# A NON-LINEAR MODEL FOR THE EXPERIMENTAL MEASUREMENT OF UTILITY

*by Patrick Suppes and Karol Walsh*
*Stanford University*

Reprinted from BEHAVIORAL SCIENCE
Vol. 4, No. 3, July, 1959
*Printed in U.S.A.*

Reprinted from BEHAVIORAL SCIENCE
Vol. 4, No. 3, July, 1959
*Printed in U.S.A.*

**When a person has to choose between two or more situations, each involving certain risks, he presumably weighs the chances of the possible outcomes against the "worth" of these outcomes to him. His estimation of the chances is governed by his "subjective probabilities," his relative preferences for the outcomes by his "utility function." This paper explores ways of estimating the latter so as to predict individuals' choices among various bets, assuming the chances are known.**

# A NON-LINEAR MODEL FOR THE EXPERIMENTAL MEASUREMENT OF UTILITY*

*by Patrick Suppes and Karol Walsh*

*Stanford University*

## THE MODEL

THE experiment described in this paper was constructed to measure the utility, to an individual, of small amounts of money[1], and to predict on the basis of the measured utilities of these offered amounts the individual's future choices with respect to similar offers of money. The experimental design is based upon previous work done by Davidson, Suppes and Siegel (1957), but no familiarity with that work is presupposed.

The model described here is designed to eliminate a specific criticism raised against the linear-programming model, discussed in Chapter III of Davidson, et al (1957). The criticism, which will be explained in detail after the basic structure of the present model has been outlined, is that the linear-programming model leads to a violation of the sure-thing principle, that is, it permits the choice of an action which, whatever the state of nature, is no better than, and for at least one state of nature is not as good as, a second action which is available. Put another way, the purpose of the present paper is to offer a non-linear model for measuring utility which has the following desirable property: it is compatible with imperfect discrimination but at the same time incompatible with any violation of the sure-thing principle.

Because this seemed the best approach, the structure of the model is presented below in terms of the basic experimental situation (to be described in more detail in the next section). Seven values (small amounts of money) $a$, $b$, $c$, $d$, $e$, $f$, and $g$ are chosen, in ascending order, and offers are given the subject in the form of a game matrix:

| *A* | *B* |
|---|---|
| $x$ | $z$ |
| $y$ | $w$ |

where $x$, $y$, $z$ and $w$ range over the selected values with the restriction that $x > z > w > y$. From seven original values, 35 offers of the above form can be generated. (The general combinatorial formula is $\binom{n}{4}$ where $n$ is the number of original values chosen). The subject is instructed to choose a column, *A* or *B*, which is called an option, and then the row value is decided by the throw of a fair die.[2] It is assumed that the subject will choose that option of the two whose ex-

* Herman Rubin contributed substantially to the formulation of the model basic to this paper. We are also indebted to Jacob Marschak for several constructive criticisms of an earlier draft. The experiment reported here was supported under Contract NR 171-034, Group Psychology Branch, Office of Naval Research.

[1] The model also applies when outcomes are not monetary.

[2] By a procedure to be described later, the subject is assured that the probability of obtaining either row value is ½.

pected value is larger. Thus from a subject's choice, say $A$, an inequality of the form

$$\text{(1)} \qquad \tfrac{1}{2}u(x) + \tfrac{1}{2}u(y) \geq \tfrac{1}{2}u(z) + \tfrac{1}{2}u(w)$$

results, where $u$ is the subject's utility function. Inequality (1) says, simply, that the expected value of option $A$ is greater than or equal to the expected value of option $B$. Thirty-five such inequalities are obtained, one for each choice of an offer. The $\tfrac{1}{2}$'s cancel out, and for further work the inequalities are converted into equivalent interval inequalities of the form

$$\text{(2)} \qquad u(x) - u(z) \geq u(w) - u(y).$$

The quantity $u(x) - u(z)$ represents a "utility interval," the difference to the subject between the utility of the offer $x$ and the next closest value in the inequality, $z$.

