7

RANDOM ORDERINGS AND STOCHASTIC THEORIES OF RESPONSES*

(1960)

Co-author: H. D. Block

I. THE PROBLEM

In interpreting human behavior there is a need to substitute 'stochastic consistency of choices' for 'absolute consistency of choices'. The latter is usually assumed in economic theory, but is not well supported by experience. It is, in fact, not assumed in empirical econometrics and psychology.

The stochastic approach brings out the affinity between the phenomenon of choice and the more general psychological phenomenon of response to physical stimuli or, for that matter, to questionnaires, or of the action of 'judges' who compare the performances of individuals.[1]

Let A be the set of all alternatives (or actions) ever to be considered by a given subject. Let F be a subset of A. Let $(a; F)$ denote a 'multiple choice' (also called 'first choice' out of a set), i.e., the following observed fact: When forced to choose one element of F, the subject chose a. In the theory of choice, F is called the feasible set. When F consists of 2, 3,... elements, the multiple choice is called binary, ternary,....

In another language, the set F is associated with a stimulus and consists of all possible responses of the subject to that stimulus; A is then interpreted as the set of all possible responses to all stimuli. In experiments on so-called preferences and attitudes, the stimulus consists in offering a list of alternative menus, alternative political candidates, alternative answers to an item in the questionnaire. This list is then the feasible set F. When a 'judge' makes paired comparisons, the set F is a pair of individuals. In experiments on perceptional discrimination, physical stimuli – sounds, lights, or weights – impinge upon the subject; his response consists in stating that one of those objects is louder, or brighter, or heavier than the other(s). The set F is then the set of the n possible responses (or $n+1$ if 'I don't know' statements are permitted).

Another example of F is the set of all family budgets open to a family with a fixed amount of money, given the prices of goods (see Hotelling [2]). Each element of F is a different way of allocating the money among different goods.

In this paper – subject only to some attempted generalizations in Section VII – we shall call the observed choice, or response, $(a; F)$ the *basic* or *direct observation*. These observations can be counted, and the constraints on (i.e., the relations among) the resulting numbers (which may possibly be only 1's or 0's) we shall call *conditions*. Some of these conditions we shall call *directly testable*. Only those conditions are directly testable which involve the basic observations only. We shall encounter conditions that are not directly testable because they involve certain entities (constructs) that are not directly observable. An important class of these constructs will be called *utilities* (identical with *sensations* in psychophysics, *values* in anthropology, ethics, and older economics). These will be defined in a variety of ways, but will always denote certain real numbers, constant or random, associated with the elements of the set A.

However, to involve basic observations only is a necessary but not a sufficient property of a directly testable condition. For example, if the set A of alternatives is finite but of a size unknown, or too large to be exhausted by experiment, a constraint involving all alternatives simultaneously – such as Condition (o) in Section IV and Condition (P_n) in Section V – cannot be tested.

The purpose of this paper is to study the logical relations between various conditions describing consistency of responses, and in particular the logical relations between certain conditions that are not directly testable and others that are. This will tell whether a given class of observational results would or would not justify the acceptance (or the rejection) of a theory.

Our particular way of defining the class of basic observations and, correspondingly, of the directly testable conditions is to some extent arbitrary. Depending on the range of possible experiments and other observations, it may be preferable to define the class more narrowly, e.g., by including binary choices only. Or we might define this class more broadly. Following the practice of psychologists, we might admit the ranking, by the subject, of three or more objects as an observable fact, although the subject's observed action consists in this case of a verbal statement. (In the case of two objects, ranking and choosing is the same thing.) We might even admit as observable the subject's verbal statements of the relative 'intensity' of his preferences.

Such variations of the domain of testability will be tentatively discussed in Section VII. In the rest of the paper, by using a particular demarcation of the class of directly testable conditions (the one most closely corresponding to the nature of economic observations), we are able to carry out a reasonably complete analysis of the relevant logical relations. The study may thus serve as a start when similar attempts are made under another definition of basic observations.

In particular, our operational approach seems to be unable to handle the following distinction that appears natural on grounds of common sense and may be important for predictions. If out of the pair $F=(a, b)$ of desirable objects a man chooses sometimes a and sometimes b, our introspection tells us that we may ascribe this to either or both of two different 'causes':

1. He may have difficulty in perceiving all the relevant characteristics of the objects, as when a and b are two 10-ton carloads of the same merchandise but the exact quantity (in pounds, net of package) or, for that matter, all the differences in quality cannot be ascertained.
2. Even if he knew exactly the differences in the characteristics of the two objects, he might find them almost equally 'desirable' or containing the promise of equal 'satisfaction', and he will vacillate as a result.

To disentangle the two 'causes' – call them 'perceptibility' and 'desirability' (anticipated 'satisfaction') – may be important if one wants to predict how people will act if perceptibility is kept constant while desirability varies, or vice versa. Economists in particular have mostly considered men's decisions under the assumption of perfect perceptibility, also called perfect information. In the more recent developments of economic decision theory, imperfect information is introduced, but a clear, and probably fruitful, distinction is kept between varying the nature of (so-called) information and varying the satisfaction (the 'payoff') attainable through the choice of action. Economists[2] use the term 'utility' as interchangeable with desirability (satisfaction), and thus independent of perceptibility. Our concept is coarser. All the various definitions of utility given in this paper will be related to the empirical entities, called 'alternatives'. Each of these is identified precisely but combines the two aspects, information and desirability, in some unknown though presumably not-too-changeable fashion. Thus when two carloads are offered again, and, say, the firm's quality-control facilities (or, in a similar example, a housewife's abilities to discriminate between cuts of meat) have not changed, predictions from previous behavior observations can be made. If discrimination is learned, through experience, in some more-or-less-determined manner, the model has to be modified, of course, but we have not attempted to bring in learning.

Stochastic theory introduces the probability $p(a; F)$ of the basic observation $(a; F)$, and it is assumed that inferences about constraints on these probabilities can be made from a finite number of basic observations. These constraints may be described with the help of parameters of certain distributions, and some of these parameters have been called, a century ago, sensations (Fechner [3], see especially pp. 70–103 of the 1889 edition). We shall call them *constant utilities*. Various types of constant utilities have been proposed. They will be studied in Section II. In addition, there emerge naturally the concepts of *random utilities* and *random orderings* to be treated in Section III. Sections IV and V continue the study of directly testable conditions necessary and/or sufficient for the existence of constant and random utilities (and random orderings), respectively. Section VI sums up the main logical relations between the random and the various forms of constant utilities, and includes a very incomplete attempt to relate our various models to the work of other authors. As already mentioned, a generalized experimental situation – the combination of ranking and choosing – will be formulated in Section VII. Section VIII discusses the memory effect and other difficulties of experiments on choice and ranking, difficulties perhaps encountered less in psychophysics proper. We shall postpone the description of Section IX until the end of the present introductory section. Section X deals tentatively with a special statistical problem: how to test statistically our directly testable conditions – i.e., the properties of the probabilities $p(a; F)$ – when (to avoid the memory effect and to avoid assuming that all subjects have identical distribution properties) each set F can be offered only a few times, possibly only once.

Clearly, non-stochastic theory is a special, strong form of stochastic theory. And whenever a theory of choices or responses is being submitted to a statistical test there is, explicitly or not, an underlying model of stochastic behavior of the subject (or possibly of the error-making observer).

We shall conclude this section with a brief examination of the non-stochastic, or absolute, theory of choice. We shall formulate it in terms of our 'basic observations', and not, as usual, in terms of the subject's verbal preference or indifference statements. Define[3] for all a, b in A:

(1.1) '$a > b$' (a preferred to b) means 'never $(b; F)$ if F contains a',
'$a \gtrsim b$' means 'not: $b \succ a$',
'$a \sim b$' (indifference) means '$a \gtrsim b$ and $b \gtrsim a$'.

Thus, if the subject forced to choose between a and b chooses b nine (but not ten) times out of ten, the non-stochastic theory calls him indifferent. We have always $a \gtrsim b$ or $b \gtrsim a$ since the subject is forced to choose. Therefore the relation '$\gtrsim$' is said to induce complete weak ordering on A provided the following condition (testable by basic observations) is satisfied:

(1.2) if $a \gtrsim b$ and $b \gtrsim c$, then $a \gtrsim c$ (transitivity).

If, in addition, the set A is finite, $A = (a_1, \ldots, a_n)$, we can associate with each of its elements a_i an integer r_i, $1 \leqq r_i \leqq n$, called *rank*, such that

(1.3) $r_i \leqq r_j$ if $a_i \gtrsim a_j$.

The vector $r = (r_1, \ldots, r_n)$, which can also be regarded as an integer-valued function on the set of integers $N = (1, \ldots, n)$, is called a ranking on N. If by some arbitrary convention ties are excluded, all r_i's are different integers and r is a permutation. Clearly any strictly decreasing monotone function on the integers $r_1, \ldots, r_n$ induces a real-valued function ω on A which is order-preserving in the sense that

(1.4) $\omega(a_i) \geqq \omega(a_j)$ *if and only if* $a_i \gtrsim a_j$.

The function ω, called the ('*ordinal*') *utility function*, is unique up to increasing monotone transformations.

If the set A of alternatives is not finite, an order-preserving function ω on A need not exist. However, Debreu [5] proved that an order-preserving function exists if A and the ordering relation '$\gtrsim$' satisfy a certain rather weak condition[4] that may justify the assumption of ordinal utility functions over the space of commodity-bundles.[5]

In the stochastic models that will follow, the (testable) transitivity condition (1.2) and the 'ordinal' utility function ω made possible by it will be suitably generalized. But, in addition, some stronger testable conditions and, correspondingly, more strictly measurable utility functions will arise naturally.

A final remark: The case when the set A includes wagers so that choices are, in general, made under uncertainty, is more general than that of choices among sure alternatives. This case has been often treated, ever since Daniel Bernoulli [8] and, for that matter, Marshall (Note IX, p. 843, [9]), by ordering the wagers according to their 'expected utility'. This leads to a non-stochastic utility that is more strictly measurable than ω: It is unique up to increasing linear transformations. This model, too lends itself to stochastic generalizations, as will be briefly discussed in Section IX.

II. STOCHASTIC CONCEPTS OF CONSTANT UTILITIES

In general, $A=(a, b, \ldots)$ will continue to denote the set of distinct alternatives, and F the feasible subset. For mathematical ease, we shall assume A to be finite, unless otherwise stated, and identify it with $N=(1, \ldots, n)$. A feasible subset will be $M \subseteq N$. The probability that the subject forced to choose an element of M chooses i, denoted previously by $p(i; M)$, can be written more briefly thus: $i(M)$. Clearly

$$(2.1) \qquad i(M) \geqq 0\,; \qquad \sum_{i \in M} i(M) = 1\,.$$

When $M=(i, j, k, \ldots)$, $i(M)=i((i, j, k, \ldots))$ will be written simply $i(i, j, k, \ldots)$. Then $i(M)$ will be called the binary, ternary..., probability when M consists of 2, 3, ..., distinct elements. The binary probability will be sometimes written in still shorter forms: $i(i, j)=ij$, $i \neq j$. It will prove convenient to define $ii=\frac{1}{2}$ so that, using (2.1), always

$$(2.2) \qquad ij + ji = 1\,,$$

whether i and j are distinct or not.

For easier reference, the various *conditions* will be labeled by (more-or less suggestive) letters, thus: (x). *A theorem* is an implication-relation between conditions (the 'hypothesis' and the 'conclusion'); by using arrows several theorems can be combined into one. In addition to the usual signs $\rightarrow$ ('implies') and $\leftrightarrow$ ('implies and is implied by'), we shall also use $\mapsto$ ('implies but is *not* implied by'). When $(x) \mapsto (y)$, (x), is said to be stronger than (y); rejection of (y) forces rejection of (x), but acceptance of (y) is inconclusive. We shall also use the sign $\vdash\!\dashv$ for 'does not imply nor is implied by'.

Each of the following three conditions (w), (v), (u), arranged in a sequence of increasing strength, constrains the set of probabilities $i(M)$ by postulating the existence of some real vector of order n (a real-valued function on the set N), called utility vector (utility function), and denoted by w, v, u, respectively. The stronger the constraint, the more strict is the sense in which the utility vector in measurable; that is, the smaller is the group of transformations under which the vector remains indeterminate.

CONDITION (w). *There is a constant real vector* $w = (w_1, \ldots, w_n)$ *such that*

$$(2.3) \qquad w_i \geqq w_j \quad \text{if and only if} \quad ij \geqq \tfrac{1}{2},$$

where w_i may be called the *weak* utility of i; and w the weak utility function on N.

CONDITION (v). *There is a constant real vector* $v = (v_1, \ldots, v_n)$ *and, associated with it, a distribution function* φ_v, *strictly increasing* [*except when its value is* 0 *or* 1][6], *such that*

$$(2.4) \qquad \varphi_v(v_i - v_j) = ij\,; \qquad \varphi_v(0) = ii = \tfrac{1}{2}\,.$$

Then v_i may be called a *strong* utility of i; and v, a strong utility function on N.

CONDITION (u). *There is a constant positive vector* $u = (u_1, \ldots, u_n)$ *such that for any* i, j *in* M

$$(2.5) \qquad u_i/u_j = i(M)/j(M)\,, \quad \text{all} \quad M \subseteq N,$$

where u_i may be called the *strict* utility of i; and u the strict utility function on N.

Clearly the weak utility function w is *unique up to an increasing monotone* *transformation.* It is analogous to the function w of the non-stochastic model (Section 1), with '$a_i \gtrsim a_j$' interpreted as '$ij \geqq \frac{1}{2}$'. We can call the cases $ij > \frac{1}{2}$ and $ij = \frac{1}{2}$, stochastic preference and indifference, respectively. These concepts are implicitly used in experimental work as when, e.g., Mosteller and Nogee [10] define indifference as the case when the subject chooses one of the two offered alternatives half of the time. As in the non-stochastic model, (w) implies a testable

CONDITION (t) (transitivity).

$$\text{(2.6)} \qquad \textit{If} \quad ij \geqq \tfrac{1}{2} \quad \textit{and} \quad jk \geqq \tfrac{1}{2}, \quad \textit{then} \quad ik \geqq \tfrac{1}{2}.$$

Since always $ij \geqq$ or $\leqq \frac{1}{2}$, transitivity guarantees a (complete weak) ordering on N: the alternative i is ranked above, below, or on the level of j according as $ij >$, $<$, or $= \frac{1}{2}$. Since N is finite we, have (as in Section I):

THEOREM 2.1. $(w) \leftrightarrow (t)$ *provided the set of alternatives is finite.*

If we admit infinite sets of alternatives and redefine the condition (w) accordingly, we have $(w) \rightarrow (t)$. But if we denote by (D) Debreu's Condition (note 4) and replace '$a \gtrsim a'$' and '$ij \geqq \frac{1}{2}$' by '$aa' \geqq \frac{1}{2}$' we can write

THEOREM 2.2. $[(D), t] \mapsto (w) \mapsto (t)$.

