COWLES FOUNDATION DISCUSSION PAPER NO. 22

Experimental Tests of Stochastic Decision Theory*

Donald Davidson and Jacob Marschak**

January 10, 1957

* Research undertaken by the Applied Mathematics and Statistics Laboratory, Stanford University, under contract Nonr 225(17), NR 171-034 and by the Cowles Commission for Research in Economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research. The paper was presented at the Symposium on Measurement co-sponsored by the Section L of the American Association for the Advancement of Science and the Philosophy of Science Association, December 29, 1956.

** The authors were helped by discussions with G. Debreu, E. Fels, L. Hurwicz, R. Radner, H. Raiffa, R. Savage and P. Suppes.

Contents

I. PRIMITIVE AND DEFINED NOTIONS.

1. A set X of the states of the world. Its subsets are called <u>events</u>, denoted by $E, F, ..$ and forming a set $\mathcal{E}$.

2. A set A of <u>alternatives</u>. A includes sure outcomes as well as wagers. If a and b are in A, and E is an event, then aEb denotes the following <u>wager</u>: the subject will receive a if E occurs; he will receive b if E does not occur. In particular, $aEa=a$; and $aXb=a$.

3. The <u>probability</u> $P(a,b)$ that the subject, forced to choose between a and b, chooses a. Therefore $0 \leq P(a,b) = 1-P(b,a) \leq 1$.

4. We define (assume) $P(a,a) = 1/2$.

5. <u>Definitions</u>. If $P(a,b) = 1$ we say a is <u>absolutely preferred</u> to b. If $1/2 < P(a,b) < 1$ we say a is <u>stochastically preferred</u> to b. If $P(a,b) = 1/2$ we say a and b are <u>stochastically indifferent</u>.

<u>Some cases of absolute preference</u>:

(i) if a and b are quantities of a good (e.g., money) and $a > b$ then $P(a,b) = 1$. (This may be, in fact, regarded as the definition of a good). More generally:

(ii) if (a_1,a_2) and (b_1,b_2) are each a pair of quantities of two goods and $a_1 > b_1$, $a_2 \geq b_2$, then (a_1,a_2) is absolutely preferred to (b_1,b_2). The latter pair is called <u>inadmissible</u>.

(iii) if a_1,a_2,b_1,b_2 are any elements of A and $P(a_1,b_1) = 1 = P(a_2,b_2)$ then, for any event E, a_1Ea_2 is absolutely preferred to b_1Eb_2. The latter wager is called <u>inadmissible.</u>

(iv) if a,b are any elements of A such that a is absolutely preferred to b, and if $E \supset F$ (i.e., event E happens whenever F happens but not conversely), then aEb is absolutely preferred to aFb.

6. *Definition*. If $P(a,c) = P(c,b)$ we say that c is a *stochastic midpoint* between a and b.

7. *Definition*. If there exists a stochastic mid point between any a and b in A, we say A is *stochastically continuous*. (This definition suggested by G. Debreu is tentative; it may have to be replaced by a stronger one to make the important conjecture 4 in Sec. II valid).

II. GENERAL STOCHASTIC THEORY OF CHOICE.

1. *Definition*. For a given subject, a real valued function u on A is called his *utility function* if, for any a,b,c,d in A,

$$u(a) - u(b) \gtreqless u(c) - u(d) \quad \text{according as}$$

$$P(a,b) \gtreqless P(c,d) \text{ provided } P(a,b),\ P(c,d) \text{ are not 0 or 1.}$$

(The proviso will be understood in all following propositions; and precautions must be taken to avoid cases of absolute preference in experiments).

This definition of utility corresponds to one of the several (partly independent) definitions of sensation used in psychophysics since Fechner: viz., the definition implied in the much used adage: "~~equal~~ equally often noticed differences (in sensation) are noticed equally often unless they are noticed always or never" (ascribed by Guilford to Fullerton and Cattell).

