Additivity, Utility, and Subjective Probability[1,2]

AMOS TVERSKY[3]

The University of Michigan, Ann Arbor, Michigan

The additive conjoint measurement model is applied to the study of decision making under certainty and risk. A data matrix is called additive if it is possible to rescale its cell entries such that their order is preserved and that every rescaled entry can be expressed as a sum of its row and column components. It is shown that the SEU model, according to which individuals attempt to maximize their subjectively expected utility, is equivalent to additivity for a specified class of risky choices.

In the experimental study eleven prisoners bid for both risky and riskless offers. Additivity is confirmed by the data supporting the independence between utility and subjective probability. Two alternative variants of the SEU model are used to derive subjective probability and utility functions for each subject. In order to account for the data, one needs either (a) a positive utility for gambling or (b) subjective probability functions where complementary events do not sum to unity. Neither variant is compatible with classical utility theory but both are successful in predicting an independent set of data. Relationships to existing data and implications for future research are discussed.

I. THEORY

Underlying most decision theories are two fundamental notions: the maximization principle and the decomposition hypothesis. The former asserts that people choose the alternative they consider best according to some criterion of worth; the latter states that the worth of an alternative can be decomposed into basic independent components.

In order to develop a theory of choice from these general principles, specific assumptions about the maximized function and the composition rule are made. In general, the maximized expression is a real-valued function termed utility which is assumed to reflect the observed choices and to preserve the hypothesized structure of the preference space.

[3] Present address: The Hebrew University of Jerusalem, Israel.

175

DECISIONS UNDER CERTAINTY

In the context of riskless choice, people are assumed to rank the alternatives according to their utilities, where the utility of a commodity bundle or a multiattribute alternative equals the sum of the utilities of its components. Despite well-known counter-examples of interacting commodities such as left and right shoes, this additive model prevailed in the economic theory of consumer behavior and has generated extensive theoretical work (Luce and Suppes, 1965; Samuelson, 1953).

In contrast to the theoretical and practical interest in the additive composition model, very little experimental work has been done to test it directly, although it has been assumed to hold in many models of scaling and data analysis. The earlier literature has been critically reviewed by Edwards (1954). Since then, Gulliksen (1955) has tested additivity among several other laws of utility combination, in a pair-comparisons study of food preferences. Thurstone and Jones (1957) scaled preferences for articles such as a table lamp or a brief case by having subjects choose between singles and/or pairs of items. The scale values were determined by setting the sum of the scale values of the single articles equal to the scale value of the pairs of articles. This additive assumption yielded a fairly close fit.

Adams and Fagot (1959) reported a study in which subjects were instructed to choose among job applicants varying in intelligence and ability to handle people; each attribute had four levels. Only 6 out of the 24 subjects satisfied the additive model. All the violations, however, were due to intransitivity of the pair-comparisons.

No attempt has been made to specify the psychological conditions under which additivity holds or fails to hold. Moreover, all choices in the above studies were hypothetical, i.e., no payoff was dependent on the subjects' choices. Finally, no appropriate statistics were developed to describe, test, or evaluate deviations from additivity.

DECISIONS UNDER RISK

Most models of risky choice are based on maximization of some form of expectation. Although this principle dates back to Bernoulli, it was not until it was reformulated by von Neumann and Morgenstern (1944) that it attracted psychologists' attention. Savage (1954) constructed the subjective expected utility (SEU) model, according to which people act (or should act) to maximize their subjective expected utility which equals the sum of the utilities of the outcomes weighted by their subjective probability of occurrence. Formally, let G be a gamble with outcomes $o_1, ..., o_n$ obtained contingent upon events $e_1, ..., e_n$; and let G' be a gamble with outcomes $o'_1, ..., o'_n$ obtained contingent upon events $e'_1, ..., e'_n$. Then there exists a real-valued utility function u, defined on outcomes, and a subjective probability function s, defined on events, such that G' is not preferred to G if and only if

(1.1)
$$\sum_{i=1}^{n} u(o_i) s(e_i) \geqslant \sum_{i=1}^{n} u(o'_i) s(e'_i).$$

The subjective expected value (SEV) and the expected utility (EU) models are defined by replacing subjective probabilities or utilities respectively by objective probabilities or monetary values. Thus, the models differ in whether probability and/or value are regarded as objective or subjective.