Condition (v), the existence of 'strong' utilities, may be made plausible by a physical analogy. Because of random variations in the properties of the air, metals, and so forth, the lowest voltage needed to produce a spark in a given direction is random; assume then that the probability of the occurrence of the spark is the larger, the larger the voltage. Condition (v) has been used in psychophysics (and more recently in the scaling of attitudes) ever since Fechner [3]. It is stated in the psychologists' adage (Guilford, [11]): "*Equally often noticed differences* (on the 'sensation' scale) *are equal* [*unless noticed always or never*]." A similar thought might unify the recent attempts to measure the 'power' of person X over person Y: on the one hand, by the probability that Y obeys X; and on the other hand, by the difference in the payoff (utility) to Y in the case of his obeying, as compared with his disobeying, X.[7]

If (v) is satisfied by some v, φ_v, then clearly for *any increasing linear transform* $v' = \alpha + \beta v$ there exist a $\varphi_{v'}$ satisfying that condition. (If the set of alternatives is continuous, and the distribution φ_v strictly monotone, these are the only admissible transformations.[8]) The scale unit for v may be chosen by setting, for example, the quartile φ_v^{-1} $(.25)=1$; this or a similarly chosen unit is sometimes called the 'just noticeable difference', we believe. Moreover, because of (2.4), φ_v may be (and, in the continuous case, must be) so chosen that

$$(2.7) \qquad \varphi_v(\zeta) + \varphi_v(-\zeta) = 1$$

for all real ζ. Then φ_v is anti-symmetrical about the point $(0, \frac{1}{2})$; the median and the mean (if it exists) are zero. This is satisfied if φ_v is normal, as assumed by Fechner, whose test of normality was, however, rather crude. Normality of φ_v is also assumed in most textbooks; we do not know whether a formal test of this assumption has been developed. The 'logistic' distribution function $\varphi_v(\zeta)=1/(1+e^{-}\zeta)$ which, as we shall presently see, is required by the next stronger condition (u), also has the property (2.7).

THEOREM 2.3. $(v) \mapsto (w)$.

Proof. Sufficiency. Assume (v) and choose i, j with $v_i \geqq v_j$; then

$$ij = \varphi_v(v_i - v_j) \geqq \varphi_v(0) = \tfrac{1}{2}\,;$$

hence (w) is satisfied, with $w=v$.

No necessity. Let $N=(1, 2, 3)$ and assume $12>13>23>\frac{1}{2}$. Then (w) is satisfied, with $w_1>w_2>w_3$; but no linear function of w will satisfy (v); for, by (v), if $\frac{1}{2}<23$ then $0=v_1-v_1<v_2-v_3$, $v_1-v_2<v_1-v_3$; hence, by (v), $12<13$, which contradicts the assumption.

As to the 'strict' constant utility condition (u), it has been postulated, and developed most fully, by R. D. Luce [15].[9] We shall now give some of his results. Because of (2.1), Condition (u) can be rewritten in the form

$$(2.8) \qquad i(M) = u_i / \sum_{j \in M} u_j\,, \quad \text{all} \quad M \subseteq N\,.$$

It follows that the vector u is *unique up to a positive factor* $\lambda=u_i/i(N)$. With N fixed, a convenient normalization is $\lambda=1=\sum_{i\in N} i(N)$; $u_i=i(N)$.

Putting $M=(i, j)$, $i(M)=ij$, we see that the probability of a binary choice is related to the ratio u_i/u_j by

$$(2.9) \qquad \psi(u_i/u_j) = ij\,, \quad \text{where} \quad \psi(x) = \begin{cases} x/(1+x)\,, & x>0\,,\\ 0, & x \leqq 0\,. \end{cases}$$

Clearly ψ is a distribution function, for it is increasing, and $\psi(-\infty)=0$, $\psi(\infty)=1$. It is analogous to φ_v of condition (v), with ratios replacing differences, but is restricted to a particular form. Moreover, by putting v_k equal to $\log u_k$, we recover condition (v) itself, with φ_v a 'logistic curve':

(2.10) $\quad \varphi_v(\zeta) = 1/(1 + e^{-\zeta})$.

This proves

THEOREM 2.4. (u)$\mapsto$(v),
for (v) imposes constraints on binary choices only.

We shall now introduce two *directly testable* conditions equivalent to (u). First consider

CONDITION ($p.r$) (constancy of probability-ratios). $i(M)>0$ *for every* i *in* M; *and for every* i, j *in* N, *the ratio* $i(M)/j(M)$ *is constant over all sets* $M \subseteq N$ *that contain* i *and* j.

Condition ($p.r$) follows directly from (u); conversely, rewriting ($p.r$) as

(2.11) $\quad i(M)/i(N) = j(M)/j(N) = \sum_{k\in M} k(M)/\sum_{k\in M} k(N)$,

putting $k(N)=u_k$ and using (2.1), we obtain (u). Now consider

CONDITION ($c.p$) (probability of choice as a conditional probability).

(2.12) $\quad 0 < i(M) = i(N)/\sum_{k\in M} k(N) \quad$ all i in $\quad M \subseteq N$.

The fraction on the right-hand side, is, of course, the conditional probability that the element i of N is chosen *when it is known that the element chosen belongs to the subset M of N*. On the other hand, the quantity on the left-hand side is the probability that the element i of N is chosen *when the subject is forced to make his choice from the subset M of N*. Clearly these two quantities are not identical; ($c.p$) is a verifiable, empirical proposition which may or may not have intuitive appeal.[10] Condition ($c.p$) follows directly from (u) written in the form (2.8); on the other hand, on replacing i by j in (2.12) and comparing the two equations, one obtains ($p.r$) which we have seen to be equivalent to (u). We summarize this and the previous results of this section in the following theorem, placing the directly testable conditions in the lower line:

THEOREM 2.5.

$$\begin{array}{ccccc}(u) & \mapsto & (v) & \mapsto & (w)\\ \updownarrow & & & & \updownarrow\\ (p.r) & \mapsto & (c.p) & \mapsto & (t)\,,\end{array}$$

provided that the Debreu condition is satisfied. (This includes the case of a finite set of alternatives.)

III. RANDOM ORDERINGS AND RANDOM UTILITIES

In the non-stochastic theory briefly reviewed in Section I, the preference statement '$a_i \succ a_j$' assigns probability 1 to a certain class of choices. In the stochastic models of Section II, the corresponding statement '$ij > \frac{1}{2}$' merely sets a lower bound on that probability and is therefore weaker; however, some of those models are weaker than others.

A different way of weakening the non-stochastic theory consists in making the ('ordinal') utility function w a random one; i.e., by defining a probability measure on the space of all real-valued functions on A. In the case of the economic example of Section I, Wald's problem of evaluating the (non-stochastic) utility function w of consumers from their observed choices $(a; F)$ is replaced by the following more general one: Estimate the probability distribution P on the space of random utility functions U, using the probabilities $p(a; F)$ already estimated from the observations $(a; F)$, or using these observations directly. Any random utility function will be regarded as equivalent to any of its increasing monotone transforms.

However, the choice probabilities $p(a; F)$ may or may not be consistent with the existence of a random utility function. That is, the following condition (defining, in fact, that function) may or may not be fulfilled:

CONDITION (U) (existence of random utilities). *There is a random vector $U=(U_1, \dots, U_n)$ unique up to an increasing monotone transformation and such that, for any i in $M \subseteq N$,*

(3.1) $\Pr\{U_i \geqq U_j, \text{ all } j \text{ in } M\} = i(M)$.

U_i is called *random utility of i*. Putting $M=(i, j)$,

(3.2) $\Pr\{U_i \geqq U_j\} = ij$.

By (2.2), it follows that, for all $i \neq j$,

(3.3) $\Pr\{U_i = U_j\} = 0$;

for if $\Pr\{U_i = U_j\} > 0$, then by (3.2) $ij + ji = \Pr\{U_i \geqq U_j\} + \Pr\{U_i \leqq U_j\} = \Pr\{U_i \geqq U_j\} + \Pr\{U_i < U_j\} + \Pr\{U_i = U_j\} > 1$, contradicting (2.2).

When the set of alternatives is finite, we can consider one particular monotone transform of U, the ranking $r = (r_1, \ldots, r_n)$ on the set $N = (1, \ldots, n)$. Because of (3.3), no ties will occur. Hence R, the set of all rankings on N, consists of the $n!$ permutations on N. We shall denote by i_r the element of N that has rank i when the ranking is r; i.e., $r_{i_r} = i$. Using this notation, we write a given ranking $r = (r_1, \ldots, r_n) = 1_r 2_r \ldots n_r$ (without commas and parentheses). Thus, if $n = 3$, the ranking $r = 312$ means that $r_1 = 2$, $r_2 = 3$, $r_3 = 1$, and is therefore identical with $r = (2, 3, 1)$.

Denote by R_{iM} the set of all rankings on N in which i is the first among all elements of $M \subseteq N$:

(3.4) $R_{iM} = \{r \mid r_i \leqq r_j, \text{ all } j \in M\} \quad i \in M$.

For example, if $N = (1, 2, 3)$, $M = (1, 2)$, $i = 1$, then $R_{iM} = (123, 132, 312)$. Clearly, for every $M \subseteq N$,

(3.5) R_{iM} *and* R_{jM} *are disjoint for all* $i \neq j$ *in* M;

and if we denote by R^M the set of all permutations on M, we have

(3.6) $\bigcup_{i \in M} R_{iM} = R^M; \qquad \bigcup_{i \in N} R_{iN} = R^N = R$.

The following condition will be presently shown to be equivalent to (U).

CONDITION (P) (existence of a probability distribution of rankings consistent with probabilities $i(M)$). *There are $n!$ numbers $P(r)$ such that*

(3.7) $P(r) \geqq 0, \qquad \sum_{R_{iM}} P(r) = i(M), \qquad i \in M \subseteq N$.

It follows, using (3.6), (2.1), that $\sum_R P(r) = 1$. We can call $P(r)$ the *probability of the ranking* r.

THEOREM 3.1. $(U) \leftrightarrow (P)$.

Proof. If we assume Condition (U) and define $P(r) = \Pr\{U_{1_r} > U_{2_r} > \cdots > U_{n_r}\}$, we obtain Condition ($P$). Conversely, assume ($P$) and define the random vector U thus: For any non-random real vector $s=(s_i, \ldots, s_n)$ let

$$\text{(3.8)} \qquad \Pr\{U=-s\}=\begin{cases}P(s), & s \in R,\\ 0, & s \notin R,\end{cases}$$

where R is the set of permutations on N. Then Condition (U) is satisfied.

THEOREM 3.2. *The existence of random utilities U does not imply the existence of weak constant utilities w, let alone the existence of strong or strict utilities, v or u.*

Proof. Let $N=(1, 2, 3)$ and $0<\alpha<\frac{1}{3}$. Let the probabilities of rankings $P(r)$ be

$$P(123)=P(231)=P(312)=\tfrac{1}{6}+\alpha>0\,,$$
$$P(321)=P(213)=P(132)=\tfrac{1}{6}-\alpha>0\,.$$

(Then $\sum P(r)=1$.) Let the probabilities of choices $i(M)$, $M \subseteq N$, be

$$\text{(3.9)} \qquad \begin{aligned}&1(N)=2(N)=3(N)=\tfrac{1}{3},\\ &1(1,2)=2(2,3)=3(1,3)=\alpha+\tfrac{1}{2},\\ &2(1,2)=3(2,3)=1(1,3)=-\alpha+\tfrac{1}{2}.\end{aligned}$$

Then $1(N)=P(123)+P(132)$, $1(1,2)=P(123)+P(132)+P(312)$, etc., and Condition ($P$) – or its equivalent, ($U$) – is satisfied. But Condition (t) is not: $1(1,2)>\frac{1}{2}$, $2(2,3)>\frac{1}{2}$, but $1(1,3)<\frac{1}{2}$. This proves the present theorem, using Theorem 2.5.[11]

We shall show that the following conditions are necessary for (P), and therefore for its equivalent, (U):

CONDITION (e) (effect of enlarging the feasible set). *If $L \subseteq M \subseteq N$, then*

$$\text{(3.10)} \qquad i(M) \leqq i(L)\,.$$

CONDITION (e_3). *For any three elements i, j, k of N,*

$$\text{(3.11)} \qquad i(i,j) \geqq i(i,j,k); \textit{ or, equivalently,}$$

$$\text{(3.11')} \qquad i(i,j,k) \leqq \min(ij, ik)\,.$$

Clearly (e_3) is a special case of (e); it was exemplified numerically in (3.9), for $n=3$. Applying to (P) the fact that $R_{iM}=\{r \mid r_i \leqq r_j, j \in L\} \cap$

$\cap \{r \mid r_i \leqq r_j, j \in M-L\} \subseteq R_{iL}$, we obtain

THEOREM 3.3. $(P) \rightarrow (e) \mapsto (e_3)$.[12]

THEOREM 3.4. *The existence of random utilities U is not implied by the existence of strong constant utilities v, let alone by the existence of weak constant utilities w.*

Proof. Denote by x, y, z the distinct generic elements of $N=(1, 2, 3)$. There exist three numbers (for example, $v_1=0.2$, $v_2=0.1$, $v_3=0$) such that

$$\begin{aligned} 0 &< v_z - v_y + \tfrac{1}{2} = \beta_{xy} < 1\,, \\ 0 &< v_1 - v_2 = \gamma_1 < 1\,, \\ 0 &< v_2 - v_3 = \gamma_2 < 1\,, \\ 0 &< v_3 - v_1 + 1 = \gamma_3 < 1\,. \end{aligned}$$

Since $\beta_{xy}+\beta_{zx}=1=\sum_x \gamma_x$, we can put $x(x, y)=\beta_{xy}$, $x(x, y, z)=\gamma_x$. Then Condition (v) is satisfied, with $\varphi_v(\zeta)=\zeta+\frac{1}{2}$. But since $\gamma_3-\beta_{31}=\frac{1}{2}>0$, $3(1, 2, 3)>3(1, 3)$, Condition (e) and therefore also the stronger Condition (P) and its equivalent (U) are contradicted.