From the seven strictly ordered values $a$, $b$, $c$, $d$, $e$ $f$ and $g$, there result six atomic intervals, $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$ and $\kappa$, as described in the figure below:

| | | | | | |
|---|---|---|---|---|---|
| $\alpha$ | $\beta$ | $\gamma$ | $\delta$ | $\epsilon$ | $\kappa$ |

$a \quad b \quad c \quad d \quad e \quad f \quad g$

Thus the inequality

$$\text{(3)} \qquad u(c) + u(d) \geq u(a) + u(e)$$

may be expressed as

$$\text{(4)} \qquad u(c) - u(a) \geq u(e) - u(d)$$

and finally as

$$\text{(5)} \qquad \alpha + \beta \geq \delta.$$

In general the 35 inequalities resulting from the subject's responses do not have a solution. A solution may be obtained if the inequalities are modified as follows: the larger side of each inequality is multiplied by the minimum positive quantity $\eta$, $\eta \geq 1$, such that a solution of the inequalities exists. Thus all inequalities like (5) are written as

$$\text{(6)} \qquad \eta\alpha + \eta\beta \geq \delta.$$

The existence of a unique minimum $\eta$ is easily proved; in the experiment discussed in this paper for each subject an $\eta$ was used which was within .1 of the minimum. In finding a solution for the atomic intervals, the normalizing restriction was made that at least one interval be equal to one. Unfortunately, as is usually the case with inequalities, the solution is not unique; here the set of solutions is a convex polyhedron. Due to limited computational facilities, the first solution found had to suffice, with the exception of one subject.

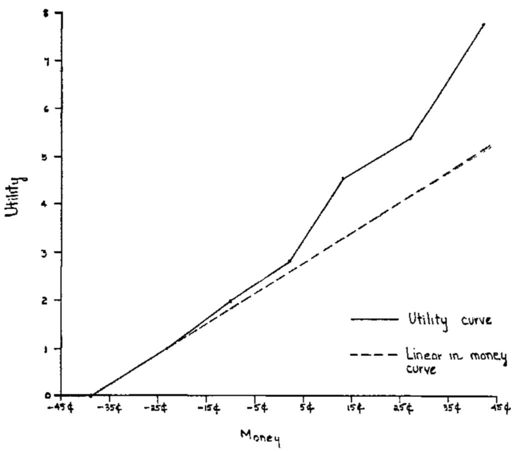



FIG. 1. Utility Curve for Subject #4.

Using the interval values obtained, a utility curve is constructed for each subject, with monetary values plotted on the abscissa and utility values on the ordinate. Figure 1, which is the graph for Subject No. 4, is typical.[3] Using these graphs, predictions are made for another set of alternatives, similar to the first in form but containing many other values besides the original seven. In other words, these graphs provide a means of linear interpolation from the utilities of the seven original values to other values given in the "prediction sessions." The solutions obtained for the intervals $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$ and $\kappa$ indicate how far apart the seven given values are in utility (with respect to an arbitrary unit) and from the graph obtained from this information the relative utilities of other values can easily be found.

The basic hypothesis of the model can be stated as follows: Let $x$, $y$, $z$ and $w$ be four amounts of money such that $x > z > w > y$ and let the given options be $A = (x, y)$ and $B = (z, w)$. Then if $\alpha = u(x) - u(z)$ and $\beta = u(w) - u(y)$ are two intervals resulting from the options and $\alpha/\beta > \eta$, the option chosen must be $A = (x, y)$; if $\beta/\alpha > \eta$, the

[3] However, *shapes* of the graphs varied considerably from subject to subject.

option chosen must be $B = (z, w)$. If $\alpha/\beta \leq \eta$ and $\beta/\alpha \leq \eta$, then the choice is said to be probabilistic, i.e., it is predictable only in the sense that it satisfies the following two postulates:

(i) if $u(x) - u(z) \geq u(w) - u(y)$,
then $P(A) \geq P(B)$,

and

(ii) if $(u(x) - u(z))/(u(w) - u(y)) \leq \eta$,
then $P(A) < 1$,

(where $P(A)$ is the probability that option $A$ will be chosen). The postulates seem intuitively satisfactory. The first says, simply, that if the utility of option $A$ is greater than or equal to the utility of option $B$, then the probability that option $A$ will be chosen is greater than or equal to the probability that $B$ will be chosen. Postulate (ii) says that if the utility ratio of option $A$ to option $B$ is less than or equal to $\eta$, then the probability that option $A$ will be chosen is less than one. It is implied that as the ratio gets larger the probability that $A$ will be chosen approaches one.

