TEL-AVIV

THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH

TEL-AVIV UNIVERSITY

RAMAT AVIV ISRAEL

מכון למחקר כלכלי ע״ש ד״ר ישעיהו פורדר ז״ל

ע״י אוניברסיטת תל־אביב

EXPECTATION AND VARIATION IN LONG RUN DECISIONS*

by

Itzhak Gilboa**

Working Paper No.10-87
January, 1987

* I wish to thank Professor David Schmeidler for his guidance. I am also grateful to Professor Andrew MasColell and Professor Elhanan Helpman for comments and references.

** Department of Economics, Tel-Aviv University

This research was supported by funds granted to the Foerder Institute for Economic Research by the **NUR MOSHE FUND**

THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH
Faculty of Social Sciences, Tel-Aviv University
Tel-Aviv, 69978, Israel.

ABSTRACT

Long-run decisions are decisions which determine the individual's payoffs in several periods in the future. This paper examines the theoretical foundations of the prevalent "weighted average" assumption, and suggests a larger class of decision rules, which take into account the effects of the payoffs variation.

The "weighted average" assumption is a special case of the generalized model, a case in which the decision maker is variation neutral. Similarly, we define and characterize variation aversion and variation-liking, and show an example of the economic implications of these properties.

LONG-RUN DECISIONS

by: Itzhak Gilboa

1. Introduction

There is a large class of economic problems in which a decision maker is asked to choose one alternative out of a choice set, each of the elements of which determines his payoff (or expected payoff) in several periods in the future. We shall refer to these problems as "long-run problems".

Examples in which this structure is explicit are, for instance, models of investment, labor planning and all problems which may be formulated as repeated games. There are, however, many more examples in which the same structure is implicitly assumed. In fact, it seems that one can hardly think of a "real life" decision problem (whether under certainty or uncertainty) that may be satisfactorily represented as a single-period problem. (In Savage (1954), for instance, the individual's single choice is among acts which provide, for each state of nature, a complete description of the aspects of the world relevant to him - at **any** point of time in the future.)

It seems that the vast majority of the economic literature assumes that there exists an "instantaneous utility function" u and a long-run function U such that each alternative $f = (f_1, f_2, \ldots)$ is assessed by $U(u(f_1), u(f_2), \ldots)$. In many cases, no restrictions are set upon U (apart from monotonicity, quasi-concavity and so forth). In most of the cases in which U's functional form is specified, it is assumed to be a weighted average $\Sigma_i p_i u(f_i)$. (Common weights are $p_i = (1-\beta)\beta^{i-1}$ for $0 < \beta < 1$ and $p_i = \frac{1}{T}1_{\{i \le T\}}$ for $T \ge 1$.)

Such a functional is discussed and axiomatized in Koopmans (1960) and Koopmans, Diamond and Williamson (1964). A more general (non-separable) functional is axiomatized in Kreps and Porteus (1979), and non-separable preferences are also used in Lucas and Stokey (1984).

The purpose of this paper is to examine the theoretical foundations of the separable functional form from a different viewpoint, and to provide an axiomatization of a slightly more general form, which is quite simple and seems to us more suitable for some of the possible applications.

2. Motivation

It is well known that Savage's formulation of the decision-under-uncertainty problem may be interpreted in other ways as well: In Savage (1954) an **act** is a function from the **states of the world** into the set of **consequences**. But if one chooses to replace the states of the world by individuals in a population, the resulting problem is a social choice problem; if the states are replaced by criteria and the consequences by grades - we obtain a model of a multi-criteria decision problem; and if the "states of the world" are interpreted as points of time - we end up with a long-run decision problem.

It follows then, that Savage (1954) - and Anscombe-Aumann (1963) - provide axiomatic foundations for the equivalents of the Expected Utility paradigm in the other contexts, that is to say, for the hypothesis that the relevant functional U is a weighted average of $\{u(s)\}_{s\in S}$ where S is the domain on which acts are defined.

However, even in their "native" interpretation - namely, in the realm of decisions under uncertainty - the classical models have been seriously attacked. (See, for instance, Allais (1952) and Ellsberg (1961).) It is, therefore, our task to examine their axioms carefully before applying them to any other field of decision theory.

Let us first consider a simple example: A decision maker is faced with a choice problem, and his decision determines his payoffs in the next four periods. At each period his payoff may be either high (H) or low (L). He has four alternatives which are (H,H,L,L), (H,L,H,L), (L,H,L,H) and (L,L,H,H). Suppose that in some sense (which is quite vague to us at this point) the "intrinsic value" of all the periods is constant. However, there is a certain "cost of adjustment" incurred by any change in the payoff level. For example, the payoffs may be the standard of living levels, and the costs - the socio-psychological costs of changing the social status. One may also think of a firm, for which the payoffs are the revenue (and profit) levels, the variation of which involve some organizational cost. In these cases it seems plausible that the decision maker's preference relation $\geq$ would satisfy

$$(H,H,L,L) \sim (L,L,H,H) > (H,L,H,L) \sim (L,H,L,H)$$

(where $>$ means strict preference and $\sim$ means equivalence.).

It is not hard to see that such preferences do not comply with the weighted average hypothesis. That is to say, there do not exist a utility function $u = (u(H), u(L))$ and a measure $p = (p_1, p_2, p_3, p_4)$ such that the preferences discussed above may be explained by maximization of the integral of u w.r.t. (with respect to) p.