2. *Definition*. We say the the *quadruple condition* is satisfied if, for any a,b,c,d in A,

$$P(a,b) \geq P(c,d) \text{ if and only if } P(a,c) \geq P(b,d).$$

Note that this condition is, in principle, testable from observations. From the two definitions 1 and 2 follows easily the

3. *General quadruple theorem*: *The quadruple condition is necessary for the existence of a utility function.*

4. It can be proved that the quadruple condition is equivalent to the following: If $a_1, a_2, a_3, b_1, b_2, b_3$ are in A and $P(a_1,a_2) \geqq P(b_1,b_2)$ and $P(a_2,a_3) \geqq P(b_2,b_3)$ then $P(a_1,a_3) \geqq P(b_2,b_3)$. In this form, the condition should help to prove the following conjecture:

5. Special quadruple theorem. If A is stochastically continuous then the quadruple condition is not only necessary (as in 3.) but also sufficient for the existence of a utility function.

6. Definitions. We say that weak stochastic transitivity is satisfied if, for any a,b,c in A,

6.1 $P(a,b) \geqq 1/2$ and $P(b,c) \geqq 1/2$ imply $P(a,c) \geqq 1/2$.

We say that strong stochastic transitivity is satisfied if, for any a,b,c in A,

6.2 $P(a,b) \geqq 1/2$ and $P(b,c) \geqq 1/2$ imply $P(a,c) \geq \text{Max}\,[P(a,b), P(b,c)]$.

(We shall sometimes omit the word stochastic). The terms are due to S. Vail. Clearly 6.2 implies 6.1 .

7. It can be proved that 6.2 is equivalent to the condition

$P(a,b) \geqq 1/2$ if and only if $P(a,c) \geqq P(b,c)$.

8. In a more symmetrical notation, writing $P(a,b) = p_1$, $P(b,c) = p_2$, $P(c,a) = p_3$ (note the reversal in p_3 !):

8.1: Weak transitivity: p_1, p_2, p_3 not all $> 1/2$ or $< 1/2$.

8.2: Strong transitivity:

$p_1 \geqq 1/2$ if and only if $p_2 + p_3 \leqq 1$,

$p_2 \geqq 1/2$ if and only if $p_3 + p_1 \leqq 1$,

$p_3 \geqq 1/2$ if and only if $p_1 + p_2 \leqq 1$.

9. <u>Theorem on stochastic transitivity</u>: <u>Strong</u> (and therefore also <u>weak</u>) <u>stochastic transitivity is a necessary condition for the existence of a utility function</u>. This follows from 3, putting c=d and remembering that P(c,c) = 1/2.

10. If the quadruple condition is satisfied then (whether A be continuous or not) the sizes of intervals between utilities can be completely ordered if the probabilities of choices are estimated from observations. If strong transitivity is satisfied, then probabilities help to compare the sizes ~~can be compared~~ for any two intervals with a common endpoint subtrahend or subtractor. If weak transitivity is satisfied then the utilities can be completely ordered. For, in fact

11. $u(a) \geq u(b)$ if and only if $P(a,b) \geq 1/2$.

The following two theorems are also useful:

12. If $u(a) \geq u(b)$ and $P(a,c) = 1/2$ then $u(c) \geq u(b)$,

a substitution theorem.

13. If c is a stochastic midpoint between a and b then

$$u(c) = 1/2 \cdot u(a) + 1/2 \cdot u(b).$$

14. <u>Testing the theory</u>. If, to avoid the occurence of absolute preferences, all alternatives offered are money-wagers (rather than money-amounts), the testing of the quadruple condition, or of transitivity conditions, may lead to the <u>rejection</u> of the existence of a utility function on the set of money-wagers, since those conditions are necessary. Moreover, if one assumes that the set of money-wagers is stochastically continuous, the test of the quadruple condition may also lead to the <u>acceptance</u> of the existence of a utility function on money-wagers, since, under continuity, that condition is also conjectured to be sufficient.

On Table 1, the upper triple of cards contains three pairs of identical columns. Each column represents a wager; each wager is used as an alternative (a or b or c). The nonsense-syllables marking the rows are the events (E,F,...) determining the outcome of the wager: these syllables are marked on the faces of a die.

III. STOCHASTIC THEORY OF CHOICE BETWEEN SUBJECTIVELY EVEN-CHANCE WAGERS

The following postulate is taken from F.P. Ramsey's work, replacing his notion of indifference (usual in economic theory) by the notion of stochastic indifference.

