28**(1) 229–240.

Karni, E.: 1989, 'Generalized Expected Utility Analysis of Multivariate Risk Aversion', *International Economic Review* **30**(2), 297–305.

Machina, M.: 1982, ' "Expected Utility" Analysis without the Independence Axiom', *Econometrica* **50**(2), 277–323.

Machina, M.: 1983, 'Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom', in *Foundations of Utility and Risk Theory with Applications*, edited by B. Stigum and F. Wenstøp, D. Reidel Publishing Co., Dordrecht.

Machina, M.: 1984, 'Temporal Risk and the Nature of Induced Preferences', *Journal of Economic Theory* **33**(2), 199–231.

Machina, M.: 1987, 'Choice under Uncertainty: Problems Solved and Unsolved', *Journal of Economic Perspectives*, **1**(1), 121–154.

Machina, M.: 1988, 'Cardinal Properties of "Local Utility Functions" ', in *Risk, Decision and Rationality*, edited by B. Munier, D. Reidel Publishing Co., Dordrecht.

Machina, M. and Neilson, W.: 1987, 'The Ross Measure of Risk Aversion: Strengthening and Extension', **55**(5), 1139–1149.

Pratt, J.: 1964, 'Risk Aversion in the Small and in the Large', *Econometrica* **32**(1–2), 122–136.

Röell, A.: 1986, 'Portfolio Choice in Quiggin and Yaari's Rank-Order Model of Choice under Uncertainty', manuscript, London School of Economics.

Rothschild, M. and Stiglitz, J.: 1970, 'Increasing Risk: I. A Definition', *Journal of Economic Theory* **2**(3), 225–243.

Wang, T.: 1993, 'L_p-Fréchet Differentiable Preference and "Local Utility" Analysis', *Journal of Economic Theory* **61**(1), 139–159

Department of Economics,
University of California, San Diego,
La Jolla, CA 92093, U.S.A.