In fact, this example is mathematically equivalent to the famous Ellsberg Paradox (Ellsberg (1961)), which challenged Savage's Sure-Thing Principle. We shall now turn to discuss this Principle.

3. The Sure-Thing Principle

Roughly speaking, the Principle (axiom P2 in Savage (1954)) says that, if two possible acts yield the same consequences whenever an event A occurs, the preferences between them should be determined only by the values they assume outside A. Or, formally, if f, g, f' and g' are four acts (functions from the set of states of the world S to the set of consequences X), and

$$f(s) = g(s) \qquad f'(s) = g'(s) \qquad s \in A$$
$$f(s) = f'(s) \qquad g(s) = g'(s) \qquad s \in A^c$$

- then $f \geq g$ iff $f' \geq g'$.

In the original interpretation of the model, the Principle seems almost unobjectionable: By definition, exactly one of all possible states of the world actually obtains, hence there cannot eventually be any situation in which the decision maker is affected by the consequences attached by his act to other states of the world.

Of course, the Principle has proved to be objectionable after all. (This is an empirical fact.) However, we claim that in other interpretations its foundations are considerably weaker to start with. Let us consider, for example, the social choice interpretation of the same decision problem discussed in section 2. Assume H and L represent high and low salaries, respectively, and let the domain be (men in city A; women in city A; men in city B; women in city B). It is perfectly reasonable to assume that a central planner should have no bias towards either city, nor towards either sex, but would strictly prefer that the poorly-paid employees will not concentrate in any single city.

In the context of long-run decisions we have already seen that the Sure Thing Principle may not be as compelling an axion as it purports to be in its original context. There is, however, a slight modification of it which seems to be a reasonably sound foundation for our theory.

4. Variation-Preserving Sure-Thing Principle

In the context of long-run decisions, as opposed to the other ones, there is a natural linear ordering on the domain of the acts: time points (or periods) are ordered by their very definition, while states of the world and individuals in a society are not. This additional structure imposed on our model allows us to reject the Sure Thing Principle as formulated, without renouncing the gist of its essence.

Let us consider the example of section 2 once more. The prospect (H,H,L,L) is preferred to (L,H,L,H), but if we replace the payoffs in periods 2 and 3 by (L,H), we obtain the prospects (H,L,H,L) and (L,L,H,H) respectively, with the latter preferred to the former. One may observe that the replacement of (H,L) by (L,H) in periods 2 and 3 is "biased" in a certain sense: it increases the variation of (H,H,L,L) but decreases that of (L,H,L,H). If the variation of the acts should play any role in our theory, there is no reason to wonder at the preference reversal. But if we restrict ourselves to such changes which **do not** affect the variation asymmetrically, we may expect the Principle to hold.

A simple way in which we can assure that changing f to f′ will have the same effect on the variation of f as changing g to g′ will have on that of g - is to restrict the scope of discussion to changes over time intervals on the edges of which f and g coincide. The weaker axiom which results will be called **Variation-Preserving Sure Thing Principle**. It is easy to see that it allows the "preference reversal" of section 2, since the time interval under discussion (periods 2 and 3) does not satisfy our additional condition: on its edges (periods 1 and 4) the two relevant acts ((H,H,L,L) and (L,H,L,H)) do **not** coincide.

5. A Description of the Model and the Results

We shall use the framework of Anscombe-Autmann (1963) rather than that of Savage (1954), since it allows for a finite domain and a denumerable one with a continuous measure. Our whole discussion will, as a matter of fact, be restricted to these two cases. (The generalization to continuous time, for instance, meets the difficulties of axiomatizing measurability and continuity of various function which are endogenous in the model. The author is not aware of any set of reasonably intuitive axioms which may guarantee the desired technical properties.)

Since we reject the Sure Thing Principle, we cannot adhere to Anscombe-Aumann's model; we therefore turn to its generalization suggested by Schmeidler (1982 and 1986) using the concept of non-additive measures. On this basic model we impose the Variation-Preserving Sure Thing Principle and, roughly speaking, we obtain the following result:

There are a utility function u, and two functions p, δ on the set of periods $\{s_1, s_2, \ldots, s_n\}$, such that the preference relation is represented by the functional

$$U(f) = \Sigma_{i=1}^{n} p(s_i)u(f(s_i)) + \delta(s_i) \mid u(f(s_i)) - u(f(s_{i-1})) \mid .$$

That is to say, there exists an "intrinsic value" $p(s_i)$ to each period s_i, and the first element in the summation is the expected utility w.r.t. the measure $p(\cdot)$. The second element for each period s_i is an extra cost/bonus incurred by the mere variation of the function $u(f(\cdot))$.

A formal formulation and a proof of the representation theorem are to be found in section 6. The extension to an infinite horizon is contained in section 7. A by-product of this section is a characterization of continuous measures in Schmeidler's model, and a proof that the Choquet Integral (See Choquet (1955)) is continuous in the appropriate sense.

Given the functional form (for a finite or infinite horizon), one may ask what is an individual's attitude towards variation. Indeed, it turns out that we may define and characterize the notions of variation aversion, variation neutrality and variation liking. This is done in section 8.

Finally, section 9 provides an example of the implications of our model. It shows that even if two "identical" individuals play a repeated zero-sum game, the super-game need not be zero-sum; that is to say, such two individuals may have a positive surplus of cooperation.