Before the rationale of making predictions is discussed, two remarks may be made about the non-linear model just outlined. First, an important property of this model is that $\eta$ is invariant under a linear transformation of the utility function. For this reason it is meaningful to compare the $\eta$ values of different individuals. If, for example, $\eta_1 < \eta_2$ it makes sense to say that individual 1 is more discriminating than individual 2. In fact, not only more discriminating, but more rational; for it would not be difficult to construct a two-person game in which 1 would have a clear advantage over 2 because of his finer discriminating ability. (A similar kind of example is discussed in detail in von Neumann and Morgenstern [1947, pp. 614–616].) For obvious reasons we call $\eta$ a *discrimination parameter*. It should be clear that discrimination is perfect if and only if $\eta = 1$.

Second, the non-linear model meets and eliminates the criticism raised against the linear-programming model of violating the "sure-thing" principle. An example may be used to illustrate the problem and show how the non-linear model succeeds where the linear model does not. The essential difference between the linear and non-linear models is that in order to obtain a solution for a set of inequalities the first model *adds* the minimum constant $\theta$, $\theta \geq 0$, to the larger side of each inequality to make a solution possible[4] instead of multiplying by the minimum positive constant $\eta$. Suppose, then, that the linear model is applied, and the following results are obtained (where $u$ is the subject's utility function):

$$u(50¢) = 49$$
$$u(45¢) = 46$$
$$u(-45¢) = -47$$
$$u(-50¢) = -51$$
$$\theta = 8$$

Then since

(7) $u(45¢) + u(-50¢) + 8 > u(50¢) + u(-45¢)$

the linear-programming model predicts that with probability greater than zero, the subject will choose option $B$ when the options are:

| A | B |
|---|---|
| 50¢ | 45¢ |
| −45¢ | −50¢ |

where the subjective probability of each row is 1/2. Such behavior seems extremely unlikely, since a person with any intelligence would immediately choose more money (option $A$) rather than less.

But suppose the non-linear model is applied to the above example. Then if the same absurd result were to hold, it would follow that, for some $\eta$, $\eta \geq 1$

(8) $\eta(u(45¢) - u(50)) \geq u(-45¢) - u(-50¢)$,

[4] To find a minimum $\theta$ is to minimize a linear expression, $\theta$, subject to linear inequalities. This is called a linear programming problem; hence our designation of the model. To minimize $\eta$, on the other hand, is to minimize a linear expression subject to non-linear inequalities.

or

(9) $\eta(46 - 49) \geq -47 - (-51)$

or

(10) $-3\eta \geq 4,$

which obviously cannot be true for any positive $\eta$. It is easily seen that the same result would obtain whenever $A = (x, y)$, $B = (z, w)$, and $x > z$, $y > w$. Thus under the non-linear model, it can be concluded that the choice of an option like $B$ above can occur only with probability zero, which result satisfies intuitive demands.

The reason that the $\theta$ solution was applicable to the linear-programming experiment of Davidson, et al. (1957) and not to the present one is to be found in the fact that in the former experiment the items offered as outcomes were phonograph records, not amounts of money, and although some might feel the necessity for it, there is really no objective principle, similar to the one that a rational person must prefer a sure gain of more money to less money, which says that an option consisting of music by Bach and Corelli should be preferred to one consisting of works by Bartok and Stravinsky. However, it might reasonably be argued against the applicability of the $\theta$-model, even to choices among phonograph records, that if a subject strictly orders six records in preference $a$, $b$, $c$, $d$, $e$ and $f$ ($f$ the highest) and if options $(e, f)$ and $(c, d)$ are offered him, then he should be judged just as irrational in choosing $(c, d)$ as in choosing the corresponding (45¢, −50¢) as opposed to (50¢, −45¢).

As indicated by the model, three kinds of predictions are made: (1) using $\eta$, (2) without $\eta$ and (3) stochastic, or probabilistic, predictions. For illustration, again let $x$, $y$, $z$ and $w$ be amounts of money such that $x > z > w > y$ and let option $A = (x, y)$ and option $B = (z, w)$. Predictions of the first kind are performed as follows: The ratio $(u(x) - u(z))/(u(w) - u(y))$ is found, call it $\alpha/\beta$. Then the three mutually exclusive and exhaustive possibilities are:

(a) if $\alpha/\beta > \eta$, option $A$ is predicted,
(b) if $\beta/\alpha > \eta$, option $B$ is predicted,
(c) if $\alpha/\beta \leq \eta$, and $\beta/\alpha \leq \eta$, then the prediction is probabilistic.