In the work of Thurstone [16] the following condition condition seems to be used:

CONDITION (*s.n*) (symmetrical normal). *There is a normal random vector $U=(U_1,\ldots, U_n)$ that satisfies Condition (U) and has all variances equal ($\sigma_{ii}=\sigma^2$, say) and all covariances equal ($\sigma_{ij}=\rho\sigma^2$, say).*

Clearly this implies that both (U) and (v) are satisfied: let $v_i=EU_i$ and let φ_v, at all points $\zeta=v_i-v_j$, be normal, with mean zero and variance $2\sigma^2(1-\rho)$. Even the following, weaker condition is sufficient for the conjunction [(U), (v)]:

CONDITION (s) (symmetry of adjusted random utilities). *There is a random vector U satisfying Condition (U), and a constant vector $v=(v_1,\ldots, v_n)$ such that $V=U-v$ has a distribution function symmetric in its arguments.*

If U, v satisfy this condition, then any pair of increasing linear transforms $\alpha+\beta U, \alpha+\beta v (\alpha, \beta$ scalars) will satisfy it. Clearly ($s.n$)$\mapsto$(s). Moreover,

THEOREM 3.5. $(s.n) \mapsto (s) \mapsto [(U), (v)]$.

Proof. Sufficiency of (s). If (s) is true, so is (U); and $i(i,j)=\Pr\{U_i \geqq U_j\}=\Pr\{V_j-V_i \leqq v_i-v_j\}=$ the value of a distribution function at v_i-v_j; because of symmetry, this function is the same for all i, j and can be written as φ_v, satisfying (v).

No sufficiency of [(U), (v)]. *Let* $N=(1, 2, 3)$. Let $v_i=i+3$, and let V_1, V_2, V_3 be independently distributed with $\Pr\{V_i=3\}=p_i=1-\Pr\{V_i=-3\}$, where $p_1=1-p_3=p>\frac{1}{2}$, $p_2=\frac{1}{2}$. If $U=v+V$, then $\Pr\{U_1>U_3\}=p^2$; $\Pr\{U_1>U_2\}=p/2=\Pr\{U_2>U_3\}$; $\Pr\{U_1>U_2, U_2>U_3\}=p^2/2$, etc. Hence U satisfies Condition (U), if $1(1, 2)=2(2,3)=p/2$; $1(1, 3)=p/2$; $1(1, 2, 3)=p^2/2$, etc. Then Condition (v) is also satisfied, with

$$\varphi_v(v_1-v_2)=\varphi_v(v_2-v_3)=\varphi_v(-1)=p/2>p^2= \\ =\varphi_v(-2)=\varphi_v(v_1-v_3).$$

Yet no linear transformation of v or monotone transformation of U can make the distribution of V symmetric, since the p_i are not all equal. If, on the other hand, the set of alternatives is continuous, let V have a joint normal distribution with zero mean and a constant variance τ^2 of differences V_x-V_y: $\tau^2=\sigma_{zz}+\sigma_{yy}-2\sigma_{xy}$, all x, y, where $\sigma_{11}=\frac{1}{2}$ and (for $x, y\neq 1$) $\sigma_{xz}=1$, $\sigma_{x1}=0$, $\sigma_{xy}=\frac{1}{4}$. If $U=V+EU$ satisfies condition (U), then (v) is also satisfied, with $v=EU$ and φ_v normal with zero mean and variance $\tau^2=\frac{3}{2}$. But no translation can make the distribution function of V symmetric in its arguments.

Mosteller [17] dropped Thurstone's symmetry assumption but maintained normality. We do not know whether tests have been developed for the joint condition [(U), (v)] in general, or for the special case of joint normality with constant variance of differences.

We shall now discuss another condition that is also stronger than the conjunction [(U), (v)] and is of particular interest, being equivalent to Condition (u) of Section II (existence of strict constant utilities). For a given permutation r of the set N define

$$\text{(3.12)} \qquad \pi(r)=\prod_{j=1}^{n-1} j_r(j_r, (j+1)_r, \ldots, n_r).$$

For example, $\pi(312)=3(1, 2, 3)\cdot 1(1, 2)$. Suppose this were the probability of the ranking 312. This would exemplify

CONDITION (π) (probability of ranking as the product of probabilities of successive first choices). *Condition* (P) *is true and, for every ranking* r *of* N, $P(r)=\pi(r)$ *as defined in* (3.12).

Substituting $\pi(r)$ for $P(r)$ in (3.7), we can put Condition (π) into a form involving the 'basic' probabilities (those of 'first choices') only:

$$(3.13) \quad i(M) = \sum_{r \in R_{iM}} \pi(r), \quad \text{for all } M \text{ and } i \in M.$$

THEOREM 3.6. $(u) \leftrightarrow (\pi)$.

Proof (outline). *Sufficiency of* (u). Define

$$(3.14) \quad \Omega(r) = \prod_{j=1}^{n-1} u_{j_r} / \sum_{k=j}^{n} u_{k_r}.$$

To prove that (u) implies (π), or (2.8) implies (3.13), is to prove the identity

$$(3.15) \quad \sum_{r \in R_{iM}} \Omega(r) = u_i / \sum_{h \in M} u_h,$$

for all $i \in M$, all $M \subseteq N$, and any positive numbers $u_1, \ldots, u_n$.[13] Without loss of generality let $M = (1, \ldots, m)$, $m \leqq n$, and, as in (3.6), denote by R^M the set of all permutations of M. Then $r^M = (r_1^M, \ldots, r_m^M) \in R^M$. The set $R = R^N$ is partitioned into m subsets of the form

$$R_{iM}^N = (r^N \mid r_i^N \leqq r_j^N; \text{ all } j \in M), \quad i \in M;$$

the set R^N is also partitioned into $m!$ subsets of the form

$$R^N(r^M) = (r^N \mid r_i^N \leqq r_j^N \text{ if } r_i^M \leqq r_j^M; \text{ all } i.j \in M).$$

Similarly, R^M is partitioned into subsets $R^M(r^L)$, $L \subseteq M$. One obtains, by induction on m, first the identities

$$\sum_{r^M \in R^M} \Omega(r^M) = 1, \qquad \sum_{r^M \in R^M(r^L)} \Omega(r^M) = \Omega(r^L),$$

and then (3.15).

Necessity of (u). To show that (3.13) implies (2.8), let $u_i = i(N)$, all $i \in N$. Because of the identity (3.15), all numbers $i(M)$ that satisfy (3.13) also satisfy the system of equations

$$(3.16) \quad i(M) = \sum_{r \in R_{iM}} u_{1_r} \prod_{j=2}^{n-1} j_r(j_r, \ldots, n_r), \quad \sum_{i \in M} i(M) = 1.$$

One then proves the uniqueness of this solution: Starting with $M = N$, then taking all sets M with $n-1$ elements, then with $n-2$ elements and so on, one can solve (3.16) for all $i(M)$ and one finds that they are all given by (2.8).

THEOREM 3.7. *If* (u) *is true, then* (P) *and* (U) *hold, with* $P(r)=\pi(r)=$ $=\Omega(r)$. This is clearly a corollary to Theorem 3.6, using Theorem 3.1.

Some interest attaches to a condition that, in effect, appears in the work of Luce [15]. In the same way as Condition (u) treats the probabilities of *first* choices, $p(i;M)=i(M)$, this new condition treats the probabilities of *last* choices. We may denote these by $p'(i;M)=i'(M)$ and remark that last choices are not 'basic observations' in our sense. (However, see Section VII on considering the rankings themselves, and hence also the last choices, 'observable'; see also Marschak [20], Section XI.)[14]

CONDITION (u'). *There exists a constant positive vector* $u'=(u_i',\ldots,u_n')$ *such that for any* i, j *in* M, $u_i'/u_j'=i'(M)/j'(M)$.

It is possible, however, to prove

THEOREM 3.8. *If* $n>2$, (u') *is inconsistent with* (u) *unless* $i(N)=i'(N)=$ $=1/n$ *for all* $i\in N$.

Proof (outline). Define, for each ranking $r=1_r2_r\ldots n_r$ the inverted ranking $r^*=n_r(n-1)_r\ldots 2_r1_r$ (hence $r^{**}=r$); and define

$$\Omega'(r)=\prod_{j=1}^{n-1} u'_{j_r}/\sum_{k=j}^{n} u'_{k_r}\,.$$

By Theorem 3.7. if (u) is true then $P(r)=\Omega(r)$. By the same theorem, if (u') is true then $P(r^*)=\Omega'(r)$, and hence $P(r)=P(r^{**})=\Omega'(r^*)$. Therefore if both ($u$) and ($u'$) are true,

$$(3.18)\qquad \Omega(r)=\Omega'(r^*).$$

Define the ranking $s=2_r1_r3_r\ldots(n-1)_rn_r$. Since $\sum u_i=1=\sum u_i'$ without loss of generality, we have $\Omega(r)/\Omega(s)=(1-u_{2_r})/(1-u_{1_r})$; $\Omega'(r^*)/\Omega'(s^*)=$ $=u'_{2_r}/u'_{1_r}$. Since 1_r2_r are arbitrary, we have by (3.18) $u_i'/u_j'=(1-u_i)/(1-u_j)$; and summing on i, we find $u_j=1-(n-1)\,u_j'$. Now define $t=1_r2_r\ldots(n-2)_r$ $n_r(n-1)_r$; then by (3.18)

$$\Omega(r)/\Omega(t)=u_{(n-1)_r}/u_{n_r}=\Omega'(r^*)/\Omega'(t^*)$$
$$=(1-u'_{(n-1)_r})/(1-u'_{n_r})\,;$$

and in general $u_i/u_j=(1-u_i')/(1-u_j')=[1-(n-1)\,u_i']/[1-(n-1)\,u_j']$; or $(n-2)\,u_i'=(n-2)\,u_i'$. Hence, for $n>2$, we have $u_i'=u_j'$ and $u_i=u_j$; so that $i(N)=i'(N)=1/n$ for all i.

THEOREM 3.9. $(u)\mapsto[(U,v)]\leftarrow\!\dashv(s)$.

Proof. We have $(u)\mapsto(U)$ by Theorems 3.6 and 3.1, since (π) is strictly stronger than (P). Moreover, $(u)\mapsto(v)$, and $(s)\mapsto(U,v)$ by Theorems 2.4 and 3.5. It remains to prove that (U,v) does not imply (u). Consider any normal distribution symmetric in its arguments. Then, by Theorem 3.5, (U) and (v) are both satisfied; and the corresponding φ_v is normal, contrary to the implication (2.10) of (u).

Some of the results of this and the preceding sections are worth summarizing in

THEOREM 3.10.
$$\begin{tikzcd}
(u) \arrow[r, mapsto] & {[(U),(v)]} \arrow[r, mapsto] \arrow[d, mapsto] & (U) \arrow[dl, |-|] \arrow[d, |-|] \\
& (v) \arrow[r, mapsto] & (w).
\end{tikzcd}$$

IV. FURTHER DIRECTLY TESTABLE CONDITIONS FOR THE EXISTENCE OF CONSTANT UTILITIES

We shall deal here with 'strong' utilities v only; directly testable conditions for the existence of 'strict' utilities u and 'weak' utilities w were given in Section II. Directly testable conditions for (v) – like those for (w) and unlike those for (u) – can, of course, involve binary choices only.

Unless otherwise stated, the set of alternatives will be assumed finite: $N=(1,\ldots,n)$. It is easy to see that we can rewrite

CONDITION (v). *There is a real vector $v=(v_1,\ldots,v_n)$ such that*

(4.1) *if* $hi\geqq jk$ *then* $v_k-v_i\geqq v_j-v_k,$

provided hi and jk are not both 0 *or* 1.

This implies that $hi\geqq\frac{1}{2}$ if, and only if, $v_k-v_i\geqq 0$. It is seen immediately that (v) implies each of the following conditions:

CONDITION (t) (transitivity). *If $ij\geqq\frac{1}{2}$ and $jk\geqq\frac{1}{2}$, then $ik\geqq\frac{1}{2}$.*

CONDITION (t_s) (strong transitivity). *If* $ij \geqq \frac{1}{2}$ *and* $jk \geqq \frac{1}{2}$, *then* $ik \geqq \max(ij, jk)$.

CONDITION (t_s^*). *If* $ij \geqq \frac{1}{2}$, *then* $ix \geqq jx$, *all x in N*.

CONDITION (t_s^{**}). *If there exists k such that* $ik \geqq jk$, *then* $ij \geqq \frac{1}{2}$.

CONDITION (6_w) (weak condition on 6-tuples). *If* $i_1i_2 \geqq j_1j_2$ *and* $i_2i_3 \geqq \geqq j_2j_3$, *then* $i_1i_3 \geqq j_1j_3$.

CONDITION (m_w) (weak condition on m-tuples). *For any* $m \leqq n$, *if* $i_ki_{k+1} \geqq j_kj_{k+1}$, $h = 1, \ldots, m-1$, then $i_1i_m \geqq j_1j_m$.

CONDITION (q) (quadruple condition). *If* $hi \geqq jk$, *then* $hj \geqq ik$.

CONDITION (6_s) (strong condition on 6-tuples). *If* $i_1i_2 \geqq i_2j_3$ *and* $i_2i_3 \geqq \geqq j_1j_2$, *then* $i_1i_3 \geqq j_1j_3$.

THEOREM 4.1.
$$\begin{array}{ccccccccccc} (v) & \mapsto & (6_s) & \mapsto & (q) & \mapsto & (6_w) & \mapsto & (t_s) & \mapsto & (t) \\ & & & & & & \updownarrow & & \updownarrow & & \\ & & & & & & (m_w) & & (t_s^*) \leftrightarrow (t_s^{**}). & & \end{array}$$

Proof. Obviously (v) implies all the other conditions, and obviously $(t_s) \rightarrow (t)$. To prove $(t_s) \rightarrow (t_s^*)$, let $ij \geqq \frac{1}{2}$ and apply (t_s) to three cases: (1) $jx \geqq \frac{1}{2}$; (2) $jx < \frac{1}{2}$, $xi \geqq \frac{1}{2}0$ (3) $jx < \frac{1}{2}$, $xi < \frac{1}{2}$. To prove the converse, let $ij \geqq \frac{1}{2}, jk \geqq \frac{1}{2}$; by ($t_s^*$) $ix \geqq jx, jy \geqq ky$; put $x = k$, $y = i$. To prove $(t_s^*) \leftrightarrow (t_s^{**})$, apply a contradiction argument, using (2.2). To prove $(6_w) \rightarrow (t_s)$, let $ij \geqq \frac{1}{2}, jk \geqq \frac{1}{2}$ and apply (6_w) with $i = i_1, j = i_2 = j_2, k = i_3$; and (1) $j_1 = j, j_3 = k$; (2) $j_1 = i, j_3 = j$. Condition (m_w) implies, and is itself obtained by successive application of (6_w). To prove $(q) \rightarrow (6_w)$ is straightforward. To prove $(6)_s \rightarrow (q)$, let $hi \geqq jk$, $i_1 = h$, $j_1 = i_2 = i$, $i_3 = j_2 = j$, $j_3 = k$.