6. The Model and the Finite-Horizon Representation Theorem

6.1. Schmeidler's model and result

Let X be a non-empty set of **consequences** and let Y denote the set of finite-support distributions over X ("lotteries"):

$$Y = \{y:X \to [0,1] \mid y(x) \neq 0 \text{ for finitely many } x\text{'s in } X \text{ and } \Sigma_{x\in X}\, y(x) = 1\}.$$

Let S be a nonempty set of **points of time** and let Σ be an algebra of subsets of S. Let F denote the set of **acts**, which is a subset of the functions from S to Y, including all the constant functions. We assume that a binary (preference) relation $\geq$ is given on F, such that F is exactly the set of all Σ-measurable bounded functions with respect to $\geq$: Given $\geq \subseteq F \times F$ we define $\geq \subseteq Y \times Y$ by identifying a lottery y in Y with the constant act which assumes the value y over all S. An act $f \in F$ is **Σ-measurable** if $\{s \mid f(s) > y\}$, $\{s \mid f(s) \geq y\} \in \Sigma$ for all $y \in Y$. We may then assume that

$$F = \{f: S \to Y \mid f \text{ is } \Sigma\text{-measurable and there are } \underline{y}, \bar{y} \in Y \text{ s.t. } \underline{y} \leq f(s) \leq \bar{y} \text{ for all } s \in S\}.$$

Linear operations are performed on F pointwise. Two acts $f, g \in F$ are **comonotonic** iff there are no $s, t \in S$ for which $f(s) > f(t)$ and $g(s) < g(t)$.

Schmeidler's axioms are:

A1. **Weak Order:** $\geq$ is complete, reflexive and transitive.

A2. **Comonotonic Independence**: If $f, g, h \in F$ are pairwise comonotonic and $\alpha \in (0,1)$, then $f \geq g$ iff $\alpha f + (1-\alpha)h \geq \alpha g + (1-\alpha)h$.

A3. **Continuity**: If $f, g, h \in F$ satisfy $f > g > h$ then there are $\alpha, \beta \in (0,1)$ such that

$$\alpha f + (1+\alpha)h > g > \beta f + (1-\beta)h.$$

A4. **Monotonicity**: If $f, g \in F$ satisfy $f(s) \geq g(s)$ for all $s \in S$, then $f \geq g$.

A5. **Nondegeneracy**: There are $f, g \in F$ such that $f > g$.

Axiom A2 deserves deliberation. In Anscombe-Aumann (1963), it is assumed in a stronger form, without the comonotonicity condition. Their theorem proves that such a preference relation is representable by Expected Utility maximization, hence it satisfies the Sure Thing Principle, which is too restrictive for our theory.

Indeed, their independence axiom states that the preferences between f and g will not be changed if they are both mixed with h, **even if** this mixing affects their variation asymmetrically (say, increases that of f but decreases that of g). However, if, in A2, we consider only triples of acts f, g, h which are pairwise comonotonic, their mixing cannot alter their variation in a biased way, and the weakened independence axiom which results is a reasonable one.

We now define a (non-additive) **measure** v on (S,Σ) to be a function $v: \Sigma \to [0,1]$ which satisfies:

(i) $v(\phi) = 0$; $v(S) = 1$

(ii) $A \subset B \subset S \Rightarrow v(A) \leq v(B)$.

For a Σ-measurable and bounded real function $\varphi: S \to R$, **the (Choquet) integral** of φ (on S) w.r.t. v is

$$\int_S \varphi dv = \int_{-\infty}^{0} [v(\{s \mid \varphi(s) > t\}) - 1]dt + \int_0^{\infty} v(\{s \mid \varphi(s) > t\})dt.$$

When no confusion is likely to arise, the subscript "S" will be omitted and we shall denote the integral simply by $\int \varphi dv$.

We now quote

Schmeidler's Theorem: $\geq$ satisfies A1-A4 iff there are an affine utility $u: Y \to \mathbb{R}$ and a measure v on (S,Σ) such that

$$f \geq g \iff \int u(f)dv \geq \int u(g)dv \quad \forall f,g \in F.$$

Furthermore, if A1-A5 hold, then u is unique up to a positive linear transformation (p.l.t.) and v is unique.

6.2. The model

We now introduce additional assumptions. First of all, we assume that there exists a linear order $\gg$ defined on the points of time. Next we assume (until section 7) that S is finite and that $\Sigma = 2^S$. W.l.o.g. (without loss of generality) we assume that $S = \{s_i\}_{i=1}^n$ where $s_{i+1} \gg s_i$ For $1 \leq i \leq n-1$. We extend $\gg$ to subsets of S as follows: For $A \subset S$ and $s \in S$, $s \gg (\ll)\ A$ if $s \gg (\ll)\ t$ for all $t \in A$. Similarly, for $A, B \subset S$, $B \gg (\ll)\ A$ if $B \gg (\ll)\ t$ for all $t \in A$. $A, B \subset S$ are **separated** if $\exists s \in S$ such that $A \ll s \ll B$ or $B \ll s \ll A$.

A subset $A \subset S$ is **an interval** if there are $1 \leq i, j \leq n$ such that $A = \{s_k \in S \mid i \leq k \leq j\}$. In this case A will also be denoted by $[s_i, s_j]$.

Unless otherwise stated, we shall assume that $\geq$ satisfies A1-A5 and refer to u and v provided by Schmeidler's theorem. Furthermore, w.l.o.g. we assume that $\sup\{u(y) \mid y \in Y\} > 1$ and $\inf\{u(y) \mid y \in Y\} < 0$, and for each $\alpha \in [0,1]$, we choose $y_\alpha \in Y$ such that $u(y_\alpha) = \alpha$.