Predictions of the second kind fall into the following three categories, also mutually exclusive and exhaustive:

(a) if $\alpha/\beta > 1$, option $A$ is predicted,
(b) if $\beta/\alpha > 1$, option $B$ is predicted,
(c) if $\alpha/\beta = 1$, then the choice is indeterminate.

Predictions of the third kind are made for those cases in which the ratios $\alpha/\beta$ or $\beta/\alpha$ (whichever is larger) are less than or equal to $\eta$. Each $\eta$ is partitioned into three intervals:[5]

$[1, 1 + (1/3)(\eta - 1)]$
$[1 + (1/3)(\eta - 1), \quad 1 + (2/3)(\eta - 1)]$
$[1 + (2/3)(\eta - 1), \eta]$

and the relative frequency of correct predictions to the total number of predictions in each interval is computed. (The choice predicted in each case is the option having the greater utility.) It is expected, as a result of the probabilistic postulates, that each successive interval will contain a larger number of correct predictions than the preceding one. In other words, as the ratio of $\alpha$ to $\beta$ becomes larger (but not larger than $\eta$) it is assumed that the option with the greater utility will have a greater probability of being chosen.

## EXPERIMENTAL PROCEDURE[6]

The subjects were eight randomly chosen sailors from Moffett Field Air Base. At the beginning of the first session the subject was told that he was being asked to participate in a "gambling" experiment whose purpose was to give some insight into the manner in which people make choices involving various offers of money. The subject was given $2.00 credit to assure him that he would not *lose* any money in the experiment. He was told that he would be allowed to keep whatever amount of money he won during the six sessions of the experiment and that in previous runs the average profit to a subject

[5] The choice of three intervals rather than some other number was determined by the number of observations available.

[6] For a detailed description of a similar experimental design see Chapter II of Davidson et al., (1957).

for two hours of play (six sessions)[7] was about $2.50.

The subject was assured that the experiment was strictly descriptive, i.e., that its purpose was to indicate how people do act under certain conditions, not to measure, according to some norm, a particular subject's actions. It was remarked, also, that the results of the experiment would remain confidential, and each of the subjects was asked not to discuss the experiment with any other person who might also be participating in it, since each person was to be tested separately, and independent results were desired.

After these general remarks had been made, the subject was given a trial set of alternatives similar to the following example:

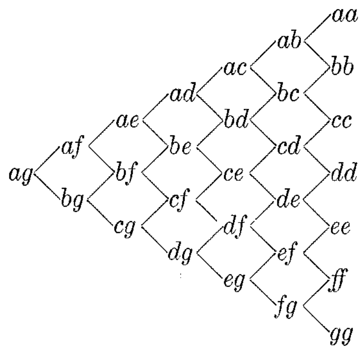
| | *A* | *B* |
|---|---|---|
| ZOJ | 6¢ | −5¢ |
| ZEJ | −4¢ | 6¢ |

He was told to pick column *A* or column *B*. Then a die with nonsense syllables ZOJ or ZEJ, each appearing on three sides, was tossed by the subject from a shaker box to determine the actual outcome. This particular die was systematically alternated with two others having different nonsense syllables on their faces.

Several trial alternatives similar to the above example were given to make certain that the subject understood the game and that he had a subjective probability of 1/2 for the chance events ZOJ and ZEJ (for the latter hypothesis choice of option *A* was required in the above example). Then the subject was asked if, in order to save time, he would not like to run through the rest of the alternatives (these were the real experimental ones) choosing all his options, *A* or *B*, first and tossing for his wins later. He was told that he would not be able to toss for all the options presented—this would take too long—but that he would be allowed to pick ten numbers at random from an envelope containing as many numbers as there were offers in the session, and that he could roll the die for these ten options. Thus the six sessions were run, the subject choosing his options first and tossing for them later. All eight subjects returned for their six sessions.