The absence of implications is proved by counterexamples:[15] (t) does not imply (t_s) since, with $n = 3$ and $12 > 13 > \frac{1}{2}$ and $23 > \frac{1}{2}$, (t) but not (t_s) is satisfied. To prove that (t_s) does not imply (6_w), let $n = 4$ and $\frac{1}{2} < 43 < 32 < < 21 < 31 < 42 < 41$. To prove that ($6_w$) does not imply ($q$), let $n = 4$, and $\frac{1}{2} < 32 < 43 < 21 < 31 < 42 < 41$. To prove that ($q$) does not imply ($6_s$), let $n = 5$ and $\frac{1}{2} < 21 < 54 < 32 < 43 < 53 < 31 < 42 < 41 < 52 < 51$. To prove that ($6_s$) does not imply ($v$), a counterexample (with $n = 9$) was constructed, but must be omitted here.

All the conditions given so far are necessary but not sufficient for the existence of v. With the exception of (m_w), they all involve binary choices among elements of subsets of fixed size m; $m=3$ in (t) and (t_s); $m=4$ in (q); $m=6$ in (6_w) and (6_s). The exception (m_w) is only apparent as (m_w) is equivalent to (6_w). Therefore all those conditions can be applied even when the size of N (the number n of all alternatives) is unknown. The following two conditions can be applied only if the size of N is known.

CONDITION (m_s) (strong condition on m-tuples). *For any* $m \leqq n$, $i_1 i_m \geqq j_1 j_m$, *if* $i_k i_{1+k} \geqq j_{r_k} j_{1+r_k}$, $k=1,\ldots,m-1$, *where* $r=(r_1,\ldots,r_{m-1})$ *is any permutation of the set* $(1,\ldots,m-1)$.

CONDITION (o) (ordering of the probabilities of binary choices). *Arrange all the* $ij \geqq \frac{1}{2}$ *in a sequence of* $n(n-1)/2$ *inequalities (involving only n distinct alternatives):* $i_1 i_2 \geqq i_3 i_4 \geqq \cdots \geqq \frac{1}{2}$; *then the corresponding system of* $n(n-1)/2$ *inequalities in n distinct real numbers*

$$v_{i_1} - v_{i_2} \geqq v_{i_3} - v_{i_4} \geqq \cdots \geqq 0$$

must have a solution.

The following needs no proof.

THEOREM 4.2. $(v) \leftrightarrow (o) \rightarrow (m_s) \mapsto (6_s)$.

We have not investigated whether (m_s) implies (o).

Thus, for a finite set N of alternatives we have the condition (o) [perhaps also (m_s)?] sufficient and necessary for the existence of a strong utility vector v, and testable if N has known size. For finite N of unknown size our testable conditions are necessary only.

A summary of the main results of this section, assuming the set of alternatives finite, is given by

THEOREM 4.3.
$$\begin{array}{ccccccccccc} (v) & & & & & & (w) & & & & \\ \updownarrow & & & & & & \updownarrow & & & & \\ (o) & \rightarrow & (m_s) & \mapsto & (q) & \mapsto & (6_w) & \mapsto & (t_s) & \mapsto & (t). \end{array}$$

If the set of alternatives is ascribed some continuity properties (to be defined presently), some of the necessary testable conditions become also sufficient. We give two examples. Suppose one postulates

CONDITION $(s.c)$ (stochastic continuity). *For any elements a, b, c of the set A of alternatives such that $ab<\lambda<ac$, there is an element d in A with $ad=\lambda$.* Debreu [24] has proved

THEOREM 4.4. $[(s.c), (q)] \mapsto (v)$.

Or one may postulate a stronger

CONDITION $(d.b.p)$ (continuous differentiability of binary probabilities). *The set, A, of alternatives is representable by a real interval, and $p(a, b)=p(a; (a, b))$ is continuously differentiable with $\partial p/\partial b<0$, $\partial p/\partial a>0$.*

In Block and Marschak [21] proof was given of

THEOREM 4.5. $[(d.b.p), (6_w) \mapsto (v)$.

To show this, one first proves a

LEMMA. *Let $x, y \in [0, 1]$ and let $f(x, y)$ be continuously differentiable with $\partial f/\partial y<0$, $\partial f/\partial x>0$. Then, in order that f admit the representation $f(x, y)= =h(u(x)-v(y))$,* where *h, u, v are monotone increasing functions of one variable, it is necessary and sufficient that there exist functions $\alpha(x)$, $\beta(y)$ such that $(\partial f/\partial x)/(\partial f/\partial y)=\alpha(x)/\beta(y)$.*

V. FURTHER DIRECTLY TESTABLE CONDITIONS FOR THE EXISTENCE OF RANDOM ORDERINGS

Directly testable sufficient conditions for (U), the existence of random utilities (or, equivalently, for (P), the existence of random orderings), were given in Section III. It was shown, for example, that the probabilities $P(r)$ exist if the ratios $i(M)/j(M)$ do not depend on M. We have also shown that this condition is not necessary. On the other hand, the directly testable condition (e) is necessary: if the $P(r)$ exist, then $i(M) \leq i(L)$ for any $i \in L \subseteq M \subseteq N$. We shall presently show that this condition is not sufficient. We have not found directly testable conditions that are both sufficient and necessary for the existence of $P(r)$ on N. The following partial results may, however, prove useful. They still use finite sets of alternatives only, of known or unknown size. The case of continuous sets has not been explored.

Condition (e_3) defined in Section III (3.11), as a special case of Condition (e), can also be conveniently denoted by (P_3), being the first link in a sequence of Conditions (P_m), $m=3,\ldots,n$, which we shall state presently. Let us first state

THEOREM 5.1. *If* $n=3$, *then* $(P_3)\leftrightarrow(P)$.

Proof. Sufficiency of (P_3). Denote by x, y, z the generic distinct elements of $N=(1, 2, 3)$, define $Q(zxy)=x(x, y)-x(x, y, z)$, and assume (P_3) or its equivalent, (e_3). Then by (3.11), $Q(zxy)\geqq 0$; and by (2.2), (2.1), the sum of the six numbers $Q(zxy)$ is equal to $3-2=1$; hence (P) is satisfied, with $P(xyz)=Q(xyz)$.

Necessity of (P_3). By Theorem 3.3.

CONDITION (P_4). *For any four distinct elements* w, x, y, z *of* N,

$$w(w, x, y) \geqq w(w, x, y, z); \tag{5.1}$$

$$w(w, x) - w(w, x, y) \geqq w(w, x, z) - w(w, x, y, z). \tag{5.2}$$

THEOREM 5.2. $(P) \rightarrow (P_4) \mapsto (P_3)$.

Proof. Add (5.1) and (5.2) to obtain (P_3) $(=e_3)$ from (P_4). The converse is clearly not true. Moreover, if (P) is true, then

$$w(w, x, y) - w(w, x, y, z) = P(zwxy) + P(zwyx), \tag{5.3}$$

$$\begin{aligned} w(w, x) - [w(w, x, y) + w(w, x, z)] - w(w, x, y, z) = \\ = P(zywx) + P(yzwx); \end{aligned} \tag{5.4}$$

since the expressions on the right-hand side are non-negative, (P_4) is satisfied. The converse is not true, since (P) may involve subsets with more than four elements.

THEOREM 5.3. *If* $n=4$, $(P_4)\leftrightarrow(P)$.

Proof. Sufficiency of (P_4). If x, y, z, w are generic distinct elements of $N=(1, 2, 3, 4)$, then (5.3), (5.4) is a system of 24 equations in as many unknowns, $P(r)$, non-negative and adding up to 1; the knowns being the $x(M)$, $M\subseteq N$. Each permutation r occurs exactly once in the set (5.3) and exactly once in the set (5.4). Thus the system breaks into six sets of four equations; each set involves exactly four permutations r (none of which appears in the other five sets) and is of the form

$$\begin{aligned}(5.5)\qquad \alpha &= P(r^{(1)}) + P(r^{(2)}),\\ \beta &= P(r^{(2)}) + P(r^{(3)}),\\ \gamma &= P(r^{(3)}) + P(r^{(4)}),\\ \delta &= P(r^{(4)}) + P(r^{(1)}),\end{aligned}$$

where the known α, β, γ, δ are non-negative. They obey the restriction $\alpha+\gamma=\beta+\delta$, so that the four equations are linearly dependent. A non-negative solution for the four $P(r)$ is obtained by putting $0=P(r^{(4)})$ or $0=P(r^{(2)})$ according as $\beta>\gamma$ or not. One then verifies (P) for all $x(M)$ and obtains $\sum_R P(r)=1$.

Necessity of (P_4). By Theorem 5.2.

THEOREM 5.4. *When* $n=4$, $(P_4)\mapsto(e)$.

Proof. Sufficiency. By Theorems 5.3 and 3.3.

No necessity. (P_4) implies

$$(5.6)\quad w(w, x) - w(w, x, y) \geqq w(w, x, z) - w(w, x, y, z) \geqq 0\,;$$

when $n=4$, (e) is the pair of inequalities

$$(5.7)\qquad w(w, x) - w(w, x, y) \geqq 0\,,$$

$$(5.8)\qquad w(w, x, z) - w(w, y, z) \geqq 0\,,$$

which are implied by, but do not imply, (5.6).

Conditions (P_4) and (P_3) are special cases of

CONDITION (P_m). *Let* $y, z, x_1, \ldots, x_{m-2}$ *denote m generic distinct elements of a set M consisting of m elements,* $3\leqq m\leqq n$. *Then for any h,* $3\leqq h\leqq m$,

$$\begin{aligned}(5.9)\qquad S_h = y(y, z) &- [y(y, z, x_1) + y(y, z, x_2) + \cdots +\\ &+ y(y, z, x_{h-2})] + [y(y, z, x_1, x_2) + y(y, z, x_1, x_3) +\\ &+ \cdots + y(y, z, x_{h-3}, x_{h-2})] \ldots \pm y(y, z, x_1, \ldots, x_{h-2}) \geqq 0.\end{aligned}$$

THEOREM 5.5. $(P)\rightarrow(P_n)$, *with*

$$(5.10)\qquad S_m = \sum_r P(1_r 2_r \ldots (m-2)_r\, yz), \quad m = 3, \ldots, n\,;$$

the sum being taken over all permutations r of the m-element-set M which end in yz.

Proof. Assume (P) and count the number of times, with sign, that any particular permutation is included in S_m. A term of the type $\ldots y \ldots z \ldots$, where $k \geqq 1$ elements besides z follow y, has the coefficient

$$(5.11) \quad 1-\binom{k}{1}+\binom{k}{2}-\ldots+\binom{k}{k}=(1-1)^k=0.$$

and so the only remaining terms will be as specified on the right-hand side of (5.10).

We have seen that, in addition, $(P_n) \rightarrow (P)$ for $n \leqq 4$. We do not know whether this is true for $n > 4$. And we note that (P_n) is not testable if n is unknown.

An interesting condition involving binary choices only was formulated independently by several authors (perhaps first by Guilbaud [25]):

CONDITION (c_3) (a condition on cycles of 3 elements).[16] *For any three distinct elements i, j, k of N,*

$$(5.12) \quad 1 \leqq ij + jk + ki \leqq 2.$$

THEOREM 5.6. $(e_3) \mapsto (c_3)$.

Proof. Sufficiency. Let $M=(i, j, k) \subseteq N$. By (e_3), $ij \geqq i(M)$, $jk \geqq j(M)$, $ki \geqq k(M)$; add; apply (2.1), (2.2).

No necessity. Use the example of Theorem 3.4, where $ij+jk+ki=\frac{3}{2}$, so that the binary probabilities obey (c_3); but one of the ternary probabilities contradicts (e_3). We shall now prove

THEOREM 5.7. *Condition (c_3) is sufficient and necessary for the binary probabilities to be consistent with Condition (e_3).*

Proof. Sufficiency. Denote by x, y, z the generic distinct elements of $M=(i, j, k) \subseteq N$. We have to show that if the six binary probabilities xy satisfy (c_3), then there exist three numbers γ_z that have the properties required by (e_3) of the ternary probabilities, viz., $\sum_z \gamma_z = 1$, $0 \leqq \gamma_z \leqq$ $\leqq \min(xy, xz)$ for all x. Clearly such numbers exist if

$$(5.13) \quad s = \min(ij, jk) + \min(ji, jk) + \min(ki,)\, kj \geqq 1.$$

Let ij be minimal, and hence, by (2.2), ji maximal, among the six xy. If $kj < ki$, then, by (2.2), $s = ij + 1 \geqq 1$; if $kj \geqq ki$, then (c_3) implies $s \geqq 1$.

Hence, (c_3) guarantees the existence of the desired numbers γ_z (in general not unique).

Necessity. By Theorem 5.6.

Theorem 5.7 may be of importance if binary probabilities are the only observable ones (see Section VII). For the same reason, the following theorems, 5.8 and 5.9, will be of interest. We shall first state the relations between (e_3), (c_3), and the two transitivity conditions (t_s) and (t) of Section IV:

THEOREM 5.8.
$$\begin{tikzcd} (t_s) \arrow[r, |-|] \arrow[d, mapsto] \arrow[dr, mapsto] & (e_3) \arrow[d, mapsto] \arrow[dl, mapsto] \\ (t) \arrow[r, |-|] & (c_3). \end{tikzcd}$$

Proof. $(t_s) \mapsto (t)$, $(e_3) \mapsto (c_3)$ by Theorems 4.1, 5.6. To show that (t_s) implies (c_3) assume, for some i, j, k in N, $ij+jk+ki<1$, contradicting (c_3); then $ij+jk<1$, and by (2.2) $ij<kj$; so that, if (t_s) [and therefore, by 4.1, (t_s^{**})] holds, then $ik<\frac{1}{2}, ki>\frac{1}{2}$, and by symmetry $ij>\frac{1}{2}, jk>\frac{1}{2}$, hence $ij+jk+ki>\frac{3}{2}$, contradicting the assumption. To show that (c_3) does not imply (t), let alone (t_s), consider $N=(i, j, k)$, and make $ij=jk=ki=0.4$, thus satisfying (c_3) but not (t). On the other hand, the case $N=(i, j, k)$, $ij=jk=0.1$, $kj=0.6$ shows that (t) does not imply (c_3). Finally, to see that $(t_s) \vdash\!\dashv (e_3)$ and $(t) \vdash\!\dashv (e_3)$, note that (t_s) and (t) constrain only binary probabilities while (e_3) constrains ternary as well as binary ones, and its constraint on the latter is identical with (c_3), which was just shown not to imply (t_s) or (t).