For convenience, extend any act f to the domain $S \cup \{s_0, s_{n+1}\}$ by $f(s_0) = f(s_{n+1}) = y_0$ for all $f \in F$. (We assume also that $s_{n+1} \gg S \gg s_0$.)

We may finally formulate

A6. **Variation Preserving Sure Thing Principle**: Suppose hat $A = [s_i, s_j]$ with $1 \le i \le j \le n$, and that

$f, f', g, g' \in F$ satisfy

$$f(s) = f'(s) \qquad g(s) = g'(s) \quad \text{for} \quad s \in A^c$$

$$f(s) = g(s) \qquad f'(s) = g'(s) \quad \text{for} \quad s \in A$$

and

$$f(s_k) = g(s_k) = f'(s_k) = g'(s_k) \quad \text{for} \quad k = i-1,\ j+1.$$

- Then $f \ge g$ iff $f' \ge g'$.

The Main Theorem: $\ge$ satisfied A6 iff there are $p, \delta: S \to R$ such that

(i) $p(s_i) \ge |\delta(s_i)| + |\delta(s_{i+1})|$ for $i < n$,

$p(s_n) \ge |\delta(s_n)|$, and $\delta(s_1) = 0$

and

(ii) $\int u(f)dv = \Sigma_{i=1}^{n} p(s_i)u(f(s_i)) + \delta(s_i)\ |u(f(s_i)) - u(f(s_{i-1}))|$

for all $f \in F$.

Moreover, if A6 holds then p and δ are unique.

6.3. Proof of the theorem

Let us first assume A6 holds. We begin with

Lemma 1 Suppose $A, B \subset S$ are separated. Then

$$v(A \cup B) = v(A) + v(B)$$

Proof: Suppose w.l.o.g. that $A \ll s \ll B$. Now assume that there exists $C \subset S$ such that: (i) $C \gg s$; (ii) $C \cap B = \phi$; and (iii) $v(B \cup C) > v(B)$.

Let $\alpha = v(B)/v(B \cup C)$ and define $f_1, f_2, f_3, g_1, g_2, g_3 \in F$ as follows:

$$f_1(t) = \begin{cases} y_1 & t \in B \\ y_0 & \text{otherwise} \end{cases} \qquad g_1(t) = \begin{cases} y_\alpha & t \in B \cup C \\ y_0 & \text{otherwise} \end{cases}$$

$$f_2(t) = \begin{cases} y_1 & t \in B \\ y_\alpha & t \in A \\ y_0 & \text{otherwise} \end{cases} \qquad g_2(t) = \begin{cases} y_\alpha & t \in A \cup B \cup C \\ y_0 & \text{otherwise} \end{cases}$$

and

$$f_3(t) = \begin{cases} y_1 & t \in B \cup A \\ y_0 & \text{otherwise} \end{cases} \qquad g_3(t) = \begin{cases} y_1 & t \in A \\ y_\alpha & t \in B \cup C \\ y_0 & \text{otherwise.} \end{cases}$$

Since $\int u(f_1)dv = \int u(g_1)dv$, we obtain $f_1 \sim g_1$. By A6, $f_2 \sim g_2$ and $f_3 \sim g_3$ must also hold. Hence $\int u(f_2)dv = \int u(g_2)dv$ and $\int u(f_3)dv = \int u(g_3)dv$. The first equality implies

$$(1-\alpha)v(B) + \alpha v(A \cup B) = \alpha v(A \cup B \cup C)$$

and the second one yields

$$v(A \cup B) = (1-\alpha)v(A) + \alpha v(A \cup B \cup C).$$

Hence $\quad v(A \cup B) = (1-\alpha)v(A) + (1-\alpha)v(B) + \alpha v(A \cup B)$,

whence $\quad v(A \cup B) = v(A) + v(B)$.

We now turn to the case in which there does not exist a subset C as required. We define $S' = S \cup \{s^*\}$ and $\gg'$ on $S' \cup \{s_0, s_{n+1}\}$ by $s_{n+1} \gg' s^* \gg' s_n \gg' \ldots \gg' s_1 \gg' s_0$. For $A \subset S'$, let $v'(A) = v(A \cap S) + \epsilon \, 1_{\{s^* \in A\}}$ for a fixed $\epsilon > 0$. Now let $F' = \{f\colon S' \to Y\}$ and define $\geq'$ on F' by $\int_{S'} u(\cdot)dv'$. By Schmeidler's theorem, $\geq'$ on F' satisfies axioms A1-A5. However, it also satisfies A6 since

$$\int_{S'} u(f)dv' = \int_S u(f)dv + u(f(s^*)) \cdot \epsilon \quad \text{for} \quad f \in F'.$$

Considering A and B as subsets of S', there exists $C = \{s^*\}$ which satisfies our conditions. Therefore $v'(A \cup B) = v'(A) + v'(B)$. But $s^* \in (A \cup B)^c$ and our conclusion follows. □

By Lemma 1 we know that v is completely determined by its value on the intervals (since any $A \subset S$ is the disjoint union of finitely many separated intervals). As there are $\binom{n+1}{2}$ intervals, there are no more than $\binom{n+1}{2}$ degrees of freedom in specifying v. However, the next lemma shall prove that this upper bound is not the best one one may obtain:

Lemma 2: Let A and B be two intervals such that $A \cap B \neq \phi$. Then $v(A \cup B) + v(A \cap B) = v(A) + v(B)$.