In outline, these were the actual mechanics of the experiment. A word now about the values offered and how these were determined. The amounts −39¢, −23¢, −10¢, 2¢, 13¢, 27¢ and 42¢ were the values chosen to be paired in various combinations of offers for the first two sessions. The first two sessions formed the part of the experiment from which the utility curves and discrimination parameters were obtained. The fifth and sixth sessions formed the part whose choices of options were predicted on the basis of information obtained from the first part.

The options of the first two sessions were chosen in the following manner: The letters *a*, *b*, *c*, *d*, *e*, *f* and *g* were selected to represent the values −39¢, −23¢, −10¢, 2¢, 13¢, 27¢ and 42¢ respectively. Then a lattice was constructed as follows on the basis of transitivity and the sure-thing principle:




and each pair of values on the lattice was combined with all other pairs of values which could not be reached by an ascending or descending line from the pair in question. [For example, *be* was taken with *af* and *cd* to form the alternatives (−23¢, 13¢) and (−39¢, 27¢), and (−23¢, 13¢) and (−10¢, 2¢).] All "non-gambles," that is, offers of the form: *ad* with *bb* or *ad* with *cc* were

[7] Actually the results of only four sessions, the first two and the last two, were used to test the above-described model; results of the other two sessions were for a model designed to measure utility of gambling (See Royden, Suppes, & Walsh, 1959).

eliminated. Seven randomly chosen offers were repeated, and all offers, totalling 42, were presented in random order to the subject. The 70 pairs of options in the "prediction" part of the experiment were composed in a manner similar to the first—no option dominated, or was greater than or equal in both values to both values of its companion—but many more values were used.

**RESULTS**

In Table I(a) are recorded the actual number of utility predictions which were correct, within $\eta$, and wrong for each subject. There are only 53 wrong, out of 560, as opposed to an expected 280 if predictions had been made by chance, and there are 128 correct. As might be anticipated, subjects with smaller $\eta$'s (*S*s 1, 2 and 4) had the greater number of right predictions. This outcome agrees with the interpretation of $\eta$ as a measure of discrimination. In Table I(b) are summarized the results of predictions made without $\eta$ but using the utilities read from the subjects' graphs (type 2 predictions, described earlier).

The results of Table I are compared to what are called "actuarial" predictions, recorded in Table II. These predictions are computed in the same way as the utility predictions made with and without $\eta$ [Table II(a) and Table II(b)], using as "utility values" the offered amounts of money themselves instead of the corresponding utility values interpolated from the graphs. For example, for the following offer $A$: (2¢, −10¢) and $B$: (17¢, −23¢), the differences

$$-10 - (23) = 13 \quad \text{and} \quad 17 - 2 = 15$$

instead of

$$u(-10) - u(-23) \quad \text{and} \quad u(17) - u(2)$$

would be computed and the ratio 15/13 compared to $\eta$, if the prediction were to be made using $\eta$, and compared to 1 if the prediction were to be made without $\eta$.

It is immediately noticeable that the use of actual monetary values as utilities fails to provide a sensitive enough scale for effective predictive results. The data of Table II(a) show that almost all predictions fall within $\eta$, (all but 10 in fact), if subject #2 is excluded. Further, a comparison of Table I(b) to Table II(b) shows that the utility model, without using $\eta$, yields 306 correct predictions to the actuarial model's 257, and 222 wrong predictions to the actuarial model's 191. This is a gain of 49 in correct predictions as opposed to a loss of 31 in wrong predictions.

In spite of the sizable discrimination parameters of some of the *S*s ($\eta$ = 1.9 for *S*s #3, 5, 8), only one of the eight (#2) had a utility function which was linear in money. It is slightly puzzling that #2 gave perfect actuarial responses in the first two sessions but not in the final two predictive sessions, as may be seen from Table II.