The following condition on binary probabilities, due to Georgescu-Roegen [22], generalizes (c_3):

CONDITION (c_m). *For any $x_1, \ldots, x_m$ in N,*

$$1 \leqq x_1x_2 + x_2x_3 + \cdots + x_{m-1}x_m + x_mx_1 \leqq m-1. \tag{5.14}$$

A proposition stronger than (c_m) and – as will be proved presently – equivalent to (c_3) is

CONDITION (c^*). (c_m) *holds for all $m \leqq n$.*

THEOREM 5.9. $(c_3) \leftrightarrow (c^*) \mapsto (c_m)$.

Proof. We have only to prove that (c_3) implies (c^*) and is not implied by (c_m), $m>3$. First show that the conjunction $[(c_3), (c_{m-1})] \to (c_m)$: by (c_{m-1}),

$$(5.15) \qquad 1 \leqq x_1x_2 + x_2x_3 + \cdots + x_{m-1}x_1 \leqq m-2\,;$$

and, by (c_3)

$$1 \leqq x_{m-1}x_m + x_mx_1 + x_1x_{m-1} \leqq 2\,,$$

or, by (2.2)

$$(5.16) \qquad 0 \leqq x_{m-1}x_m + x_mx_1 - x_{m-1}x_1 \leqq 1\,;$$

adding (5.15) and (5.16), we obtain (c_m); now put $m=4, 5, \ldots$. To prove that (c_m) does not imply (c_3) let $m=n=4$; if the matrix $[x_ix_j]$ is

$$\begin{bmatrix} 0.5 & 0.8 & 0.2 & 0.5 \\ 0.2 & 0.5 & 0.8 & 0.6 \\ 0.8 & 0.2 & 0.5 & 0.6 \\ 0.5 & 0.4 & 0.4 & 0.5 \end{bmatrix},$$

then (c_4) is satisfied, but (c_3) is not, for $x_1x_2+x_2x_3+x_3x_1=2.4>2$.

The main results of this section can be summarized in

THEOREM 5.10. *If $n \geqq m > 3$, then*

$$\begin{array}{ccccccccc} (P) & \to & (P_n) & \mapsto & (P_m) & \mapsto & (e_3) & \leftrightarrow & (P_3) \\ \downarrow & & & & & & \downarrow & & \\ (e) & \mapsto & (c_m) & \leftarrow\!\dashv & (c^*) & \leftrightarrow & (c_3) & \leftarrow\!\dashv & (t_s) \mapsto (t)\,; \end{array}$$

if $n=3$ or 4, then $(P) \mapsto (P_n)$.

VI. A SURVEY

Theorem 3.10 summarized the relations between conditions postulating the existence of random orderings, and of constant utilities of various degrees of strength. It is also useful to summarize the relations between Conditions (U), (u), (v), and (w) and some of the testable conditions, mainly from Sections IV and V:

THEOREM 6.1. *For A finite,*

$$\begin{tikzcd} (c.p) \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] & (u) \arrow[r, mapsto] \arrow[d, mapsto] & (v) \arrow[r, mapsto] \arrow[dr, mapsto] & (q) \arrow[r, mapsto] & (t_s) \arrow[r, mapsto] \arrow[d, mapsto] & (t) \arrow[r, leftrightarrow] & (w) \\ (\pi) \arrow[r, mapsto] & (U) \arrow[rr, mapsto] \arrow[ur, mapsto] & & (e) \arrow[r, mapsto] \arrow[ur, mapsto] & (c_3). \arrow[ur, mapsto] & & \end{tikzcd}$$

The following attempt to relate our various conditions to the hypotheses stated by earlier authors is very incomplete. Not only could we not do justice to the work of many psychologists; we have also failed to study thoroughly some of the statisticians' work.

In Section II, the ideas of R. D. Luce, imposing a very strong constraint on the probabilities of multiple choices of any order, were identified with Condition (u); it can also be found in an early unpublished paper of Törnqvist [26]; and was stated orally (but taken back as unrealistic) by Debreu. R. A. Bradley and M. E. Terry [27] and L. R. Ford [28] applied this condition to binary choices only; in this form it is studied by J. Marschak [20] in some detail. A weaker constraint on the probabilities of binary choices is that of Fechner [3], identified with Condition (v). This condition, the still weaker Condition (w), and an intermediate case [the 'strong transitivity', (t_s)], was also studied by C. Coombs [29]; H. Gulliksen [30]; the late S. Valavanis-Vail [31], to whom we owe the important Condition (t_s); W. Edwards [32]; A. G. Papandreou *et al.* [33]; K. May [34]; J. M. Davis [35]; and D. Davidson (in collaboration with one of the present authors) [36].

In Section III, we identified an idea of Thurstone's with our Condition ($s.n$) (somewhat weakened in the subsequent work of Mosteller): a special form of combining the model (v) of strong constant utilities with the model (U) of random orderings. Conditions which can be identified with (U) were studied by G. Th. Gilbaud [25], N. Georgescu-Roegen ([22], [37]), D. Rosenblatt [38], R. D. Luce [15], J. Marschak [39], and others.

M. G. Kendall [40], G. L. Mallows [41], I. R. Savage [42], E. L. Lehmann [43], H. D. Brunk [44] and other statisticians (we owe these references to Lehmann) have studied rank-order statistics, 'ranking models', and paired comparisons, starting with problems of statistical inference. For example, given a sample of binary choices, estimate the 'underlying' (constant) ranking (perhaps a vector-parameter of the distribution of our U?); or test the hypothesis that a certain parameter is higher with one than with another of a pair of parent populations. To our regret, we have not had the opportunity to do the important job of stating the relations (if any) between our network of conditions and the stochastic models that may underlie the work of these authors. We shall briefly deal with statistical inference in Section X.

VII. VARYING THE DOMAIN OF TESTABILITY

As remarked in Section I, the definition of the class of basic observations, and therefore also of the class of testable conditions, depends on the range of possible experiments and other observations. Moreover, Condition (*o*) of Section IV exemplifies the case of a constraint that involves all elements of the set of alternatives simultaneously; if this set is finite and of unknown size, such a condition is not testable, whatever the definition of the basic observation.

In this paper, the basic observations have been defined as $(a; F)$, the actual *choice* of a single element a out of a feasible subset F of alternatives. It was suspected that the subject's verbal statements about how he *ranks* the elements of a given subset of three or more alternatives were less reliable, for the characterization and prediction of actual behavior, than his actual choices. But this suspecion may have been unnecessary. Psychologists do ask their subjects to rank three or more objects, according to the subject's intensity of perception, or his preferences, etc. They also ask 'judges' to rank individuals according to some characteristics. Presumably some consistency is assumed to prevail among all the rankings of any offered subsets of alternatives, and also between these rankings and the actual choices from any offered subset. In a stochastic model involving a finite set of alternatives $N=(1, 2, 3, 4)$, this would mean that, for example, a subject who, *at a given time*, ranks N in the order 1234, *would*, if asked to rank the subset (2, 3, 4), produce the ranking 234; and if asked to choose from the subset (2, 3, 4), *would* choose 2. The experimenter's faith in such consistency may be increased if the subject is told that the verbal statement about ranking will commit his actual choice when he is later presented with a pair of alternatives (a procedure suggested in a non-stochastic context by W. A. Wallis as quoted by L. J. Savage [45]), or, for that matter, with any subset of alternatives.[17] With the rankings (on the set N of all alternatives) admitted as basic observations, our condition (P) can be reformulated as a testable one, with $P(r)$ indicating the probability of an *observable* ranking r; that is, P is a distribution about which inferences can be made from experiments on rankings.

TESTABLE CONDITION (P). *The numbers* $P(r)$ *have the property* (3.7).

Since ranking experiments can be made on any subset $K \subseteq N$, we may, denoting by $P(r^K)$ the probability of a ranking r^K of K, generalize (P) into the following

TESTABLE CONDITION (P^K). *The numbers* $P(r^K)$ *have the property* (3.7) *with K replacing N, and with* $M \subseteq K \subseteq N$.

Note that when the subset K consists of two elements, the choosing from the subset and its ranking are the same thing. But the observation of ranking within triples and larger subsets does provide additional testable constraints not covered in the previous sections.

Testability is extended still further if the following procedure is permitted. The subject is given a set $M \subseteq N$, where M consists of m objects, and is told to select l of them ($l \leqq m$) and to rank these l objects. The ranking is denoted by $r^{L,M} = (r_{a_1} \ldots r_{a_l})$ where $L = \{a_1, \ldots, a_l\}$. The condition $(P_{l,m})$ is said to hold if there exist $n!$ non-negative numbers $P(r^{N,N}) = P(r)$ whose sum is unity such that, for each set M having at least l, but not more than m elements, and each subset L of M with l elements

$$(7.1) \qquad P(r^{L,M}) = \sum P(r),$$

the summation being over all rankings r such that $r_i \leqq r_j$ if either (1) $i \in L$ and $j \in M - L$, or (2) $i \in L$, $j \in L$, and $r_i^{L,M} \leqq r_j^{L,M}$.
Note: with $l = 1$, $(P_{1,n})$ *becomes the condition* (P) *of Section* 3, since (7.1) then becomes

$$i(M) = \sum_{r \mid r_i \leqq r_j,\, j \in M} P(r).$$

Thus $(P) \leftrightarrow (P_{1,n})$.

Clearly: $(P_{l,m}) \mapsto (P_{l,m-1})$, $(l < m)$; $(P_{m-1,m}) \leftrightarrow (P_{m,m})$. Moreover:

1. If $L = (i)$, then $P(r^{L,M}) = i(M)$.
2. If $L = (i)$, $M = (i, j)$, then $P(r^{L,M}) = P(ij)$.
3. If $L = (i, j)$, $M = (i, j)$, then $P(r^{L,M}) = P(ij)$.
4. If $L = N$ or $N - 1$, then $P(r^{L,N}) = P(r) = P(1_r 2_r \ldots n_r)$.

So far, we have discussed the possibility of extending the domain of testable conditions beyond the definition given in Section I. On the other hand, the observational possibilities may be such as to make even that definition too wide.

For example, they may be such as to make binary but not ternary, etc. choices observable. As remarked orally by H. Simon, to replace the experiment – in which the subject names the heaviest of two objects placed in his two hands – by a similar experiment with three objects (taken in successive pairs, or weighed on three fingers, etc.) may change experimental conditions so drastically as to invite inconsistencies. On the other hand, no such inconsistencies need be expected in experiments where the subject tells which of the two, or three, or possibly four lights is brighter. In the economics of the actual market, binary choices are often not observable where ternary or higher ones are; for example, a speculator will (1) buy, (2) sell, or (3) do neither of these.

When the testability domain is extended, some sufficient existence conditions become irrelevant. As a simple example let N consist of three elements and first assume that both binary and ternary choices are observable but rankings are not. Theorem 5.1 gives a necessary and sufficient condition for the existence of probabilities of rankings as defined in Condition (P) of Section III. But it does not say that these probabilities will be equal to the numbers $P(r)$ of the Testable Condition (P) of the present section, numbers about which inferences could be made if rankings were observable.

On the other hand, suppose that only binary choices among the three alternatives in N are observable, but neither ternary choice nor rankings are. Then Condition (c_3) $(1 \leqq xy+yz+zx \leqq 2)$ becomes important; for, by Theorem 5.7, this condition is sufficient for the existence of the three ternary choice probabilities $i(x, y, z)$ and of the six probabilities of rankings $P(r)$; and nothing stronger can be stated with the observations at hand.

VIII. THE BOUNDARY CASE: PERFECT DISCRIMINATION OR PREFERENCE

The case when the binary probability $p(a; F)$, with $F=(a, b)$, takes the boundary values 1 or 0 is called, in psychology of perceptions, perfect (as distinct from imperfect) discrimination. In the language of the theory of choice there is then perfect (as distinct from imperfect, or stochastic) preference. The term can be extended to the case when the feasible set F contains any number of elements.

To take account of the boundary case, special provisos are needed to make conditions such as (t_s), (v), (u) consistent with empirical observations. To take the weakest of those conditions, (t_s) or its equivalent

(t_s^{**}): if there exists an element k of the set N of alternatives, such that $ik=jk$, then $ij=\frac{1}{2}$. Now, if three bodies i, j, k weigh, respectively, 1000, 10, and 6 grams, and xy is the probability that the subject says that x is heavier than y, then the observations will very likely show $ij=ik$; but they will very unlikely show $jk=\frac{1}{2}$. The conditions (t) and (w), both weaker than (t_s), are not endangered by this fact. But the stronger conditions (v) and (u) are. Without the proviso about boundary cases, made in Section II and again in Section IV, Condition (v) would imply, in our case of three bodies, $v_j=v_k$ for each of the lighter ones, because of the existence of the very heavy body i. Similarly, in Condition (u) all the (strict) utilities u_i had to be assumed positive to prevent any of the probabilities $i(M)$ from being 0 or 1.

Perfect discrimination is circumvented experimentally by presenting to the subject, at each trial, objects not too strongly differing from each other in the relevant physical dimension; and by keeping the subject uninformed of the physical dimension – for example, through blindfolding.

This approach is not sufficient, for experiments on choice, to the extent to which these are intended to reveal 'desirabilities,' with 'perceptibility' kept more or less constant (see Section 1). Ideally, one would like to give the subject all the physical information, making the choice dependent only on his 'desires'. Suppose, then, that two alternatives differ from each other, however slightly, with respect to a single characteristic which is easily identified by the subject and is completely ordered with respect to the subject's choices. He will then presumably show perfect preference. Thus, in particular, when a and b are quantities of an economic 'good' and a exceeds b, he will prefer a with probability 1, even if the difference in quantities is very small. (In fact, this may be regarded as the definition of a 'homogeneous' economic good.) Also, if a and b are two vectors of quantities of the same m goods, $a=(a_1, \ldots, a_m)$, $b=(b_1, \ldots, b_m)$, and if a dominates b, i.e., if $a_i \geqq b_i$ for all i and $a_j > b_j$ for some j, then a will be perfectly preferred to b. The same is true if the vectors a and b are interpreted as two money wagers, (where a_i and b_i are, respectively, the monetary gains awarded when event E_i happens) and a dominates b. On the other hand, if a and b are two vectors neither of which dominates the other, stochastic instead of perfect preference may occur. Accordingly, the pairs of commodity-bundles used in the experiments by Papandreou and others [33] and the pairs of wagers used in the experiments of Davidson and

Marschak [36] were such as to exclude domination.

The vacillation of the chooser between the alternatives has thus been ascribed by Simon [46] and by Quandt [47] to the presence, in each alternative, of several relevant characteristics that themselves are completely ordered, but to which the subject pays attention only one at a time. In general, we need not assume that every alternative is so reduced (albeit 'unconsciously') to a bundle of characteristics, each of them an ordered one, so that the businessman decomposes, as it were, the job-applicant into his 'personality factors', or considers a prospective plant site as a vector of numerical attributes. Yet in some important cases such vectorial representation is indeed manifest. This gives rise to disturbing discontinuities in applying our stochastic models. For example, denote by $a=(a_1, a_2)$, $b=(b_1, b_2)$, and $b'=(b_1', b_2')$, three pairs of 'homogeneous commodities' – apples and oranges, not factories and vice-presidents; let ε_1 and ε_2 be two small positive numbers, and let $b_1=b_1'=a_1+\varepsilon_1$, $b_2=a_2$, $b_2'=a_2-\varepsilon_2$. Then, very likely, the binary probability ba is 1; while $b'a$ may be quite close to $\frac{1}{2}$; this big jump in probabilities would result from a very small physical change in the alternatives offered.