Unlike the riskless case, the above theory has stimulated rather extensive experimental work aimed not only at testing the models but also at determining the form of the utility and subjective probability functions. For a review of this literature see Edwards (1954; 1961; 1962) and Luce and Suppes (1965).

To bypass the serious difficulties involved in simultaneous measurement of utility and subjective probability for each subject, researchers have derived and tested some empirical consequences of the SEU model. The most recent attempts of this kind were by Becker, DeGroot, and Marschak (1964) and by Coombs, Bezembinder, and Goode (1967). However, any test of the model based on only some of its implications is incomplete because the model may be incorrect even when the tested implications are verified. Consequently, such implications may lead to the rejection of the model, but they cannot lead to its acceptance. Testing a model by using a necessary and sufficient condition for its existence avoids this problem and permits evaluation of the goodness of fit of the model.

The present study is concerned with a simple class of gambles of the form (a, p) in which one wins a positive amount a if an event p occurs and zero if p does not occur. Let D be a data matrix whose $D(a, p)$ entry is a measure of the worth of the gamble (a, p) such as the subject's minimal selling price or the proportion of times he prefers it to a standard gamble. In order to study the relationships between such data matrices and the SEU model the following definition is introduced.

(1.2) A data matrix $D = A \times P$ is said to be *additive* if there exist real-valued functions f, g, and ϕ defined on A, P, and D, respectively, such that

(i) $\phi(a, p) = f(a) + g(p)$,

(ii) $\phi(a, p) \geqslant \phi(b, q)$ if and only if $D(a, p) \geqslant D(b, q)$ for all a, b in A and p, q in P.
If, in addition, one can choose ϕ such that $\phi(a, p) = D(a, p)$ for all (a, p) in D then D is called *strictly additive*.

Clearly, any strictly additive matrix is additive but not conversely. An additive representation of D is simply any numerical assignment satisfying Conditions (i) and (ii). Thus a data matrix is said to be additive if it is possible to rescale its rows, columns, and cell entries such that the rank order of the cell entries is preserved, and every rescaled entry equals the sum of its row and column components. Luce and Tukey (1964) refer to this measurement model as simultaneous conjoint measurement to emphasize the fact that the dependent and the independent variables are measured simultaneously. The relationships between additivity and the SEU model for the type of gambles considered are given by the following elementary result.

(1.3) THEOREM. *For gambles of the form (a, p) the* SEU *model is satisfied if and only if D is additive.*

Proof.   First assume D is additive. Hence there exist functions f, g, and ϕ such that

$$D(a, p) \geqslant D(b, q) \qquad \text{if and only if}$$

$$\phi(a, p) = f(a) + g(p) \geqslant f(b) + g(q) = \phi(b, q).$$

Let

$$u(a) = e^{f(a)} \qquad \text{and} \qquad s(p) = e^{g(p)}.$$

Hence

$$D(a, p) = D(b, q) \qquad \text{if and only if}$$

$$u(a) \, s(p) \geqslant u(b) \, s(q),$$

which satisfies the SEU model provided the utility of zero is set equal to zero.

Conversely, assume the SEU model is satisfied and set $u(0) = 0$; hence there exist utility and subjective probability functions such that

$$D(a, p) \geqslant D(b, q) \qquad \text{if and only if}$$

$$u(a) \, s(p) \geqslant u(b) \, s(q).$$

Let

$$f(a) = \log u(a) \qquad \text{and} \qquad g(p) = \log s(p).$$

Hence

$$D(a, p) \geqslant D(b, q) \qquad \text{if and only if}$$

$$f(a) + g(p) = \phi(a, p) \geqslant \phi(b, q) = f(b) + g(q),$$

which completes the proof of (1.3).

The SEV and EU models are satisfied if and only if D has an additive representation subject to the constraint that u or s correspond to monetary values or objective probabilities respectively.