1. Postulate. If, for a pair a,b in A with $P(a,b) \neq 1/2$, the event E has the property $P(aEb, bEa) = 1/2$ then it has the property $P(cEd, dEc) = 1/2$ for every pair c,d in A. We say that such an event has a (subjectively) equal chance and call E an "even-chance event". The class of all even-chance events is a subset of $\mathcal{E}$ and will be denoted by $\mathcal{E}^*$. If E is in $\mathcal{E}^*$ any wager aEb is called an even-chance wager (we shall omit the word subjective except when a reminder is called for).

2. Postulate. If E and F are both in $\mathcal{E}^*$ then, for any a,b in A,

$$P(aEb, aFb) = 1/2 .$$

The first of the two postulates says, in effect, that an event is or is not an even-chance event, regardless of the particular wagers used to identify it as such (Ramsey's "ethical neutrality" of beliefs). The second postulate says two even-chance wagers with the same pair of outcomes are stochastically indifferent. Combining the postulates with the definition of, and theorems on, utility in Sec. II, we obtain

3. Theorem. If u exists and the above postulates are true then, for any E, F in $\mathcal{E}^*$ and any a,b,c in A,

$$u(aEb) = u(bEa) = u(aFb) = u(bFa);$$

Because of these substitution relations we can write more briefly

$$aEb \equiv ab \equiv ba, \quad \text{whenever } E \text{ is an even-chance event.}$$

4. Definition. We say that the subject is unbiased if, for every wager ab that has (subjectively) even chance, and for every c with $P(ab,c) = 1/2$, $(P(a,c) = P(c,b))$,

~~$P(a,ab) = P(ab,a)$,~~

i.e., if ab is ~~a~~ indifferent to every stochastic midpoint between a and b. From this definition, and that of the utility function in Sec. II, we obtain

5. Theorem. If a subject is unbiased and has a utility function over the set A, then

5.1 For any a,b in A, $u(ab) = 1/2 \cdot u(a) + 1/2 \cdot u(b)$

(Compare 13 in Sec. II).

5.2 For any $a_1,a_2,a_3,a_4,b_1,b_2,b_3,b_4$ in A

$P(a_1a_2, b_1b_2) \geqq P(b_3b_4, a_3a_4)$ if and only if

$\sum_1^4 u(a_i) \geqq \sum_1^4 u(b_i)$.

6. Definition. We say that the octuple condition is satisfied if for every $a_1,a_2,a_3,a_4,b_1,b_2,b_3,b_4$ in A, and for all integers i,j,k,ℓ running from 1 through 4,

either $P(a_ia_j, b_ib_j) \geqq P(b_ib_j, a_ka_\ell)$

or $P(a_ia_j, b_ib_j) \leqq P(b_ib_j, a_ka_\ell)$.

Altogether, 9 pairs of distinct probabilities are thus related by concordant inequalities, provided the subscripts are all kept distinct. Note that, like the quadruple condition (and its transitivity implications), the octuple condition is, in principle, testable from observations.

7. General octuple theorem. The subject has a utility function and is unbiased only if the octuple condition is satisfied. Proof: use 5. It is open to conjecture whether (analogous to the special quadruple theorem of Section II) the octuple condition is not only necessary but also sufficient for the subject to have a utility function and to be unbiased, provided the set of alternatives is stochastically continuous.

8. From the octuple condition or from 5.2 one derives a weaker condition which involves 6 instead of 8 outcomes: For any a,b,c,d,e,f,g in A,

$$P(bf,de) \geqq 1/2 \text{ if and only if } P(ab,cd) \geqq P(ae,cf).$$

We call this condition the strong transitivity condition for utility intervals because of the following analogy with Section II: if u exists and the subject is unbiased, the left side of the condition implies

$$I \equiv u(b) - u(d) \geqq u(e) - u(f) \equiv J ,$$

using I,J to denote utility-differences (interval-sizes). Write $K \equiv u(c) - u(a)$; then the whole condition becomes the identical statement that

$$I \geq J \text{ if and only if } I - K \geq J - K .$$

This is analogous to the strong transitivity condition of Section II which, if u exists, is stated identically as:

$$u(a) \geqq u(b) \text{ if and only if } u(a) - u(c) \geqq u(b) - u(c) .$$

9. Similarly, the weak transitivity condition for intervals is defined as follows:

$$\text{if } P(bf,de) \geqq 1/2 \text{ and } P(ae,cf) \geqq 1/2 \text{ then } P(ab,cd) \geqq 1/2 ;$$

if u exists and the subject is unbiased, this condition becomes the following identity (with the interval-sizes I,J,K defined as in the preceding paragraph):

$$\text{If } I \geq J \text{ and } J \geq K \text{ then } I \geq K, -$$

just like 6.1 of Section II becomes the identity:

$$\text{If } u(a) \geq u(b) \text{ and } u(b) \geq u(c) \text{ then } u(a) \geq u(c).$$