For such reasons, the psychophysicist's proviso that the 'differences' should not be 'noticed always or never' is doubly important in designing and interpreting experiments on economic choices,[18] if one wants such experiments to throw light on desirabilities, as separated from the perceptibility aspect of the choice phenomenon.

IX. REMARKS ON CHOICE UNDER UNCERTAINTY

As visualized at the end of Section I, we shall now outline possible stochastic reformulations of the theory of choice under uncertainty, weakening in particular the 'expected utility' condition.

A wager associates each alternative (in the sense defined previously) $i=1,\ldots, n$ with an event. Denote by $q=(q_1,\ldots, q_n)$ any wager in which the event giving rise to alternative i has probability q_i (possibly zero). In its non-stochastic form, the expected utility condition says that there is a constant vector $\bar{\omega}=(\bar{\omega}_1,\ldots, \bar{\omega}_n)$, unique up to an increasing linear transformation, and such that

$$(9.1) \qquad \sum\bar{\omega}_i q_i \geqq \sum\bar{\omega}_i q_i' \quad \text{if and only if} \quad q \succsim q',$$

where $\gtrsim$ means 'preferred or indifferent'. Since the observer cannot be sure that two subjects behave as if they ascribed the same probability to any given event, not only the 'utility-numbers' $\bar{\omega}_i$ but also the probabilities q_i have to be interpreted in a subjective sense, thus: For a given subject, there exist a vector $\bar{\omega}$ and, associated with each wager, a vector q such that (9.1) is satisfied.

It has been shown[19] that this condition is implied by other conditions (called 'axioms') that apparently appeal more immediately to what our intuition recommends as rational, or consistent. As we remarked at the beginning, consistency in an absolute sense is likely to be rejected by experience.[20] But it can be replaced by more general conditions, of stochastic consistency; these imply, or are implied by, or both, statistically testable hypotheses. In particular, each of the Conditions (w), (v), (u), (U) of our Sections II and III can be combined with the expected utility principle to yield a stronger condition to be denoted by ($\bar{\omega}$), ($\bar{v}$), ($\bar{u}$), ($\bar{U}$), respectively. In formulating these conditions, we shall denote by Q the set of all wagers (i.e., the set of all vectors q of order n, with $\sum q_i = 1$, $q_i \geqq 0$, all i); the feasible set of wagers will be denoted by $F \subseteq Q$; the probabilities of choices will be denoted as before, e.g., $q(F) = p(q; F)$; $qq' = p(q; (q, q'))$.

CONDITION ($\bar{w}$). *There is a constant vector* $\bar{w} = (\bar{w}, \ldots, \bar{w}_n)$ *such that*

$$(9.2) \qquad \sum \bar{w}_i q_i \geqq \sum \bar{w}_i q'_i, \quad \textit{if} \quad qq' \geqq \tfrac{1}{2}.$$

CONDITION ($\bar{v}$). *There is a constant vector* $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_n)$ *and, associated with it, a distribution function* $\varphi_{\bar{v}}$, *strictly increasing except at its values* 1 *and* 0, *such that*

$$(9.3) \qquad \varphi_{\bar{v}}\left(\sum \bar{v}_i q_i - \sum \bar{v}_i q'_i\right) = qq'.$$

CONDITION ($\bar{u}$). *There is a constant positive vector* $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_n)$ *such that, for every* $F \subseteq Q$ *and every* q, q' *in* F,

$$(9.4) \qquad \frac{\sum q_i \bar{u}_i}{\sum q'_i \bar{u}_i} = \frac{q(F)}{q'(F)}.$$

CONDITION ($\bar{U}$). *There is a random vector* $\bar{U} = (\bar{U}_1, \ldots, \bar{U}_n)$ *such that for every* $F \subseteq Q$ *and every* q *in* F, (9.5) $\quad q(F) = \Pr\{\sum (q_i - q_i')\,\bar{U}_i \geqq 0\,; \quad all \quad q' \text{ in } F\}$.

Clearly $\bar{w}$ and $\bar{U}$ (unlike w and U) are unique up to an increasing linear transformation. So is $\bar{v}$ (by virtue of the continuity of the set Q; see note 8). And $\bar{u}$ is unique up to a positive multiplier (as is u).

The directly testable conditions of the previous sections – e.g., Conditions (t), (t_s), (q), etc. – can be appropriately reformulated, and some new ones added. A case most easily accessible to observations is that of even-chance wagers: i.e., the case when the subjective probabilities are $\frac{1}{2}$ for each of a pair of alternatives. Let iEj be the wager promising i if event E happens, and j otherwise; then the subjective probability of E (and therefore also each of the subjective probabilities q_i and q_j) is said to be equal to $\frac{1}{2}$ if the subject is indifferent between iEj and jEi. Or, using the stochastic interpretation and denoting by $(iEj \cdot jEi)$ the probability of choosing the first out of that pair of wagers, $q_i = q_j = \frac{1}{2}$ if $(iEj \cdot jEi) = \frac{1}{2}$. In principle this property is operationally ascertainable.[21] Consider an even-chance wager, with $q_1 = q_2 = \frac{1}{2}$, and another, $q_3' = q_4' = \frac{1}{2}$; the expressions $\sum q_i \bar{v}_i$ and $\sum q_i' \bar{v}_i$ in Condition ($\bar{v}$) become, respectively, $(\bar{v}_1 + \bar{v}_2)/2$ and $(\bar{v}_3 + \bar{v}_4)/2$; and so with similar expressions in ($\bar{u}$), ($\bar{w}$) and ($\bar{U}$). We can denote the even-chance wager aEb simply by $\overline{ab}$ (hence $\overline{aa} = a$), and the probability of choice between two such wagers by $(\overline{ab} \cdot \overline{cd})$. Then Condition ($\bar{v}$) can be rewritten – analogous to relation (4.1) and with the proviso that the considered probabilities of choice are not 1 or 0 – thus:

$$(9.6) \quad \begin{array}{ll} \textit{if} & (\overline{a_1 a_2} \cdot \overline{b_1 b_2}) \geqq (\overline{b_3 b_4} \cdot \overline{a_3 a_4}), \\ \textit{then} & v_{a_1} + v_{a_2} + v_{a_3} + v_{a_4} \geqq v_{b_1} + v_{b_2} + v_{b_3} + v_{v_4}. \end{array}$$

This clearly implies the following *directly testable* set of conditions:

CONDITION (8_E) (8-tuple condition on even-chance wagers). *If* $(\overline{a_1 a_2} \cdot \overline{b_1 b_2}) \geqq (\overline{b_3 b_4} \cdot \overline{a_3 a_4})$, *then* $(\overline{a_{r_1} a_{r_2}} \cdot \overline{b_{s_1} b_{s_2}}) \geqq (\overline{b_{s_3} b_{s_4}} \cdot \overline{a_{r_3} a_{r_4}})$, *where* (r_1, r_2, r_3, r_4) *and* (s_1, s_2, s_3, s_4) *are arbitrary permutations of* (1, 2, 3, 4).

THEOREM 9.1. ($\bar{v}$) → (8_E).

It can be shown that the set (8_E) of conditions is not only necessary but also sufficient for (9.6) provided the stochastic continuity condition ($s.c$) of Section 4 is satisfied with respect to the considered set of even-chance wagers.

If some of the eight alternatives in Condition (8_E) are made identical, weaker directly testable conditions arise, involving a smaller number of distinct alternatives, and necessary for Condition ($\bar{v}$). In certain conjunctions, and assuming stochastic continuity, they are also sufficient for ($\bar{v}$), over the considered set of even-chance wagers. They were studied by Davidson and Marschak [36]; Debreu's paper [54], though dealing with even-chance wagers in a non-stochastic context, is also relevant.

X. TESTING STOCHASTIC CONSTRAINTS WHEN REPLICATIONS ARE FEW[22]

All the considered conditions have been constraints on probabilities of certain responses of the subject to varying external conditions. More precisely, these probabilities have been in the form $p(a; F)$ although in Section VII more general forms were briefly discussed. These probabilities define the coordinates of the parameter space, and each directly testable condition ('hypothesis') defines a certain region of this space. The test consists in deciding, on the basis of observations, whether the parameter point – the set of probabilities $p(a; F)$ for all F – falls into that region.

A trial is easily replicated if the same feasible set F – the same set of stimuli – can be presented to a random sample of subjects *provided* it is reasonable to assume that each member of the parent population is characterized by the same relevant probability distribution of responses. This assumption has been made, in effect (with respect to choices of foods, birthday gifts, etc.) by Jones, Peryam, and Thurstone [55], Thurstone and Jones [56], and Gulliksen [57].

If this assumption of identical distributions is not justified, the hypothesis must be tested separately for each subject to find the proportion of the observed subjects that fail to satisfy certain strong conditions – such as (q) (see Section IV) and therefore (v); the proportion of subjects failing to satisfy weaker conditions – such as (t) and therefore (w); and so forth. One can thus estimate the frequency distribution of the various 'degrees of stochastic consistency' among the parent population, even though the probabilities $(a; F)$ are not the same for two persons with the same degree of consistency. Two persons may both obey (w) and both fail to satisfy (v), even though one of them tends to vegetarianism more than the other. In psychophysics, two persons may obey the Fechnerian condition (v)

and yet have different functions φ_v, e.g., have different quartiles or 'just noticeable' differences (see Section II). A frequency distribution of degrees of (stochastic) consistency, in a given culture, is a useful thing to know. One may be interested in training people towards greater consistency, because of its importance for decision-making and leadership. This is not the same thing as inculcating identical tastes.

The number of replications, *for the same subject*, of the same set F of stimuli is limited by the fact that, in the course of successive replications, the subject's characteristic probabilities $p(a; F)$ for a given F may change. Inasmuch as this may be due to fatigue, psychophysicists do not seem to have found it necessary to curtail replications too much; they regard relative frequencies with which given responses a are elicited by a fixed F (e.g., a fixed pair of sounds) as workable estimates of the probabilities $p(a; F)$.

True, in psychophysical experiments the subject must be prevented from identifying the stimuli and remembering them. For example, he must be blindfolded. If he were asked to choose the heavier of two objects (or the string producing the higher pitch) but were not blindfolded, the visual stimulus, added to that of the weight (or sound), might help him to identify and remember the objects and thus possibly influence him when the same object is presented again.

For a similar reason, successive replications may lack independence when subjects have to respond to verbal stimuli, as in questionnaires on attitudes or on voting preferences. Students of attitudes try therefore to repeat a question in a different form, preferably after a lapse of a little time. The same lack of serial independence is likely to occur in the case of choices between verbally stated economic alternatives such as commodities and wagers. It might not be true of 'blindfolded' choices between wines. But, as pointed out in Section I, the economic theorist wants to separate preferences from difficulties of perception; his ideal is to observe the decisions of people while giving them as complete information as possible.

Accordingly, experimenters on choices have used only a small number of repetitions of the same pair of alternatives (Papandreou *et al.* [33]) or have avoided such repetitions altogether (Davidson and Marschak [36]). This creates special problems in devising statistical tests of significance, and at any rate makes such tests relatively weak. To bring out the logical nature of the problem, we shall concentrate on the mathematically easiest (though empirically weakest) case when each set F is presented once only; and we shall limit ourselves to the simplest of our hypotheses, the weak stochastic transitivity condition (t).

During the course of experiment the subject is asked to choose one from each of the sets (a, b), (b, c), (a, c). Denote by p_a, p_b, p_c the binary probabilities ab, bc, ca. Then $p=(p_a, p_b, p_c)$ is a point in $[0, 1]^3$. Assume that the subject has a certain probability distribution $\mathscr{P}$ of choosing points in $[0, 1]^3$; the random choice of triples (a, b, c) leads to a random choice from $[0, 1]^3$ with the distribution $\mathscr{P}$ on it. Let $q_i=p_i-\frac{1}{2}$, $i=a, b, c$. The region of $[0, 1]^3$ where condition (t) is *not* satisfied (call it T) is characterized by the fact that all three q_i are $\geqq 0$ or $\leqq 0$ and are not all $=0$. That is, apart from a set of measure 0, all three q_i have the same sign. It is desired to estimate $\mathscr{P}(T)=\Theta$.

It may be useful to think of an analogous one-dimensional problem. A coin-making machine is characterized by an unknown probability distribution of the chance variable p (probability of a coin falling heads). One is permitted to toss coins, *each only a few times* (*perhaps only once*), in order to get evidence about the distribution of p; for example, one wants to test the hypothesis that the proportion of coins with $p>\frac{1}{2}$ (i.e., biased in favor of heads) is not more than a preassigned number. The region '$p>\frac{1}{2}$' clearly corresponds to our region T, and can be so denoted; the preassigned number is an upper bound Θ' on $\mathscr{P}(T)=\Theta$. Now, an infinite number of coins of which exactly 10% are biased for heads, would, if each is tossed once, show more than 5% (and not more than 55%) heads; hence, if out of a very large number of coins tossed, exactly $Z=5\%$ have fallen heads, all we can say is that the proportion of coins biased for heads is less than $\Theta'=10\%=2Z$. Thus, *regardless of the number of coins* we are able to toss (each once), we cannot in general make statements involving a pre-assigned upper bound Θ' on the proportion of coins belonging to a given region. If Θ' had been 15%, then we could assert $\Theta<\Theta'$, but if Θ' had been 7%, then we could not assert $\Theta<\Theta'$. Thus, if we want to be sure that we can make an assertion, we can choose Θ' only after the tossings have been observed; if we are willing to accept the possibility of making no assertion, we can preassign Θ'. Similarly, in testing the stochastic transitivity condition (t), we shall be able to post-assign but not always to pre-assign an upper bound Θ' on Θ; and this regardless of the number of triples (a, b, c) we can present – each once – to the subject. The situation there is further complicated by the fact that we have only a finite sample and we introduce other parameters t^0, t' to analyze the confidence level.

Furthermore, if each coin could be tossed twice, we would still be unable to guarantee an assertion with a pre-assigned upper bound to Θ, but the ratio $\Theta'/\bar{Z}$ of the post-assigned bound to the observed proportion of heads would become smaller. And a similar statement can be made about the consequences of duplicating the presentation of each triple (a, b, c) to the subject.