The fundamental assumption of all psychological expectation models, which is independent of any particular measurement method, is that utility and subjective probability contribute independently to the overall "worth" of a gamble. That is, judgments of desirability of outcomes are independent of judgments of likelihoods of events. More specifically, utility and subjective probability are compensatory but noninteracting. Studies and discussions of this hypothesis can be found in Edwards (1962), Irwin (1953), and Slovic (1966). Thus, if the additivity of a data matrix can

be tested, Theorem (1.3) provides a measurement-free test of expectation models which does not depend on the measurement of utility or subjective probability.

It should be pointed out that since the status quo is the zero point on the utility scale, $u(a)$ should be interpreted as a utility for a *more* dollars rather than the utility of a dollars. For a discussion of this interpretation and some related problems see Krantz and Tversky (1965).

The conditions under which additive representations exist have been studied extensively in recent years. Axiomizations which yield additive (or closely related) representations have been established by Debreu (1959; 1960); Pfanzagl (1959), Suppes and Winet (1955), and more recently by Krantz (1964) and by Luce and Tukey (1964). Necessary and sufficient conditions for additivity, for finite data matrices, have been established by Scott (1964) and by Tversky (1964). In order to state the latter result, certain constructions are introduced.

Let $D = A \times P$ be a data matrix with a, b, c in A and p, q, r in P. Each cell is represented by an ordered pair of the form (a, p) called a data element. The set of all data elements D is enriched by introducing formal sums of data elements where addition is defined component-wise. For example,

$$(a, p) + (b, q) + (c, r) = (a + b + c, p + q + r).$$

Let Δ denote the set of all such formal sums, denoted by lower-case Greek letters. Thus, Δ is the free Abelian semigroup generated by D.

Two binary relations on Δ are defined by:

(1.4) Let

$$\alpha = (a_1, p_1) + (a_2, p_2) + \cdots + (a_n, p_n)$$

and

$$\alpha' = (a_1', p_1') + (a_2', p_2') + \cdots + (a_n', p_n')$$

(i) $\alpha =_1 \alpha'$ whenever $D(a_i, p_i) = D(a_i', p_i')$ $i = 1, 2, ..., n$

(ii) $\alpha >_1 \alpha'$ whenever $D(a_i, p_i) \geqslant D(a_i', p_i')$ $i = 1, 2, ..., n,$
and $D(a_i, p_i) > D(a_i', p_i')$ for at least one i.

$\alpha \geqslant_1 \alpha'$ is defined as either $\alpha >_1 \alpha'$ or $\alpha =_1 \alpha'$. Hence, $\geqslant_1$ may be regarded as the additive closure of $\geqslant$. But since different sets of data elements may have the same formal sum, the latter do not determine their summands uniquely. Consequently, one may obtain both $\alpha \geqslant_1 \alpha'$ and $\alpha' >_1 \alpha$ for some α and α' in Δ. The following asymmetry axiom is introduced in order to exclude this possibility.

(1.5) ASYMMETRY AXIOM.   There are no α, α' in Δ such that $\alpha \geqslant_1 \alpha'$ and $\alpha' >_1 \alpha$.

The asymmetry axiom states, in effect, that the order of the formal sums is invariant with respect to different combinations. That is, if by applying (1.4) one obtains $\alpha \geqslant_1 \alpha'$, it is not possible to find a combination of data elements that yields $\alpha' >_1 \alpha$. Put differently, the axiom requires that for all α, α' in Δ, one and only one of the following holds: $\alpha >_1 \alpha'$, $\alpha =_1 \alpha'$, $\alpha' >_1 \alpha$, or α and α' are incomparable.

Alternatively, (1.5) may be formulated as an irreflexivity condition stating that there is no α in Δ such that $\alpha >_1 \alpha$. The equivalence of the above three forms is easily established and their essence is to assure that the additive closure of the observed order is a proper order relation. It has been shown (Tversky, 1964) that a data matrix of the type considered is additive if and only if it satisfies the asymmetry axiom (1.5). Furthermore, the above result can be generalized to infinite matrices and to partially-ordered data. See Tversky (1967) for a detailed discussion of this result in the context of a general theory of conjoint measurement. For some related developments see Fishburn (1967).