Proof: If $A \subset B$ or $B \subset A$, the lemma is trivial. Assume, then, w.l.o.g. that $A = [s_i, s_j]$ and $B = [s_k, s_\ell]$ where $1 \leq i < k < j < \ell \leq n$. As in the previous lemma we first assume that there exists a subset $C \subset [s_{\ell+1}, s_n]$ such that $v(B \cap C) > v(B)$. In this case, let $\alpha \in [0,1)$ satisfy $v(B) = (1-\alpha)v(A \cap B) + \alpha v(B \cup C)$ and define f_1, f_2, f_3, g_1, g_2, $g_3 \in F$ by:

$$f_1(t) = \begin{cases} y_1 & t \in B \\ y_0 & \text{otherwise} \end{cases} \qquad g_1(t) = \begin{cases} y_1 & t \in A \cap B \\ y_\alpha & t \in (B-A) \cup C \\ y_0 & \text{otherwise} \end{cases}$$

$$f_2(t) = \begin{cases} y_1 & t \in B \\ y_\alpha & t \in A-B \\ y_0 & \text{otherwise} \end{cases} \qquad g_2(t) = \begin{cases} y_1 & t \in A \cap B \\ y_\alpha & t \in (B-A) \cup (A-B) \cup C \\ y_0 & \text{otherwise} \end{cases}$$

and

$$f_3(t) = \begin{cases} y_1 & t \in A \cup B \\ y_0 & \text{otherwise} \end{cases} \qquad g_3(t) = \begin{cases} y_1 & t \in A \\ y_\alpha & t \in (B-A) \cup C \\ y_0 & \text{otherwise} \end{cases}$$

Since $\int u(f_1)dv = v(B) = (1-\alpha)v(A\cap B) + \alpha v(B\cup C) = \int u(g_1)dv$, $f_1 \sim g_1$ and A6 implies that $f_2 \sim g_2$ and $f_3 \sim g_3$. By $\int u(f_2)dv = \int u(g_2)dv$ we obtain

$$(1-\alpha)v(B) + \alpha v(A\cup B) = (1-\alpha)v(A\cap B) + \alpha v(A\cup B\cup C)$$

and the equality $\int u(f_3)dv = \int u(g_3)dv$ yields

$$v(A\cup B) = (1-\alpha)v(A) + \alpha v(A\cup B\cup C).$$

Combining the equalities we get

$$(1-\alpha)v(A) + (1-\alpha)v(B) = (1-\alpha)v(A\cap B) + (1-\alpha)v(A\cup B)$$

where $\alpha < 1$.

In case no such event C exists, one may proceed as in Lemma 1 to complete the proof. □

Note that in view of this last lemma, v is completely determined by its value on intervals of length 1 and 2. Hence there are no more than (2n-1) degrees of freedom in specifying v. We shall now proceed to represent v in a simple way which will suggest an intuitive explanation of the (2n-1) parameters.

Lemma 3: There are functions $\hat{p}$, $\hat{\delta}$: $S \to R$ such that:

(i) $\hat{p}(s_i) \geq 0 \qquad \forall i \leq n$

(ii) $\hat{\delta}(s_1) = 0$

(iii) For any $1 \leq i \leq n$ and $\ell \geq 0$ such that $i + \ell \leq n$,

$$v([s_i, s_{i+\ell}]) = \hat{\delta}(s_i) + \Sigma_{k=0}^{\ell}\hat{p}(s_{i+k}).$$

Furthermore, these two functions are unique.

Proof: Let us first define the functions: set $\hat{\delta}(s_1) = 0$, $\hat{p}(s_1) = v(\{s_1\})$, and for $2 \leq i \leq n$ define

$$\hat{p}(s_i) = v([s_{i-1}, s_i]) - v(\{s_{i-1}\})$$

and

$$\hat{\delta}(s_i) = v(\{s_i\}) - \hat{p}(s_i).$$

It is obvious that $\hat{p}$ and $\hat{\delta}$ satisfy conditions (i) and (ii), and that they are the only pair of functions satisfying:

(1) $v(\{s_i\}) = \hat{\delta}(s_i) + \hat{p}(s_i) \qquad \forall i \leq n$

(2) $v([s_i, s_{i+1}]) = \hat{\delta}(s_i) + p(s_i) + p(s_{i+1}) \qquad \forall i \leq n\text{-}1.$

All we have to show is that $\hat{p}$ and $\hat{\delta}$ also satisfy condition (iii). We use induction on ℓ. For $\ell = 1$ and $\ell = 2$ we use (1) and (2) respectively. Assuming correctness for ℓ-1, lemma 2 yields

$$v([s_i, s_{i+\ell}]) = v([s_i, s_{i+\ell-1}]) + v([s_{i+\ell-1}, s_{i+\ell}]) -$$

$$v(\{s_{i+\ell-1}\}) = \hat{\delta}(s_i) + \Sigma_{k=0}^{\ell-1}\hat{p}(s_{i+k}) + \hat{\delta}(s_{i+\ell-1}) +$$

$$\hat{p}(s_{i+\ell-1}) + \hat{p}(s_{i+\ell}) - \hat{\delta}(s_{i+\ell-1}) - \hat{p}(s_{i+\ell-1})$$

$$= \hat{\delta}(s_i) + \Sigma_{k=0}^{\ell}\hat{p}(s_{i+k}). \qquad \square$$

Now we have

Lemma 4: For any $f \in F$,

$$\int u(f)dv = \Sigma_{i=1}^{n}\hat{p}(s_i)u(f(s_i)) + \hat{\delta}(s_i)[u(f(s_i)) - u(f(s_{i-1}))]^+$$

(where $x^+ = \max\{x,0\}$ for $x \in R$.)