We shall now develop a test of Condition (t). Let

$$(10.1)\quad f(p) = p_a p_b p_c + (1-p_a)(1-p_b)(1-p_c) = \\ = \tfrac{1}{4} + q_a q_b + q_b q_c + q_c q_a,$$

where $q_i = p_i - \frac{1}{2}$. If $p \in T$ then $f(p) \geqq \frac{1}{4}$. If $p \notin T$ then $f(p) \leqq \frac{1}{2}$, for if $q_a = -\alpha, q_b = \beta, q_c = \gamma$ (α, β, γ non-negative), $f(p) = \frac{1}{4} + \beta\gamma - \alpha(\beta+\gamma) \leqq \leqq \frac{1}{4} + \beta\gamma \leqq \frac{1}{2}$, and the other possibilities are covered by symmetry. Define three random variables X_a, X_b, X_c: $X_a = 1$ if a is chosen from (a, b) and $X_a = 0$ otherwise; $X_b = 1$ if b is chosen from (b, c) and $X_b = 0$ otherwise; $X_c = 1$ if c is chosen from (a, c) and $X_c = 0$ otherwise. Define

$$(10.2)\quad Z = X_a X_b X_c + (1-X_a)(1-X_b)(1-X_c).$$

For a given vector p, $\Pr\{Z=1\} = f(p)$. Now $\Pr\{Z=1\} = \Pr\{Z=1 \mid p \in T\} \cdot \mathscr{P}(p \in T) + \Pr\{Z=1 \mid p \notin T\}\, \mathscr{P}(p \notin T)$. Hence $\Pr\{Z=1\} \geqq (\Theta/4) + 0$ and $\Pr\{Z=1\} \leqq \Theta + \frac{1}{2}(1-\Theta) = \frac{1}{2} + (\Theta/2)$. Thus Z is a binomial chance variable with mean $\mu = \Pr\{Z=1\}$ where

$$(10.3)\quad (\Theta/4) \leqq \mu \leqq \tfrac{1}{2} + (\Theta/2).$$

Let $Z_1, \ldots, Z_m$ be independent observations on Z, each based on a different triple (a, b, c). Let $\bar{Z} = \sum_1^m Z_k/m$. Our first test is based on the following considerations:

(1) Let $0 < \Theta' \leqq 1$. If $\Theta \geqq \Theta'$ then $\mu \geqq \Theta'/4$. Let V' be a binomial chance variable with $\Pr\{V'=1\} = \Theta'/4$. Then $\Pr\{Z_k = 0\} \leqq \Pr\{V'=0\}$. Let $t' \leqq \Theta'/4$. Then $\Pr\{\bar{Z} \leqq t'\} \leqq \Pr\{\bar{V}' \leqq t'\}$ where $\bar{V}'$ is the mean of m values of V'. If a confidence level α' is specified, one can choose m large enough so that $\Pr\{\bar{V}' \leqq t'\} \leqq \alpha'$. Hence:

$$(10.4')\quad \text{if } \Theta \geqq \Theta',\ \Pr\{\bar{Z} \leqq t'\} \leqq \alpha'.$$

(2) Similarly, consider Θ^0, with $0 \leqq \Theta^0 < \frac{1}{2}$. If $\Theta \leqq \Theta^0$ then $\mu \leqq \frac{1}{2} + \Theta^0/2$. Let V^0 be a binomial random variable with $\Pr\{V^0 = 1\} = \frac{1}{2} + \Theta^0/2$. Then $\Pr\{Z_k = 1\} \leqq \Pr\{V^0 = 1\}$. Let $t^0 \geqq \frac{1}{2} + \Theta^0/2$. Then $\Pr\{\bar{Z} \geqq t^0\} \leqq \Pr\{\bar{V}^0 \geqq t^0\}$ where $\bar{V}^0$ is the mean of m values of V^0. Again, with a specified confidence level α^0, we can choose large enough m to make $\Pr\{\bar{V}^0 \geqq t^0\} \leqq \alpha^0$. Hence

(10.4^0) if $\Theta \leqq \Theta^0, \Pr\{\bar{Z} \geqq t^0\} \leqq \alpha^0$.

Summarizing: If $\bar{Z} \geqq t^0$, we assert $\Theta > \Theta^0$; if $\bar{Z} \leqq t'$, we assert $\Theta < \Theta'$; if $t' < \bar{Z} < t^0$, we assert nothing. The consequences of this are shown in the following table of probabilities of occurrence:

Assertion	True state	
	$\Theta \leqq \Theta^0$	$\Theta \geqq \Theta'$
$\Theta > \Theta^0$	$< \alpha^0$	–
$\Theta < \Theta'$	–	$< \alpha'$
Nothing	–	–

where the dashes indicate probabilities that we have not estimated. This procedure protects us against extreme errors of classification, but leaves open the possibility of coming to no decision.

If m is large enough so that the normal approximation to the binomial may be used, then the constants involved in the test are determined by the following equations:

(10.5) $F(s^0) = 1 - \alpha^0, \quad F(s') = 1 - \alpha'$,

where F is the normal distribution function with zero mean and unit variance, and

(10.6′) $s' = (\Theta' - 4t')\sqrt{m}\,/\,2\sqrt{\Theta'(1 - \Theta'/4)}$,

(10.6^0) $s^0 = (2t^0 - 1 - \Theta^0)\sqrt{m}\,/\sqrt{(1 + \Theta^0)(1 - \Theta^0)}$.

The experimenter may specify the parameters $\Theta^0, \Theta', t^0, t'$ either before or after the experiment. Let us first discuss the case where the choice is made *after* he has the data. Suppose $\bar{Z}$ turns out to be quite high. He is then seeking to make an assertion $\Theta > \Theta^0$. He can do this with confidence level $1 - \alpha^0$ where α^0 is determined from (10.6^0), provided that $t^0 \leqq \bar{Z}$ and $\Theta^0 \leqq 2t^0 - 1$. He would like Θ^0 to be as high as possible and at the same time $1-\alpha^0$ to be as high as possible. As he takes Θ^0 closer to $2t^0-1$, s^0 tends to zero and α^0 tends to $\frac{1}{2}$, which is not a convincing confidence level. Thus to keep the confidence level $1-\alpha^0$ high he must take Θ^0 much less than $2t^0-1$; on the other hand, the smaller he takes Θ^0, the weaker the impact of his assertion $\Theta>\Theta^0$. Thus it is clear that he should choose t^0 as large as possible, namely Z. Now if the experimenter wants to work with a fixed confidence level $1-\alpha^0$, he should adjust Θ^0 low enough so that he achieves, by (10.6^0), his desired confidence level. If, on the other hand, there is a fixed 'tolerance level' Θ^0 which the experimenter regards as the significant indicator of inconsistency, then he would use this fixed Θ^0 and again with $t^0=Z$ find his confidence level by (10.6^0). More generally, he can strike what balance he wishes, pushing Θ^0 up while driving $1-\alpha^0$ down, keeping of course $t^0=Z$. If, instead, Z is quite low, then he wishes to assert $\Theta<\Theta'$. Here he would like Θ' as small as possible and $1-\alpha'$ as large as possible. Considerations similar to those above, applied now to (10.6′), show that he should choose $t'=Z$ and similarly make an arrangement between Θ' and $1-\alpha'$.

It may be, however, that the experimenter is doing many experiments and would like to have a fixed procedure, regardless of the data. He can specify t^0 and Θ^0 in advance, subject to $\Theta^0 \leqq 2t^0-1$, and know that his confidence level is given by (10.6^0). (The larger $2t^0-1-\Theta^0$, the larger the confidence level, but the larger he takes t^0, the more likely he will have nothing to assert, while the smaller he takes Θ^0, the weaker are the assertions he does make; this balance he settles in advance.) Similarly he selects Θ', t' in advance. Now the data from many experiments can be handled by a fixed procedure; if $Z \geqq t^0$ or $Z \leqq t'$, an assertion is made, while if $t'<Z<t^0$, no assertion is made. This procedure will be more convenient, but clearly it involves a loss of one or more of the following: confidence level, strength of assertion, and likelihood of making an assertion, as compared with the method of selecting the parameters after the experiment. However, if the experimenter is testing a large number of individuals in order to select a subset for further testing, he might find it better, as far as the data-processing is concerned, to fix the parameters in advance.

Our second method is to find a confidence interval for Θ. Let $0<\alpha<1$ and $0<\delta<1$ be given. Then it follows that for m sufficiently large $P(-\delta<Z-\mu<\delta)>1-\alpha$; i.e., $P(Z-\mu<\delta$ and $Z-\mu>-\delta)>1-\alpha$, or $P(Z<\delta+\mu$ and $Z>\mu-\delta)>1-\alpha$; from (10.3) it follows that $P(Z<\delta+$

$+\frac{1}{2}+(\Theta/2)$ and $Z>(\Theta/4)-\delta)>1-\alpha$. Thus $P(2Z-2\delta-1<\Theta<4Z+$ $+4\delta)>1-\alpha$.

While it is true that the length of the confidence interval $(2Z-2\delta-1,$ $4Z+4\delta)$ is $2Z+6\delta+1>1$, it is not centered at $\frac{1}{2}$. Hence, if Z is near zero or unity, the effective length of the interval may be quite small; e.g., if $Z=0$, the conclusion is $-2\delta-1<\Theta<4\delta$, which has an effective length of only 4δ; while if $Z=1$, the conclusion is $1-2\delta<\Theta<4+4\delta$, which has an effective length of 2δ. Since for a prescribed significance level $1-\alpha$ we do not know δ explicitly, we can use the estimate $\delta=k(\alpha)\ \sigma=k(\alpha)\cdot$ $\cdot(pq/m)^{1/2}\leqq k(\alpha)/2(m)^{1/2}$. Thus

$$P\left(2Z-\frac{k(\alpha)}{\sqrt{m}}-1<\Theta<4Z+\frac{2k(\alpha)}{\sqrt{m}}\right)>1-\alpha.$$

If the normal approximation for the binomial is used, $k(\alpha)$ is determined from the equation

$$(10.7)\qquad F(k(\alpha))=1-\alpha/2,$$

where F is the normal distribution with zero mean and unit variance.

The confidence interval may also be used for making decisions; e.g. let $0<\Theta^*\leqq\Theta^{**}<1$. If the confidence interval is contained in the interval $(0,\ \Theta^*)$, assert that $\Theta\leqq\Theta^*$; if the confidence interval is contained in the interval $(\Theta^{**},\ 1)$ assert that $\Theta\geqq\Theta^{**}$, otherwise assert nothing. In other words, if

$$Z<\frac{\Theta^*}{4}-\frac{k(\alpha)}{2\sqrt{m}}.$$

assert that $\Theta\leqq\Theta^*$; if

$$Z>\tfrac{1}{2}\left(\Theta^{**}+1+\frac{k(\alpha)}{\sqrt{m}}\right),$$

assert that $\Theta\geqq\Theta^{**}$; otherwise assert nothing. We then get the following table of probabilities.[23]

Assertion	True state		
	$0\leqq\theta<\theta^*$	$\theta^*\leqq\theta\leqq\theta^{**}$	$\theta^{**}<\theta\leqq 1$
$\theta\leqq\theta^*$	–	$<\alpha$	$<\alpha$
$\theta\geqq\theta^{**}$	$<\alpha$	$<\alpha$	–
Nothing	–	–	–

NOTES

* Research carried out mostly under contracts of the Office of Naval Research with the Cowles Foundations for Economics at Yale University and with Columbia University. The paper was completed while Jacob Marschak was Ford Visiting Research Professor at the Graduate School of Industrial Administration, Carnegie Institute of Technology. We owe much to discussions with D. Davidson, G. Debreu, M. DeGroot, L. Hurwicz, and R. D. Luce.

1 While the stochastic approach uses measurable *probabilities*, another approach, also promising to give realism and unity to the understanding of choices and responses, makes use of the measurable *time-delay* of a choice or response. The shorter the delay, the larger is the difference between certain values said to be attached by the subject to two alternative responses. A very long delay reveals a state of (almost) indifference, the 'conflict-situation': Hamlet took a very long time to decide whether to kill his uncle. The logic of this approach and its combination with stochastic theories was treated by Cartwright and Festinger [1].

2 As pointed out by James Tobin in a private communication. *Added in 1974*: Tversky [59, p. 295] regards this distinction as amenable to his experiments on multi-attribute alternatives. See also [60].

3 Preferences thus defined are 'independent of irrelevant alternatives', a principle pointed out by Hotelling's pupil Arrow [4] and much used ever since. A better name for Arrow's principle is 'irrelevance of added alternatives'.

4 The condition is, *A is perfectly separable and is ordered by the relation* '$\gtrsim$'; *and for every a' in A the sets* $(a \mid a' \gtrsim a)$ and $(a \mid a \gtrsim a')$ *are closed*. The condition is trivially fulfilled by every finite set, for it is perfectly separable and all its, subsets are closed. Also, every subset of a finite-dimensional Euclidean space is perfectly separable, and it is reasonable to assume that the sets $(a \mid a' \gtrsim a)$ and $(a \mid a \gtrsim a')$ are closed when '$\gtrsim$' means a consumer's preference over the space of a finite number of commodities.

5 Thus Abraham Wald [6] has suggested a method to evaluate the consumer's ordinal utility function ω from the basic observations $(a; F)$ – purchases made as F varies with incomes and prices – assuming a non-stochastic model and not attempting a statistical estimate. More recently, H. Theil and H. Neudecker [7] gave a stochastic generalization of the same model, specifying, as did Wald, ω to be quadratic in the quantities of the goods consumed.

6 Regarding the proviso in brackets, see Section VIII.

7 These two points of view, the empirical and the game-theoretical, are represented in somewhat modified form by Dahl [12] and by Shapley and Shubik [13], respectively. If one accepts John Harsanyi's proposal to measure the power of X over Y by "how much difference it makes to Y between being X's friend or enemy," this difference between maxmax and maxmin payoffs (see, e.g., Luce and Raiffa [14]) might be related to the probability of Y's decision to be a friend.

8 For, suppose that $\varphi_{v'}(v'(x)-v'(y)) \equiv \varphi_v(v(x)-v(y))$. Then with $\phi=\varphi_{v'}^{-1}\varphi_v$ we have (i) $v'(x)-v'(y)=\phi(v(x)-v(y))$, (ii) $v'(y)-v'(z)=\phi(v(y)-v(z))$, and (iii) $v'(x)-v'(z)=\phi(v(x)-v(z))$. Adding (i) and (ii) and using (iii) we find, letting $a=v(x)-v(y)$, $b=v(y)-v(z)$, that $\phi(a+b)=\phi(a)+\phi(b)$. The only solution of this last equation which is not unbounded in every interval is, as is well known, $\phi(a)=\beta a$. Then from (i), with y fixed, we have $v'(x)=\beta v(x)+\alpha$.