In the investigation of decision making and measurement theory, some consequences of the asymmetry axiom have been explored. Coombs, Bezembinder and Goode (1967) derived the following triple cancellation condition from the SEU model: Let $D = A \times P$ be a data matrix with outcomes a, b, c in A and events p, q, r in P. The SEU model implies that:

(1.6) If

$$D(a, q) \geqslant D(b, p),$$
$$D(d, p) \geqslant D(c, q),$$
$$D(b, r) \geqslant D(a, s),$$

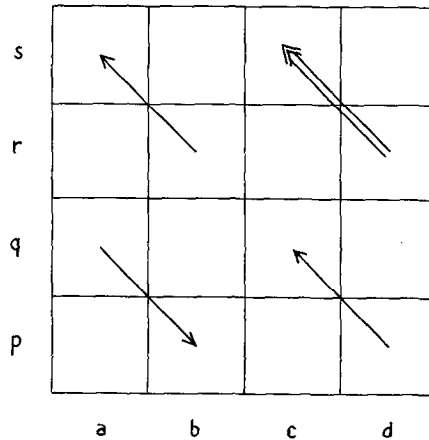then

$$D(d, r) \geqslant D(c, s).$$

FIG. 1.   A graphical illustration of triple cancellation (1.6).

This condition is illustrated in Fig. 1, where arrows indicate preference. The conclusion of the argument is denoted by a double arrow. Triple cancellation is an immediate consequence of the asymmetry axiom (1.5). For suppose (1.6) does not hold, hence $D(c, s) > D(d, r)$ and by applying definition (1.4) we obtain

$$
\begin{aligned}
\alpha &= (a, q) + (d, p) + (b, r) + (c, s) \\
&= (a + b + c + d, p + q + r + s) \\
&>_1 (b, p) + (c, q) + (a, s) + (d, r) \\
&= (a + b + c + d, p + q + r + s) = \alpha
\end{aligned}
$$

or $\alpha >_1 \alpha$. But since $\alpha =_1 \alpha$ for all α, by part (i) of (1.4), the asymmetry axiom is violated which completes the derivation of (1.6). Two other testable consequences of additivity are double cancellation and monotonicity (or independence) which were investigated by Adams and Fagot (1959), Debreu (1959) and Luce and Tukey (1964).

A data matrix is said to satisfy *double cancellation* whenever
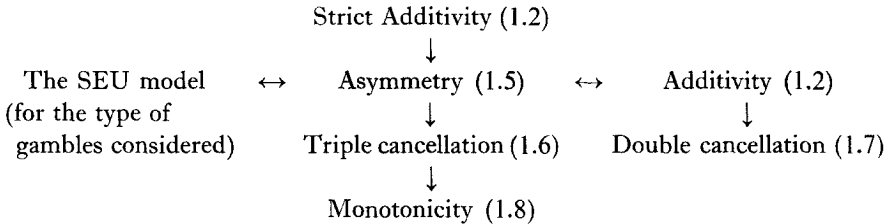
(1.7) $D(a, q) \geqslant D(b, p)$ and $D(b, r) \geqslant D(c, q)$ imply $D(a, r) \geqslant D(c, p)$.

A data matrix is called *monotone* (or independent) whenever

(1.8) $D(a, p) \geqslant D(a, q)$ if and only if $D(b, p) \geqslant D(b, q)$ and $D(a, p) \geqslant D(b, p)$ if and only if $D(a, q) \geqslant D(b, q)$.

To derive monotonicity from triple cancellation, assume $D(a, p) \geqslant D(a, q)$, but since $D(a, p) = D(a, p)$ and $D(b, p) = D(b, p)$ we obtain, by (1.6), $D(b, p) \geqslant D(b, q)$ as required. A symmetric argument applied to the second component completes the proof. The relationships between the various models and conditions are summarized by the following theorem. That implication does not hold unless indicated in the theorem can be shown by simple counterexamples.