Proof: For a given $f \in F$, let $\Pi: \{1,\ldots,n\} \to \{1,\ldots,n\}$ be a permutation such that $u(f(s_{\pi(i)})) \geq u(f(s_{\pi(i+1)}))$ for $1 \leq i \leq n-1$. By the definition of the Choquet Integral,

$$\int u(f)dv = \Sigma_{i=1}^{n}[u(f(s_{\pi(i)})) - u(f(s_{\pi(i+1)}))]v(\{s_j \mid \pi(j) \le \pi(i)\}) =$$

$$= \Sigma_{i=1}^{n}[u(f(s_{\pi(i)})) - u(f(s_{\pi(i+1)}))] \cdot$$

$$\left[\sum_{\{j|\pi(j)\le\pi(i)\}} \hat{p}(s_{\pi(j)}) + \right.$$

$$\left. \sum_{\{j|\pi(j)\le\pi(i)\}} \hat{\delta}(s_{\pi(j)}) 1_{\{u(f(s_{\pi(j)-1}))<u(f(s_{\pi(i)}))\}} \right]$$

$$= \Sigma_{i=1}^{n}\hat{p}(s_i)u(f(s_i)) + \sum_{\{k|u(f(s_k))>u(f(s_{k-1}))\}} \hat{\delta}(s_k)\Big[u(f(s_k)) -$$

$$- u(f(s_{k-1}))\Big]$$

$$= \Sigma_{i=1}^{n}\hat{p}(s_i)u(f(s_i)) + \hat{\delta}(s_i)\Big[u(f(s_i)) - u(f(s_{i-1}))\Big]^{+}. \qquad \square$$

Now we may also prove that

Lemma 5 For any $i \ge 2$, $\hat{p}(s_{i-1}) \ge \hat{\delta}(s_i) \ge -\hat{p}(s_i)$ and for $i \le n-1$ $\hat{\delta}(s_i) + \hat{p}(s_i) \ge \hat{\delta}(s_{i+1})$.

Proof: This result follows from the monotonicity axiom (A4) and Lemma 4.

$\square$

We shall now prove the Main Theorem. Let us define the functions p and δ by

$$\delta(s_i) = \frac{1}{2}\hat{\delta}(s_i) \qquad \text{for } i \le n$$

and

$$p(s_i) = \hat{p}(s_i) + \frac{1}{2}(\hat{\delta}(s_i) - \hat{\delta}(s_{i+1})) \quad \text{for } i \le n$$

(with $\hat{\delta}(s_{n+1}) = 0$ by definition.)

Let us first show that these functions satisfy our conditions. Consider condition (i). Obviously, $\delta(s_1) = \frac{1}{2}\hat{\delta}(s_1) = 0$. To see that $p(s_i) \ge |\delta(s_i)| + |\delta(s_{i+1})|$, one only has to use the definition of p and δ and Lemma 5. Next consider condition (ii). Since $x^+ = \frac{1}{2}(x + |x|)$ for all $x \in R$, Lemma 4 completes the proof.

Conversely, we have to assume that there are p and δ as required, and prove that A6 holds. However, this is quite easy. □□

7. Extension to an Infinite Horizon

Let us now suppose that $S = \{s_i \mid i \in N\}$ where $s_i \ll s_{i+1}$, and retain all other assumptions and definitions.

In the case of an infinite S, questions of continuity quite naturally arise. If Σ is a σ-algebra, we shall say that a measure v is **continuous** if, whenever $B_n \subset B_{n+1} \subset S$, $\lim_{n\to\infty} v(B_n) = v(\bigcup_{n\ge1} B_n)$ and whenever $S \supset B_n \supset B_{n+1}$ we have $\lim_{n\to\infty} v(B_n) = v(\bigcap_{n\ge1} B_n)$.

In our case, $\Sigma = 2^S$ is a σ-algebra, and it makes sense to ask when is v continuous. Our interest in this problem is not a matter of sheer curiosity (although we do not believe there is anything wrong with sheer curiosity): We cannot expect to have a "neat" representation of the Choquet Integral as an infinite series - unless v is continuous.

So let us define a topology on the set of acts F. We define it by the following notion of convergence: Let $\{f_n\}_{n\geq 1} \subset F$ and $f \in F$. We say that $\{f_n\}_{n\geq 1}$ **monotonically converges** to f if the following two conditions hold:

(i) There is a sequence $\{A_n\}_{n\geq 1} \subset \Sigma$ for which $A_n \subset A_{n+1}$ and $\bigcup_{n\geq 1} A_n = S$, such that $f_n(s) = f(s)$ for $s \in A_n$.

(ii) Either $f_n(s) \geq f_{n+1}(s) \geq f(s)$ for all $n \geq 1$ and $s \in S$

or $f_n(s) \leq f_{n+1}(s) \leq f(s)$ for all $n \geq 1$ and $s \in S$.

We now introduce another axiom:

A7 (Time Continuity) Suppose that $\{f_n\}$ monotonically converges to f and that $f > g (f < g)$. Then there exists an $n \geq 1$ such that $f_n > g (f_n < g)$.