TABLE I
PREDICTIONS MADE USING SUBJECT'S UTILITY FUNCTION

| Subject | $\eta$ | (a): with $\eta$ | | | (b): without $\eta$ | | |
|---|---|---|---|---|---|---|---|
| | | Correct Predictions | within $\eta$ | Wrong Predictions | Correct Predictions | Equal Intervals | Wrong Predictions |
| 1 | 1.4 | 38 | 30 | 2 | 61 | 1 | 8 |
| 2 | 1.0 | 35 | 14 | 21 | 35 | 14 | 21 |
| 3 | 1.9 | 8 | 57 | 5 | 31 | 3 | 36 |
| 4 | 1.2 | 42 | 9 | 19 | 45 | 0 | 25 |
| 5 | 1.9 | 4 | 65 | 1 | 45 | 1 | 24 |
| 6 | 1.5 | 0 | 67 | 3 | 29 | 4 | 37 |
| 7 | 1.62 | 0 | 70 | 0 | 28 | 8 | 34 |
| 8 | 1.9 | 1 | 67 | 2 | 32 | 1 | 37 |
| Totals | | 128 | 379 | 53 | 306 | 32 | 222 |

TABLE II
PREDICTIONS MADE USING ACTUARIAL VALUES

| Subject | (a): with $\eta$ | | | (b): without $\eta$ | | |
|---|---|---|---|---|---|---|
| | Correct Predictions | within $\eta$ | Wrong Predictions | Correct Predictions | Equal Intervals | Wrong Predictions |
| 1 | 1 | 69 | 0 | 31 | 14 | 25 |
| 2 | 35 | 14 | 21 | 35 | 14 | 21 |
| 3 | 0 | 70 | 0 | 32 | 14 | 24 |
| 4 | 3 | 61 | 6 | 30 | 14 | 26 |
| 5 | 0 | 70 | 0 | 33 | 14 | 23 |
| 6 | 0 | 70 | 0 | 35 | 14 | 21 |
| 7 | 0 | 70 | 0 | 32 | 14 | 24 |
| 8 | 0 | 70 | 0 | 29 | 14 | 27 |
| Totals | 39 | 494 | 27 | 257 | 112 | 191 |

For a more careful comparison of the utility and actuarial models it is natural to consider the prediction data of Tables I(b) and II(b), since the actuarial model made essentially no clear predictions when the discrimination parameter was considered. We want to test the null hypothesis that the predictive worth of the two models is the same when $\eta$'s are ignored. The basic idea of the test we use is to tabulate the predictive disagreements of the models for each subject, then to find for each subject the significance level for rejecting the null hypothesis, and finally to combine these significance levels and apply a $\chi^2$ test. Appropriate references to the literature for this kind of analysis are Mosteller and Bush (1954) and Moses (1956). Since such tests are not completely standard, we give a certain amount of detail. (See also our previous discussion in Royden, et al. 1959.)

For each subject the 70 predictions for the two models were compared. Offers for which the models made the same prediction were deleted. Any remaining offer for which the utility model predicted the correct choice and the actuarial model predicted the incorrect choice or a tie was assigned a plus (+). An offer for which the utility model predicted a tie and the actuarial model a loss was also assigned a plus. Offers for which the roles of the actuarial and utility models were reversed in the above description were assigned a minus (−). The data are tabulated in Table III. Let $p$ be the probability of a plus. For each subject an exact significance level was found for the null hypothesis that $p = .5$ against the alternative that $p > .5$. A one-sided test was used to provide a clear basis for combining significance levels. (A two-sided test does not properly take account of varying directionality of results.) A similar one-sided test was run against the hypothesis that $p < .5$.

The seven significance levels for each one-sided test were now combined. Since the distributions which gave rise to the individual significance levels $p(X_i)$ are not continuous, rather than use Fisher's statistic

$$(1) \qquad \sum_{i=1}^{7} - 2 \log p(X_i),$$

following the suggestion of Moses (1956), Lancaster's statistic

$$(2) \qquad \sum_{i=1}^{7} -2 \log \left( \frac{p(X_i) + p(X_i + 1)}{2} \right)$$

was used, where $X_i$ is the observed number of +'s. Statistic (2) is approximately distributed as $\chi^2$ with 14 degrees of freedom.