(1.9) THEOREM:

Strict Additivity (1.2)
↓

The SEU model ↔ Asymmetry (1.5) ↔ Additivity (1.2)
(for the type of ↓ ↓
gambles considered) Triple cancellation (1.6) Double cancellation (1.7)
↓
Monotonicity (1.8)

Although for the type of gambles considered, additivity is both necessary and sufficient for the SEU model, it does not determine utilities and subjective probabilities uniquely. The set of all additive representations of a given finite data matrix is the set of all solutions of the corresponding system of linear inequalities. Thus, each

additive solution may be regarded as a point in some n-space, where the set of all solutions forms a polyhedral convex cone, denoted $C(D)$. The resultant scale may be viewed as a multidimensional, ordered metric scale. In order to obtain a (unique) numerical assignment, additional constraints are imposed.

A solution ϕ' is said to be a *least-squares* solution if it is closest to the data point D in the Euclidean-distance sense. Formally, ϕ' is a least-squares solution if it minimizes

$$(1.10) \qquad \delta(\phi, D) = \sum_a \sum_p (\phi(a, p) - D(a, p))^2$$

where the summation ranges over all a in A and p in P.

The least-squares solution is an additive representation of the data that distorts it minimally in the least-squares sense. If the data is strictly additive then D must be in $C(D)$ and the data coincide with its least-squares solution. To establish the uniqueness of the least-squares solution, assume ϕ' is not unique; hence there exists a solution $\phi'' \neq \phi'$ such that

(i) $\delta(\phi', D) = \delta(\phi'', D)$ and

(ii) there is no solution ϕ such that $\delta(\phi, D) < \delta(\phi', D)$.

Otherwise, ϕ' is either unique or not a least-squares solution. Since ϕ' and ϕ'' are equidistant from D, by (i), they lie on the boundary of a hypersphere whose center is D. Let ϕ be any point on the open line segment (ϕ', ϕ''). Since the solution space, $C(D)$, is convex, ϕ must also be in $C(D)$; but, since ϕ is clearly inside the hypersphere, $\delta(\phi', D) > \delta(\phi, D)$ contrary to (ii) above, which establishes the uniqueness of the least-squares solution. Note that although ϕ' is unique, the corresponding scales for the two factors (f and g) are determined only up to a common additive constant.

THE UTILITY OF GAMBLING

Related to the problem of measuring utility of outcomes in a risky situation is the puzzling problem of measuring the utility of gambling. Royden, Suppes, and Walsh (1959) constructed the only available model for experimental measurement of the utility of gambling based on the assumption that people maximize the sum of the expected monetary value and the utility of gambling for the particular lottery. Although the model allows for individual parameters, its predictive power did not exceed that of the simple expected value model. The classical formulation of von Neumann and Morgenstern does not allow a specific utility for gambling.

The additive model offers a new approach to this problem. Let $D(a, p)$ be the subject's minimal selling price of the gamble (a, p) in which he wins a positive amount of money a with probability p. If the subject's prices are compatible with the SEU model, then, by (1.3), there exist functions ϕ, f and g defined on prices, outcomes and events such that:

$$\phi(D(a, p)) = f(a) + g(p)$$

where ϕ preserves the order of the cell entries. Both ϕ and f, however, are defined on the same domain, i.e., money. Hence we have two utility functions for money, one for selling prices (ϕ) and one for risky outcomes (f). The normalized difference between them, denoted $\omega(a)$, is proposed as a measure of the utility of gambling:

$$(1.11) \qquad \omega(a) = \frac{f(a) - \phi(a)}{f(a) + \phi(a)}.$$

This measure is obtained by comparing the utility of a sure-thing, e.g., a price, to the utility of a risky outcome. In the classical formulation of the SEU model these functions are identical, that is, $\phi(a) = f(a)$ for all a, and hence $\omega(a) = 0$.

An alternative measure, denoted $\theta(a)$, is based on the difference between utilities derived from risky and riskless choices. Let $f(a)$ be the utility of a derived via the additive model from risky choices, and let $h(a)$ be the utility of a derived via the additive model from riskless choices. Define

$$(1.12) \qquad \theta(a) = \frac{f(a) - h(a)}{f(a) + h(a)}.$$

Since the former measure (ω) requires interpolation between utilities whereas the latter measure (θ) can be obtained directly, the latter is used in the present study.