The following lemma applies to Schmeidler's model whenever Σ is a σ-algebra (and do not depend upon the denumerability of S or the strict order (>>) defined on it):

Lemma 6 Suppose A1-A5 hold. Then A7 is satisfied iff v is continuous.

Proof: First assume A7 holds. Suppose that $B_n \subset B_{n+1}$ $(B_n \in \Sigma)$ and let $B = \bigcup_{n\geq 1} B_n$. Let $A_n = B_n \cup B^c$

and

$$f_n(s) = \begin{cases} y_1 & s \in B_n \\ y_0 & \text{otherwise} \end{cases} \qquad f(s) = \begin{cases} y_1 & s \in B \\ y_0 & \text{otherwise.} \end{cases}$$

By A7, $\int u(f_n)dv \underset{n\to\infty}{\to} \int u(f)dv$, hence $v(B_n) \to v(B)$.

Now, consider the case $B_n \supset B_{n+1}$ $(B_n \in \Sigma)$ and denote $B = \bigcap_{n\geq 1} B_n$. Let $A_n = B \cup B_n^c$ and define $\{f_n\}_n$ and f as above. Again $\{f_n\}_n$ monotonically converges to f and the result follows.

Now assume v is continuous, and that there are f, $\{A_n\}_{n\geq 1}$ and $\{f_n\}_{n\geq 1}$ as required by A7. (I.e., $\{f_n\}_{n\geq 1}$ monotonically converges to f and $f_n = f$ on A_n.) All we have to show is that $\int u(f)dv \underset{n\to\infty}{\to} \int u(f)dv$, that is, that the Choquet Integral is continuous w.r.t. monotonic convergence.

Since f and f_1 are bounded, the whole sequence $\{f_n\}_{n\geq 1} \cup \{f\}$ is uniformly bounded. Without loss of generality we may assume that it is bounded by y_1 and y_0, hence $\int u(f_n)dv = \int_0^1 v(\{s \mid u(f_n(s)) > t\})dt$ for all $n \geq 1$ and $\int u(f)dv = \int_0^1 v(\{s \mid u(f(s)) > t\})dt$.

Suppose $f_n(s) \leq f_{n+1}(s) \leq f(s)$ for all $n \geq 1$ and $s \in S$. (The other case, namely $f_n(s) \geq f_{n+1}(s) \geq f(s)$, is proved symmetrically.) For every $t \in [0,1]$,

$$\{s \mid u(f_n(s)) > t\} \subset \{s \mid u(f_{n+1}(s)) > t\} \subset \{s \mid u(f(s)) > t\}$$

and

$$\bigcup_{n\geq 1} \{s \mid u(f_n(s)) > t\} = \{s \mid u(f(s)) > t\}.$$

Hence $v(\{s \mid u(f_n(s)) > t\}) \underset{n\to\infty}{\to} v(\{s \mid u(f(s)) > t\})$ monotonically. Since $[0,1]$ is compact, $v(\{s \mid u(f_n(s)) > t\})$ uniformly converges (as a function of t) to $v(\{s \mid u(f(s)) > t\})$.
This implies that $\int u(f_n)dv \underset{n\to\infty}{\to} \int u(f)dv$. □

We may now prove

Theorem 2 Assume $\geq$ satisfies A1-A5. Then A6 and A7 hold iff there are $p, \delta: S \to R$ such that:

(i) $p(s_i) \geq |\delta(s_i)|+|\delta(s_{i+1})|$ for $i \geq 1$ and $\delta(s_1) = 0$.

(ii) For any $f \in F$,

$$\int u(f)dv = \Sigma_{i=1}^{\infty} p(s_i)u(f(s_i)) + \delta(s_i) \mid u(f(s_i)) - u(f(s_{i-1}))|.$$

Proof: For the "only if" part, let p and δ be defined as in the proof of the Main Theorem. Let there be given $f \in F$ and assume w.l.o.g. that $y_1 \geq f(s) \geq y_0$ for all $s \in S$. Define $A_n = \{s_i \mid i \leq n\}$ and

$$f_n(s) = \begin{cases} f(s) & s \in A_n \\ y_0 & \text{otherwise} \end{cases}$$

By A7, $\int u(f_n)dv \underset{n\to\infty}{\to} \int u(f)dv$. But $\int u(f_n)dv = \Sigma_{i=1}^{n} p(s_i)u(f(s_i)) + \delta(s_i) \mid u(f(s_i)) - u(f(s_{i-1})) \mid$ and (ii) follows. Hence (i) is also valid as in Lemma 5.

Now let us prove the "if" part. A6 is proved as in the Main Theorem. To see that A7 holds as well, one only has to notice that since the series in (ii) converges for all $f \in F$, v is continuous. □□

8. Definition and Characterization of Variation Aversion

In this section we shall assume that $n \geq 3$ (or $n = \infty$) and that on top of A1-A7, the following strong monotonicity axiom is satisfied:

A5* (Strong Monotonicity): For $f, g \in F$, if $f(s) \geq g(s)$ for all $s \in S$ and $f(s) > g(s)$ for some $s \in S$, then $f > g$.

Note that A5* implies that $p(s_i) > 0$ for all i.