9a. The content of 8. and 9. is illustrated in Table 1 by the specimen cards in the second line. (The interchange of rows or of columns in each card does not matter) If the subject obeys theorem 2.1 then, whenever he chooses a wager, i.e., a column of the card, he compares the sizes of two intervals, represented by the rows of the card. On the three cards considered, there are three pairs of identical rows: they correspond to the three intervals I, J, K just discussed.

10. Henceforth we shall refer to the transitivity conditions of Section II as transitivity conditions for alternatives. With p_1, p_2, p_3 appropriately defined, the symmetrical formulations in 8., Section II can be carried over from "alternatives" to "intervals". These definitions are, as suggested in the preceding paragraphs: (and note the reversal in p_3 !)

$$p_1 \equiv P(bf,de);\ p_2 \equiv P(ae,cf);\ p_3 \equiv P(cd,ab).$$

11. Money wagers. If the outcomes are money gains or losses the conditions (i),(iii) of 5., Section I, will apply: the subject will have absolute preference for wager ab over wager $a'b'$ if $a > a'$ and $b \geq b'$. This is avoided if the six money amounts that occur in the six wagers just listed (see 10.) obey the following inequalities:

$$b < d < e < f < c < a.$$

12. Testing the theory. One has first to test the hypothesis (irrelevant for the theory of the preceding Section but important for the present one) that

all chance events underlying the designed money-wagers have subjectively even chance, for each subject. Only subjects passing this test can be further used. Having satisfied ourselves on this account, we can test whether the octuple condition and the transitivity conditions are satisfied. If they are not, the subject is not unbiased or has no utility function on the set of money-wagers; and none, therefore, on the set of money-amounts. If they are satisfied, the subject may or may not have a utility function on money-amounts. Should the mathematical conjecture prove true that, under a continuity assumption, the octuple condition is sufficient (and not only necessary) for the unbiasedness and the existence of a utility function, its testing might lead to the acceptance of the utility function on money-amounts.

IV. OUTLINE OF EXPERIMENT.

1. Subjects: 17 students of Stanford University.

2. Alternatives offered: money-wagers. Each specimen card on Table 1 shows two wagers, one in each of the two columns of the card. The chance event was: that a cube will come up with one of the three faces marked by a certain nonsense-syllable.

3. The subject's response consisted in stating which of the two wagers he accepted; and in some of the trials, selected at random, he had actually to play and win or lose money.

4. For the purposes of Section III it was important to ascertain whether the chance-event in question was actually given an even chance by the subject. However, the hypothesis that $P(aEb,bEa) = 1/2$ was not tested. Instead, a subsidiary postulate was made:

If m, m', m'', n are money amounts and $m - m' = m'' - m = h$, and $P(nEm, mEn) = 1/2$, then and only then, for any postive h, however small,

$$P(nEm, m'En) > 1/2 \quad \text{and} \quad P(nEm, m''En) < 1/2 .$$

In previous experiments of D. Davidson, S. Siegel and P. Suppes, h was equal one cent; and the same cubes were used. A large majority of the then tested Stanford students accepted the wager nEm over $m'En$; and rejected the wager nEm in favor of $m''En$, for various values of the money-amounts. This strengthened our a priori belief that the condition of the subsidiary postulate was likely to be satisfied; therefore, we tested every cube only a few times, with each subject, to satisfy ourselves that the subjective chances were even. All 17 subjects passed this test and were therefore admitted to further experiments.