The additive model allows the utility for gambling to be positive, zero, or negative. Thus it permits study of problems concerned with the utility for gambling as well as the utility for different types of risk, which cannot be studied within the classical framework. In general, θ may assume different values for different values of a. Whenever θ is fairly uniform over the range considered, the average $\theta(a)$, denoted θ, may be regarded as an index of the utility of gambling. Note that both measures of the utility of gambling are independent of subjective probability. They are based on the fact that it is possible to obtain two different scales of value form the additive model. One cannot obtain a similar measure based on subjective probability since it is not possible to obtain two different scales of subjective probability under two utility levels without violating additivity.

In the following experimental study, the additive model is employed to test some of the basic assumptions of utility theory and to construct utility and subjective probability scales from choices under certainty and risk.

II. METHOD

Subjects. Eleven male inmates from the State Prison of Southern Michigan in Jackson, Michigan, whose ages ranged from 23 to 50 served as subjects in the experiment. They were selected, on the basis of their cooperativeness, from volunteers who had I.Q.'s above 100 and who had participated in an earlier gambling experiment. In the previous experiment, the subjects

spent approximately 20 hours choosing between two-outcome gambles using the display described below. Prisoners were chosen as subjects since nonmonetary commodities such as cigarettes and candy are used as currency in the prison and thus are effective payoffs. Also, the expected payoff for a two-hour session provided a strong incentive for the inmates because it constituted a significant amount relative to their usual income.

function $x = y$ is plotted as a dotted line in all figures. Inspection of the graphs reveals that the bids were above expected value in the risky sets (I, II, and IV) and below expected value in the riskless set (III). The bids for cigarettes (I) and candy (II) were nearly equal, with the latter slightly higher.
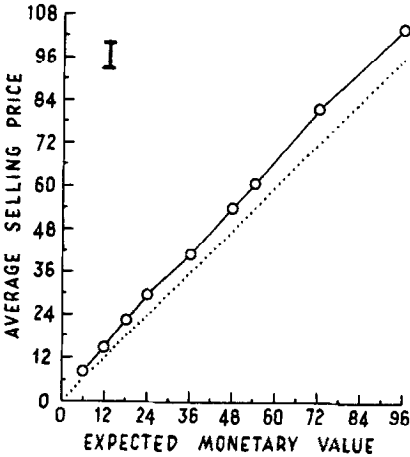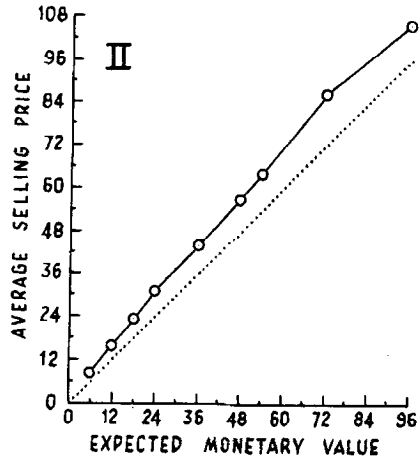

FIG. 4.   Average bids for Set I.

FIG. 5.   Average bids for Set II.

They were given no money to start with but were told that they were entitled to play or sell three options at the end of each session. They were further told that the experimenter would try to take advantage of them by letting them play the option if their price were high or by buying their right to play if their price were low.

Three options were selected in advance to be played at the end of each session. For each subject, the experimenter determined a buying price chosen randomly within one standard deviation of the expected value of the option. If the subject's selling price exceeded this predetermined value, the subject played the gamble or received the riskless option. On the other hand, if the buying price exceeded the subject's bid, the subject sold the option at the experimenter's buying price. It was pointed out to the subjects that under these conditions they could do no better than write down their "true" lowest selling price, or their indifference point. For a discussion of this point, see Becker *et al.* (1964, p. 228).

The gambling device was a wheel of fortune with a rotating spinner which stopped on one out of ten spots (see Fig. 2). The gambles were played by having each subject spin the spinner and win the amount indicated if the spinner stopped on a black spot, and nothing otherwise.