We shall say that $\geq$ is **variation averse** if the following condition holds: For all $f_1, f_2, g_1, g_2 \in F$ and $i \geq 2$, if:

(i) $f_1(s_i) = g_1(s_i) < f_2(s_i) = g_2(s_i)$

(ii) $f_1(s_j) = f_2(s_j)$ and $g_1(s_j) = g_2(s_j)$ for $j \neq i$

(iii) $f_1 \sim g_1$

and (iv) $f_1(s_{i-1}), f_1(s_{i+1}) \leq f_1(s_i)$; $g_1(s_{i+1}) \leq g_1(s_i)$ and $g_1(s_{i-1}) \geq g_2(s_i)$

- then $f_2 < g_2$.

That is to say, if $f_1 \sim g_1$ and we improve both of them on s_i (to the level $f_2(s_i) = g_2(s_i)$), but f_1's variation has increased while that of g_1 has remained constant, then the modified f_1 (namely, f_2) is less preferred than the modified g_1 (which is g_2).

Theorem 3: Let $\geq$ satisfy A1-A7 and A5*. Then $\geq$ is variation averse iff $\delta(s_i) < 0$ for all $i \geq 2$.

Proof: First suppose that $\geq$ is variation averse. Let us use the functional form of Lemma 4 (rather than that of the Main Theorem (or Theorem 2) itself) and show that $\hat{\delta}(s_i) < 0$. For a fixed $i \geq 2$, choose $0 < \gamma < \alpha < \beta < 1$ such that $(\beta-\alpha)(\hat{\delta}(s_{i-1}) + \hat{p}(s_{s-i})) = \gamma\hat{p}(s_{i+1})$.
Now define

$$f_1(s_j) = \begin{cases} y_\alpha & j = i,\ i-1 \\ y_\gamma & j = i+1 \\ y_0 & \text{otherwise} \end{cases} \qquad g_1(s_j) = \begin{cases} y_\beta & j = i-1 \\ y_\alpha & j = i \\ y_0 & \text{otherwise} \end{cases}$$

and

$$f_2(s_j) = \begin{cases} y_\alpha & j = i-1 \\ y_\beta & j = i \\ y_\gamma & j = i+1 \\ y_0 & \text{otherwise} \end{cases} \qquad g_2(s_j) = \begin{cases} y_\beta & j=i-1,i \\ y_0 & \text{otherwise} \end{cases}$$

The numbers were chosen in such a way that $\int u(f_1)dv = \int u(g_1)dv$. Hence, by variation aversion, $f_2 < g_2$. However, $\int u(g_2)dv - \int u(f_2)dv = - (\beta-\alpha)\hat{\delta}(s_i) > 0$. Hence $\hat{\delta}(s_i) < 0$.

On the other hand, if $\delta(s_i) < 0$ for all $i \geq 2$, $\geq$ is obviously variation averse, and the proof is complete. □□

The definitions and characterizations of variation liking and variation neutrality are, of course, very similar and will not be given here in detail.

9. An Example

Consider a zero-sum two-person game played infinitely many times by two players which are identical as regards their assessments of future payoffs. That is to say, there exists a single functional $U(u_1,u_2,\ldots)$ such that $U(u^I(z_1),u^I(z_2),\ldots)$ represents player I's utility if the outcome of the i-th stage is z_i, and $U(u^{II}(z_1), u^{II}(z_2),\ldots)$ represents player II's utility.

In the classical model, $U(u_1, u_2,\ldots) = \Sigma_{i=1}^{\infty} p_i u_i$. Hence the super-game itself is also zero-sum and any pair of strategies is Pareto-optimal. However, if the two players are not variation-neutral, U is no longer a linear functional and this claim is no longer true. Since $|u^I(z_i) - u^I(z_{i-1})| = |u^{II}(z_i) - u^{II}(z_{i-1})|$ for any stage i and any outcome vector $(z_1, z_2,\ldots)$, it may be the case that replacing $(z_1, z_2,\ldots)$ by $(z_1', z_2',\ldots)$ will strictly increase or decrease both players' utility levels. Similarly, two identical agents in a single commodity economy (without production) may benefit from trade.

REFERENCES

Allais, M. (1953), "Le Comportement de l'Homme Rationnel devant le Risque: Critiques des Postulats et Axioms de l'Ecole Americaine" **Econometrica** 21, 503-546.

Anscombe F.J. & R.J.Aumann (1963), "A Definition of Subjective Probability", **The Annals of Math. Stat.**, 34, 199-205.

Choquet, G. (1955) "Theory of Capacities", Ann. Inst. **Fourier, Grenoble**, 5, 131-295.

Ellsberg, D. (1961) "Risk, Ambiguity and the Savage Axioms", **Quart. J. of Econ.**, 75, 643-669.

Koopmans, T.C. (1960), "Stationary Ordinal Utility and Impatience", **Econometrica** 28, 287-309.

Koopmans, T.C., P.A.Diamond and R.E.Williamson (1964), "Stationary Utility and Time Perspective", **Econometrica** 32, 82-100.

Kreps, D.M. and E.L.Porteus (1979). "Dynamic Choice Theory and Dynamic Programming", **Econometrica** 47, 91-100.

Lucas, R.E.Jr. and N.L.Stokey (1984), "Optimal Growth with Many Consumers", **JET** 32, 139-171.

Savage, L.J. (1954) **The Foundations of Statistics**, John Wiley & Sons, New York.

Schmeidler, D. (1982), "Subjective Probability and Expected Utility without Additivity" (A Temporary Title). Working Paper, The Foerder Institute for Economic Research.

Schmeidler, D. (1986) "Integral Representation without Additivity", **Proc. of the Amer. Math. Soc.**, Vol.97, No.2.