The results are as follows. Against the hypothesis that $p > .5$, $\chi^2 = 32.48$, which for a two-sided test is significant at the .01 level. Against the hypothesis that $p < .5$, $\chi^2 = 20.98$ which (for a two-sided test) is significant at the .2 level. Two-sided tests were used on the combined results because there is no real a priori basis for fixing on a directionality. The combined significance levels obtained from the two tests indicate clear predictive superiority for the utility model. The unusually good predictions for subject #1 are the main factor in establishing this superiority. It is worth remarking that if *both* tests had resulted in a significance

TABLE III
PREDICTIVE DISAGREEMENT OF UTILITY AND ACTUARIAL MODELS

| Subject | Plus (Utility) | Minus (Actuarial) | Number of runs |
|---|---|---|---|
| 1 | 35 | 8 | 15 |
| 2 | — | — | — |
| 3 | 10 | 16 | 12 |
| 4 | 27 | 18 | 21 |
| 5 | 20 | 14 | 14 |
| 6 | 14 | 26 | 19 |
| 7 | 11 | 21 | 15 |
| 8 | 15 | 19 | 17 |

TABLE IV
PROBABILISTIC PREDICTIONS

| Subject | $[1, 1 + (1/3)(\eta - 1)]$ | $[1 + (1/3)(\eta - 1), 1 + (2/3)(\eta - 1)]$ | $[1 + (2/3)(\eta - 1), \eta]$ |
|---|---|---|---|
| 1 | .70 | .66 | .91 |
| 2 | — | — | — |
| 3 | .63 | .88 | .50 |
| 4 | .50 | — | .33 |
| 5 | .39 | .50 | .50 |
| 6 | .47 | .35 | .50 |
| 7 | .48 | .44 | .50 |
| 8 | .44 | .59 | .50 |

level $\leq .05$ the proper interpretation would seem to be that subjects should be divided into two subsets according to which model they most nearly satisfy, and no assessment of the over-all relative predictive worth of the two models should be attempted.[8]

The most obvious objection to the statistical analysis just made is that the 70 prediction observations on each $S$ are surely not independent. On the other hand, since the 70 offers were all different, strong negative or positive recency effects should not be observed. Speculation aside, one partial test of independence is a count of the number of runs of +'s and −'s. If $S$s are using some computational or strategic device over several trials (i.e., offers), one would expect the number of runs to be less than would be anticipated on a random basis. For each $S$ we tested the null hypothesis that the +'s and −'s are randomly distributed against the hypothesis that there are too few runs. The observed number of runs for each $S$ is given in Table III. The significance levels obtained were combined, using statistic (2) described above. The obtained result was $\chi^2_{14} = 10.68$, which for a one-sided test is significant at the .75 level. This result supports the hypothesis of independence.

In Table IV are found the results of checking the two probabilistic postulates on the discrimination parameter. Unfortunately, the data do not support these postulates. The predictions in Table IV were made by dividing the interval ratios which were $\leq \eta$ for each subject into three sets in the manner previously described. The option resulting in the larger utility interval was predicted as the subject's choice. In each cell of Table IV is recorded the ratio of correct predictions to the total number falling within the indicated interval. The result expected and desired was a rise from .5 (which would be anticipated assuming pure chance) in the lowest interval $[1, 1 + (1/3)(\eta - 1)]$ to some frequency close to 1 in the third interval, since at this point the numerator would be almost large enough to make the utility ratio greater than $\eta$ and hence lead to sure prediction. But only subject #1 comes close to giving an indication of the result expected, with values of: .7, .66, .91. 13 out of the 21 cell ratios were found to be greater than .5, which figure can be accounted for on the basis of chance alone.

It is worth noting that a pair of similar probabilistic postulates for the linear programming model described in Chapter III of Davidson, et al. (1957) were more satisfactorily verified.

[8] If an external criterion can be found to provide a foundation for dividing Ss into two appropriate subsets, the comparison is more interesting. An example of this is given in Royden et al., (1959).

## REFERENCES

Davidson, D., Suppes, P., & Siegel, S. *Decision making: An experimental approach.* Stanford, Calif.: Stanford Univ. Press, 1957.

Moses, L. E. Statistical theory and research design. *Ann. Rev. Psychol.*, 1956, 7, 233–258.

Mosteller, F. & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge, Mass.: Addison Wesley Publishing Co., 1954. Ch. 8.

Royden, H. L., Suppes, P., & Walsh, K. A model for the experimental measurement of the utility of gambling. *Behavioral Sci.*, 1959, 4, 11–18.

von Neumann, J. & Morgenstern, O. *Theory of games and economic behavior.* (2nd ed.) Princeton: Princeton Univ. Press, 1947.

(Manuscript received August 26, 1958)