9 Occasionally (in his Section 2.D.3, entitled 'A Generalization to Two or More Alternatives') Luce combines Condition (u) with another condition which we shall call (u')

(existence of strict *dis*utilities of *last* choices) and which is not generally consistent with (u): see our Theorem 3.8 below.

10 Luce tries to convey this appeal by an analogy with Arrow's (non-stochastic) principle which we called before the 'irrelevance of added alternatives'.

11 This counterexample is due to C. Winsten. Another interesting proof is due to P. Halmos. Three dice, A, B, C, are loaded so as to turn up with the following probabilities:

Face no.		1	2	3	4	5	6
Die:	A	0	0	.5	5	0	0
	B	0	.6	0	0	0	.4
	C	.4	0	0	0	.6	0

Let a, b, c (independent chance variables) be the number of spots turning up on A, B, C, respectively. Then (P) is satisfied: the probabilities of the six possible rankings of the numbers a, b, c according to magnitude, must be consistent with the probabilities that one of the numbers is the largest in a given pair or triple. But (t) is not satisfied: $\Pr\{a>b\}=.6>\frac{1}{2}$; $\Pr\{b>c\}=.64>\frac{1}{2}$; $\Pr\{a>c\}=.4<\frac{1}{2}$.

12 Whether or not (e) implies (P) will be left open until Section V.

13 Block and Marschak [18] searched for an appropriate urn model, and Debreu [19] found it, thus providing a shorter proof.

14 *Added in 1973*: See now Section 6.3., An 'Impossibility' Theorem, in Luce and Suppes [58].

15 For methods to construct these counterexamples see Block and Marschak [21]. Georgescu-Roegen [22] and Chipman [23] obtained a condition intermediate in strength between (t_s) and (t) by substituting min for max in (t_s). See Marschak ([20]), Section 7.)

16 Since (5.12) can be rewritten $ij+jk\geqslant ik$, this condition has also been called 'triangular'. Its further properties are studied in Marschak [20].

17 On the other hand, the following procedure will not do for our present purposes (it was suggested in an earlier paper of Marschak [39]): to construct the subject's ranking of a set by asking him to make a choice from it; then removing the alternative chosen, and asking him to make a choice from the remaining subset; and so on. In the stochastic case, the subject may shift during the procedure from one ranking to another ranking of the same set. In [39] the case was non-stochastic.

18 The case of absolute preferences is discussed by N. Georgescu-Roegen [22] in the context of the economic theory of demand (for homogeneous commodities). J. S. Chipman [23] drew attention to important anomalies in the case of statistical decisions.

19 This was done by J. von Neumann and O. Morgenstern [48], with q interpreted 'objectively'; their 'axioms' and proof were subsequently simplified by various writers, most concisely by I. N. Herstein and J. Milnor [49]. The 'subjective' interpretation of q was introduced by F. P. Ramsey [50] and made precise by L. J. Savage [45]. See also W. Edwards [32]. R. D. Luce [51] generalized the stochastic theory of choices under uncertainty by introducing random subjective probabilities.

20 See the examples of actual behavior found by M. Allais [52] and discussed also by L. J. Savage ([45], pp. 101–4.)

21 For still greater experimental ease, Davidson and Marschak [36], following the device used by Davidson, Suppes and Siegel [53], and guided by considerations of continuity, assumed that, when the alternatives i, j, i', j' are money amounts, then the

even-chance event E has also the following testable property:

$$(iEj \cdot j' Ei) = 1 = (jEi \cdot i' Ej),$$

provided $i - i' = j - j' = \varepsilon$, a small positive amount.

[22] *Added in 1974:* See now DeGroot [61].

[21] Naturally, much stronger tests are obtained if some *a priori* restrictions are admitted. Thus, in the Davidson and Marschak paper [36] two alternative hypotheses were formulated to test a stochastic constraint, viz., the transitivity condition (t) or (t_s): (1) The probability-triples (ac, bc, ca) are distributed uniformly over $[0, 1]^3$; (2) they are distributed over the region defined by Condition (t) or Condition (t_s) uniformly. Mathematical simplicity is about the only serious claim in favor of these restrictions unless one invokes Laplace's 'principle of ignorance'.

BIBLIOGRAPHY

[1] Cartwright, D. and Festinger, L., 'A Quantitative Theory of Decision', *Psychol. Rev*. **50** (1943) 595–621.

[2] Hotelling, H., 'Demand Functions with Limited Budgets', *Econometrica* **3** (1935) 66–78.

[3] Fechner, G. Th., *Elemente der Psychophysik*, Breitkopf & Härtel, Leipzig, 1859.

[4] Arrow, K. J., *Social Choice and Individual Values*, Wiley, New York, 1951.

[5] Debreu, G., 'Representation of a Preference Ordering by a Numerical Function', Chap. 11 in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), *Decision Processes*, Wiley, New York, 1954.

[6] Wald, A., 'The Approximate Determination of Indifference Surfaces by Means of Engel Curves, *Econometrica* **8** (1940) 144–75.

[7] Theil, H. and Neudecker, H., 'Substitution, Complementarity and the Residual Variation around Engle Curves', *Rev. Econ. Studies* **25** (1957) 114–23.

[8] Bernoulli, D., 'Exposition of a New Theory on the Measurement of Risk' (translated by Dr. Louise Sommer), *Econometrica*, **22** (1954) 23–36.

[9] Marshall, A., *Principles of Economics*, 8th ed. Macmillan, London, 1920.

[10] Mosteller, F. and Nogee, P., 'An Experimental Measurement of Utility', *J. Polit. Econ*. **59** (1951) 371–404.

[11] Guilford, J. P., *Psychometric Methods*, 2d ed., McGraw-Hill, New York, 1954.

[12] Dahl, R. A., 'The Concept of Power', *Behav. Sci*. **2** (1957) 201–15.

[13] Shapley, L. and Shubik, M., 'A Method for Evaluating the Distribution of Power in a Committee System', *Amer. Polit. Sci. Rev*. **48** (1954) 787–92.

[14] Luce, R. D. and Raiffa, H., *Games and Decisions*, Wiley, New York, 1957.

[15] Luce, R. D., *Individual Choice Behavior*, Wiley, New York, 1959.

[16] Thurstone, L. L., 'A Law of Comparative Judgment', *Psychol. Rev*. **34** (1927) 273–86.

[17] Mosteller, F., 'Remarks on the Method of Paired Comparisons', *Psychometrika* **26** (1951) 3–9, 203–6, 207–18.

[18] Block, H. D. and Marschak, J. 'An Identity in Arithmetic', *Bull. Amer. Math. Soc*. **65** (1959) 123–24.

[19] Debreu, G., 'On "An identity in arithmetic"', *Proc. Amer. Math. Soc*. **11** (1960) 220–21.

[20] Marschak, J., 'Binary Choice Constaints on Random Utility Indicators', *Selected Essays*, **8**.

[21] Block, H. D. and Marschak, J., 'Random Orderings', Cowles Foundation Discussion Paper No. 42, Cowles Foundation, New Haven, Conn., 1957 (duplicated).
[22] Georgescu-Roegen, N, 'Threshold in Choice and the Theory of Demand', *Econometrica* **26** (1958) 157–68.
[23] Chipman, J. S. 'Stochastic Choice and Subjective Probability', Technical Report No. 2, School of Business Administration, Univ. of Minn., Minneapolis, Minn., 1958 (duplicated).
[24] Debreu, G., 'Stochastic Choice and Cardinal Utility', *Econometrica* **26** (1958) 440–44.
[25] Guilbaud, G. Th., 'Sur une difficulté de la théorie du risque', *Colloques internationaux du centre national de la recherche scientifique* (*économétrie*) **40** (1953) 19–25.
[26] Törnqvist, L., 'A Model for Stochastic Decision Making', Cowles Commission Discussion Paper, Economics 2100, Cowles Commission, Chicago, 1954 (duplicated).
[27] Bradley, R. A. and Terry, M. E., 'Rank Analysis of Incomplete Block Designs, I. The Method of Paired Comparisons', *Biometrika* **39** (1952) 324–44.
[28] Ford, L. R., 'Solution of a Ranking Problem from Binary Comparisons', *Amer. Math. Month.* **64** (1957) (Aug. suppl.), 28–33.
[29] Coombs, C., On the Use of Inconsistency of Preferences in Psychological Measurement', *J. Exp. Psychol.* **55** (1958) 1–7.
[30] Gulliksen, H., 'Paired Comparisons and the Logic of Measurement', *Psychol. Rev.* **53** (1946) 199–213.
[31] Valavanis-Vail, S., 'A Stochastic Model for Utilities', Seminar on the application of mathematics to social sciences, Univ. Michigan, Ann Arbor, 1957 (duplicated).
[32] Edwards, W., 'Probability Preferences in Gambling', *Amer. J. Psychol.* **66** (1955) 349–64.
[33] Papandreou, A. G. (with the collaboration of O. H. Sauerlender, O. H. Brownlee, L. Hurwicz, W. Franklin), 'A test of a Stochastic Theory of Choice, *Univ. Calif. Publ. in Econ.* **16** (1957) 1–18.
[34] May, K., 'Intransitivity, Utility and the Aggregation of Preference Patterns', *Econometrica* **22** (1954) 1–13.
[35] Davis, J. M., 'The Transitivity of Preferences', *Behav. Sci.* **3** (1958) 26–33.
[36] Davidson, D. and Marschak, J., 'Experimental Tests of Stochastic Decision Theories', *Selected Essays*, 6.
[37] Georgescu-Roegen, N., 'The Pure Theory of Consumer Behavior', *Quart. J. Econ.* **50** (1936) 545–93.
[38] Rosenblatt, D., 'On Some Stochastic Process Formulations of Individual Preference and Consumer Behavior (abst.)', *Econometrica* **24** (1956) 347–48.
[39] Marschak, J., 'Norms and Habits of Decision Making under Certainty', *Selected Essays*, 5.
[40] Kendall, M. G., 'Further Contributions to the Theory of Paired Comparisons', *Biometrics* **11** (1955) 43–67.
[41] Mallows, G. L., 'Non-Null Ranking Models, I', *Biometrika* **44** (1957) 114–30.
[42] Savage, I. R., 'Contributions to the Theory of Rank Orders', *Ann. Math. Stat.* **27** (1956) 590–615; **28** (1957) 968–77.
[43] Lehmann, E. L., 'Ordered Families of Distributions', *Ann. Math. Stat.* **26** (1955) 399–419.
[44] Brunk, H. D., 'Mathematical Models for Ranking from Paired Comparisons',

Technical Report No. 11, Air Force Office of Scientific Research, Department of Mathematics, University of Missouri, Columbia, Mo., Nov. 1958 (duplicated).
[45] Savage, L. J., *The Foundations of Statistics*, Wiley, New York, 1954. Revised, Dover, 1972.
[46] Simon, H. A., 'A Behavioral Model of Rational Choice', *Quart. J. Econ.* **69** (1955) 99–118.
[47] Quandt, R. E., 'A Probabilistic Theory of Consumer Behavior', *Quart. J. Econ.* **70** (1956) 507–36.
[48] Neumann, J. von and Morgenstern, O., *Theory of Games and Economic Behavior*, 2d ed. Princeton Univ. Press, Princeton N.J., 1947.
[49] Herstein, I. N. and Milnor, J., 'An Axiomatic Approach to Measurable Utility' *Econometrica* **21** (1953) 291–97.
[50] Ramsey, F. P., *The Foundations of Mathematics and Other Logical Essays*, Kegan Paul, London 1931.
[51] Luce, R. D., 'A Probabilistic Theory of Utility', *Econometrica* **26** (1958) 193–224.
[52] Allais, M., 'Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine', *Econometrica* **21** (1953) 503–46.
[53] Davidson, D. and Suppes, P. (in collaboration with S. Siegel), *Decision Making: An Experimental Approach*, Stanford Univ. Press, Stanford, Calif., 1957.
[54] Debreu, G., 'Cardinal Utility for Even-Chance Mixtures of Pairs of Sure Prospects', *Rev. Econ. Studies* **26**(1958/9) 174–77.
[55] Jones, L. V., Peryam, D. R., and Thurstone, L. L., 'Development of a Scale for Measuring Soldiers' Food Preferences', *Food Research* **20** (1955) 512–20.
[56] Thurstone, L. L. and Jones, L., 'The Rational Origin for Measuring Subjective Values', *J. Amer. Stat. Assoc.* **52** (1957) 458–71.
[57] Gulliksen, H., 'Measurement of Subjective Values', *Psychometrika* **21** (1956) 229–44.
[58] Luce, R. D. and Suppes, P., 'Preference, Utility, and Subjective Probability', *Handbook of Mathematical Psychology*, Vol. III, Ch. 19, Wiley, 1965.
[59] Tversky, A., 'Elimination by Aspects: A Theory of Choice', *Psychological Review* **79** (1972) 281–99.
[60] Tversky, A., 'Intransitivity of Preferences', *Psychological Review* **76** (1969) 31–48.
[61] DeGroot, M. H., 'Optimal Allocation of Observations', *Annals of the Institute of Statistical Mathematics*, **18** (1966) 13–28.

ECONOMIC INFORMATION, DECISION, AND PREDICTION

SELECTED ESSAYS: VOLUME I

THEORY AND DECISION LIBRARY

AN INTERNATIONAL SERIES

IN THE PHILOSOPHY AND METHODOLOGY OF THE

SOCIAL AND BEHAVIORAL SCIENCES

VOLUME 7

JACOB MARSCHAK

ECONOMIC INFORMATION, DECISION, AND PREDICTION

Selected Essays: Volume I

PART I

Economics of Decision

D. REIDEL PUBLISHING COMPANY

DORDRECHT-HOLLAND / BOSTON-U.S.A.

Library of
Library of
Library of
Library of
Library of
Library of
Library of
Library of
Library of Congress Catalog Card Number 74–81940

The set of three volumes: ISBN 90 277 0524 0
Volume I: ISBN 90 277 0544 5
Volume II: ISBN 90 277 0545 3
Volume III: ISBN 90 277 0546 1

Published by D. Reidel Publishing Company,
P.O. Box 17, Dordrecht, Holland

Sold and distributed in the U.S.A., Canada and Mexico
by D. Reidel Publishing Company, Inc.
306 Dartmouth Street, Boston,
Mass. 02116, U.S.A.

Printed in The Netherlands by D. Reidel, Dordrecht

CONTENTS

CONTENTS OF VOLUMES II AND III

VOLUME II

PART II: ECONOMICS OF INFORMATION AND ORGANIZATION

VOLUME III

PART III: MONEY AND OTHER ASSETS

PART IV: ECONOMIC MEASUREMENTS

PART V: CONTRIBUTIONS TO THE LOGIC OF ECONOMICS