5. To study the possible effects of learning, three sessions were held with each subject; the first session was in part occupied by the pre-test just described,

6. The numbers of triples of cards (each containing a pair of wagers) offered in each session is shown in Table 2 . It was possible to devise the cards in such a way that the same card was sometimes used in a triple designed to test the "Transitivity of Alternatives" (Sec.I) and in another triple, serving to test the "Transitivity of Intervals" (Sec.III). Because of this overlap, the total number of cards presented to each subject was not 2x3x76 but much smaller.

7. Absolute preferences were, we hope, avoided by:

using wagers (and never inadmissible ones);

never using a choice between a wager and a sure thing, because of the suspicion raised by workers in the field, that certainty as such may be loved (or hated).

always chosing the money amounts so that the actuarial values of the two wagers on one card never differed by more than 4 1/2¢.

8. To avoid the effect of memory, the wager ab was given sometimes as aEb and sometimes as bEa; and was entered sometimes in the left, and sometimes in the right column. But <u>the same "pair of wagers" was never repeated</u>. This distinguished the experiment from those performed by Papandreou et al. on commodity bundles (not wagers), and by Ward Edwards (com. wagers). More experience is needed before one decides just how dangerous is the repetition of the same pair of alternatives.

V. STATISTICAL DECISION RULES

1. No statistical design has been developed to test the <u>quadruple</u> and <u>octuple</u> conditions. A test of the quadruple condition would involve four ~~four~~ probabilities of choices, and hence sixtuples of wager-pairs (cards), instead of the triples actually used so far.

2. <u>Transitivity regions</u>. Denote by $p^i = (p^i{}_1, p^i{}_2, p^i{}_3)$ the ith triple of the probabilities defined in Section II, 8; or, alternatively, in Section III, 10. p^i is a point of the unit cube U, since each component of p is between 0 and 1.

We now define two sub-regions of the unit cube U:

Region W: p^i obeys the conditions in 8.1, Sec. II (weak transitivity)

Region S: p^i obeys the conditions in 8.2, Sec. II (strong transitivity)

Hypotheses-pairs to be tested:

Hypothesis H_W: For all i, p^i is in W
" H^o_W: There exists an i such that p^i is in U-W;

Hypothesis H_S: For all i, p^i is in S
" H^o_S: There exists an i such that p^i is in U-S.

3. Only one observation was made for each triple of wager-pairs, to avoid the effect of memory. Statistically, the problem is analogous to the following: "Test the hypothesis that each coin being made by the Denver mint has a bias (not necessarily an equally strong one for all coins) in favor of falling heads; you are permitted to take a finite sample of coins and to toss each coin just once.".

4. Statistical reformulation of hypotheses:

$$\begin{cases} H_W: & p^i \text{ is distributed uniformly over } W, \text{ and } \mathrm{Prob}(p^i \text{ in } W) = 1 \\ H_o: & p^i \text{ is distributed uniformly over } U, \text{ and } \mathrm{Prob}(p^i \text{ in } U) = 1 \end{cases}$$

$$\begin{cases} H_S: & p^i \text{ is distributed uniformly over } S, \text{ and } \mathrm{Prob}(p^i \text{ in } S) = 1 \\ H_o: & p^i \text{ is distributed uniformly over } U, \text{ and } \mathrm{Prob}(p^i \text{ in } U) = 1 \end{cases}$$

5. Nature of observations. Subject chooses

a^i over b^i: yes, no

b^i over c^i: yes, no

c^i over a^i: yes, no

A cyclical observation: "yes, yes, yes" or "no, no, no" for a given triple.

	Probability of cyclical observation
If H_o is true	$\frac{20}{80} = 25.00\%$
If H_W is true	$\frac{15}{80} = 18.75\%$
If H_S is true	$\frac{11}{80} = 13.75\%$

6. Decision rule: Accept H_W if number of ~~intransitive~~ cyclical responses $= r < c$; where c is such that $\mathrm{Prob}(r < c|H_o) = \mathrm{Prob}(r \geq c|H_W)$ (= significance level), thus making the largest of the two error probabilities a minimum. Similarly for H_S.

DECISION RULES	Number of observations n = 76 (All Sessions)	26 (Session III)
Accept H_w if r is less than:	17	6
Significance level:	24 %	35 %
Accept H_s if r is less than:	15	5
Significance level:	10 %	23 %

7. Moreover: If the proportion of cyclical observations falls very much below the probability indicated in 5. for weak or strong transitivity, we shall presume that the assumption of uniform distribution over the (weak or strong) transitivity region is to be corrected; we shall have to assign lower weights to those points of the region that lie near its boundaries other than the facets of the unit cube.