The total winnings per session ranged from $0-$2.40, with an expected value of approximately $1. Because of prison regulations, money had to be deposited in the subjects' accounts. Cigarettes and candies, however, were distributed at the end of each session. Since a pack of cigarettes and a bag of candy have the same price (30¢) at the prison store, Sets I and II were matched in expected value. In general, the study was designed to maximize the number of different options having the same expected value in order to provide the most rigorous test of the theory.
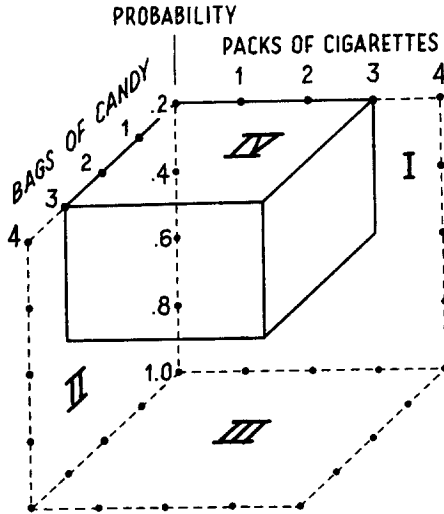

FIG. 3. A geometrical representation of the experimental design.

III. RESULTS

A gross depiction of the data is presented in Figs. 4-7 where the average selling prices, or bids, are plotted against the monetary expected value of the gambles. The

function $x = y$ is plotted as a dotted line in all figures. Inspection of the graphs reveals that the bids were above expected value in the risky sets (I, II, and IV) and below expected value in the riskless set (III). The bids for cigarettes (I) and candy (II) were nearly equal, with the latter slightly higher.
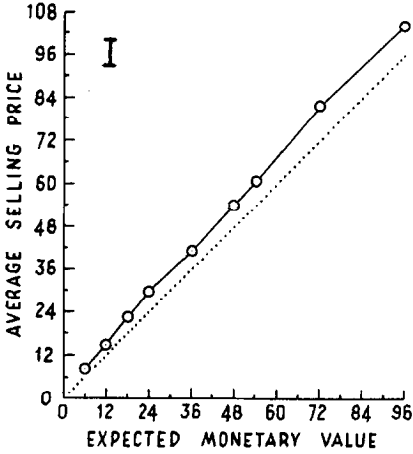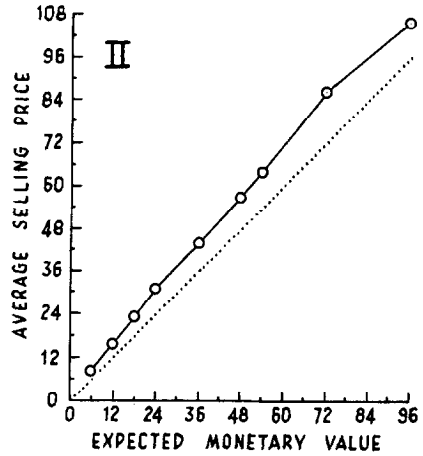
FIG. 4. Average bids for Set I.

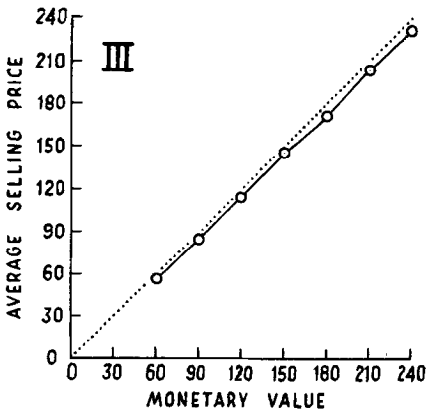FIG. 5. Average bids for Set II.

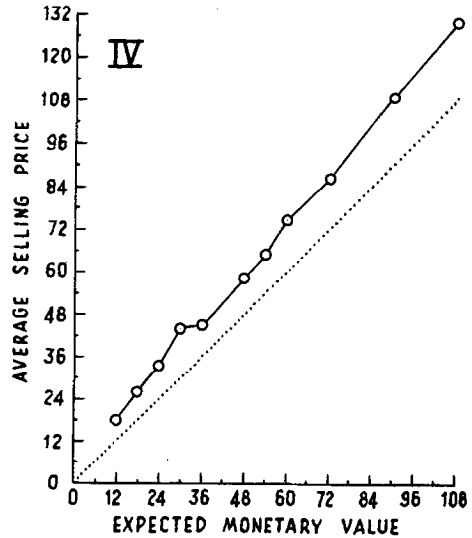FIG. 6. Average bids for Set III.