VI. RESULTS OF EXPERIMENT

1. Table II shows, from session to session, considerable leaning towards smaller frequency of cyclical responses. In the overall result, but more particularly in the last session, the number of cyclical responses falls, for a large majority of subjects, far below the expected frequency under strong transitivity.

2. Table III applies the Decision rules of 6, Section V, to the total of all sessions. For all subjects, weak transitivity had to be accepted (though at the very modest, i.e., high, significance level), and this with respect to alternatives as well as to intervals. For

all but two subjects, strong transitivity had to be accepted (at a significance level of 10%); but the two subjects were not the same with respect to alternatives and to intervals. However, the correlation of .49 between the subjects' behavior with respect to alternatives and with respect to intervals is just significant (for 17 observation-pairs) at the 5% level (Tables of Fisher and Yates).

3. The low frequency of cyclical responses, especially after a learning period, lets us believe that for most subjects the points p lie well within the transitivity regions (see 7.) and that therefore the transitivity conditions necessary for the existence of a utility function over the set of money-wagers as well as over the set of money-amounts are satisfied.

Table 1

Specimen Cards

To test Transitivity (strong, weak, none) of Alternatives:

	1	2
ZOJ	+6¢	-5¢
ZEG	-38¢	-21¢

	1	2
QUG	+6¢	-54¢
QEJ	-38¢	+22¢

	1	2
ZUH	-54¢	-5¢
ZEG	+22¢	-21¢

To test Transitivity (strong, weak, none) of Utility-Intervals:

	1	2
ZOJ	+24¢	+13¢
ZEG	-6¢	+5¢

	1	2
QUG	+38¢	+31¢
QEJ	+13¢	+24¢

	1	2
ZUH	+31¢	+38¢
ZEG	+5¢	-6¢

Table 2 - Results, Single Subjects

	Testing Transitivity of Alternatives				Testing Transitivity of Intervals			
Sessions:	I	II	III	Total	I	II	III	Total
ber of triples offered	22	28	26	76	22	28	26	76
				NUMBER OF CYCLICAL RESPONSES				
ected number under uni-rm distribution over the gion of transitivity	5.50	7.00	6.50	19.00	5.50	7.00	6.50	19.00
ak "	4.13	5.25	4.88	14.25	4.13	5.25	4.88	14.25
ong "	3.03	3.85	3.58	10.45	3.03	3.85	3.58	10.45
Subject:								
A	4	0	0	4	0	0	1	1
B	3	5	2	10	1	6	3	10
C	5	3	3	11	6	5	5	16
D	4	7	0	11	4	5	2	11
E	1	0	0	1	0	3	3	6
F	3	6	0	9	5	10	2	17
G	2	1	2	5	3	0	0	3
H	2	1	1	4	4	2	1	7
I	1	2	1	4	3	3	0	6
J	2	9	5	16	3	3	6	12
K	4	2	2	8	3	5	3	11
L	1	1	0	2	1	1	0	2
M	2	2	1	5	7	5	1	13
N	6	3	7	16	2	0	3	5
O	4	2	1	7	3	2	0	5
P	1	3	3	7	2	1	3	6
Q	7	5	2	14	1	4	2	7
Average:	3.06	3.06	1.77	7.88	2.20	2.80	2.60	8.12

TABLE 3

All sessions (76 triples ; distribution of subjects by the number of cyclical responses. (Correlation coefficient = .49)

No. of cyclical responses on intervals

No. of cyclical responses on alternatives	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	No. of Subjects	
1						E												1	Accept Strong and Weak Transitivity
2		L																1	
3																		0	
4	A					I	H											3	
5			G										M					2	
6																		0	
7					O	P												2	
8											K							1	
9																	F	1	
10										B								1	
11											D					C		2	
12																		0	
13																		0	
14							Q											1	
15																		0	Accept Weak Trans-
16					N							J						2	
17																		0	
No. of Subjects	1	1	1	0	2	3	2	0	0	1	2	1	1	0	0	1	1	17	
	Accept Strong and Weak Transitivity														Accept Weak Transitivity Only				