FIG. 7. Average bids for Set IV.

The analysis is divided into three parts. First, additivity and strict additivity of the obtained bidding matrices were tested and the expectation models were compared. Next, two variants of the additive model were applied to derive subjective probability and utility scales for each subject. Finally, the derived scales were used to predict an independent set of data (Set IV) and to compare some alternative choice models.

ADDITIVITY ANALYSIS

The test for additivity was based on the median selling prices of each of the eleven subjects, for all options from Sets I, II, and III. If we let $D(a, p)$ be the median bid for the option (a, p) then the SEU model is equivalent, by (1.3), to the additivity of D. Hence, under utility theory there exist numbers $f(a), f(b), g(p)$ and $g(q)$ such that

$$(3.1) \qquad D(a, p) \geqslant D(b, q) \qquad \text{iff} \qquad f(a) + g(p) \geqslant f(b) + g(q)$$

$$\text{for all } a, b \text{ in } A \text{ and } p, q \text{ in } P.$$

The above equation defines a system of linear inequalities for each data matrix of each subject. An IBM 7090 was programmed to solve the inequalities, if possible, or to find a numerical assignment which minimizes the number of inversions if no perfect solution could be found. A pair of cells $(a, p), (b, q)$ is said to form an *inversion* with respect to a given numerical assignment whenever:

$$(3.2) \qquad D(a, p) \geqslant D(b, q) \qquad \text{but} \qquad f(a) + g(p) < f(b) + g(q).$$

The algorithm was based on systematic eliminations of inequalities from the system (3.1) until a solvable subset, yielding the smallest number of inversions, was found. To obtain a unique numerical assignment, the least-squares solution (1.10) was approximated using a linear programming method developed in Tversky and Zivian (1966).

The other expectation models EU, SEV, and EV were tested by solving (3.1) in the manner described, subject to the constraint that $f(a)$ or $g(p)$ or both are known constants corresponding to the logarithms of value (a) and probability (p) respectively.

The riskless bidding matrices were solved under two models:

(a)   The riskless utility, or the RU model, where the utilities were considered unknowns.

(b)   The riskless value or the RV model, where the utilities were considered as known constants corresponding to the monetary values of cigarettes and candy.

Kendall's rank-order correlation (τ) between each data matrix and its solution provided a measure of the degree of additivity. τ is the difference between the proportion of pairs of cell entries which are ordered like their additive scale values and the proportion of pairs of cell entries which are ordered differently from their additive scale values. Thus, if p denotes the proportion or the probability of an inversion, then

$\tau = (1 - p) - p$, and $p = \frac{1}{2}(1 - \tau)$. The values of τ for each subject under the various models for Sets I, II, and III are presented in Table 1.

TABLE 1

RANK-ORDER CORRELATION (τ) BETWEEN THE DATA AND
THE ADDITIVE SOLUTION OBTAINED UNDER EACH MODEL

Models	SEU		SEV		EU		EV		RU	RV
Sets	I	II	I	II	I	II	I	II	III	III
Subjects										
1	1	1	.949	.916	.916	.916	.812	.776	1	1
2	1	1	.966	.966	.966	.966	.966	.966	1	1
3	1	.983	.966	.966	.950	.999	.932	.888	.983	.949
4	1	.983	.949	.949	.916	.949	.854	.854	1	1
5	.950	1	.950	.966	.949	.966	.949	.966	1	1
6	1	1	.966	.983	.966	.949	.966	.940	1	.949
7	.950	.983	.950	.966	.950	.966	.932	.966	1	.888
8	1	1	.966	.966	.966	.966	.966	.966	1	1
9	.966	1	.950	.966	.966	.966	.949	.966	.966	.880
10	1	1	.966	.966	.966	.966	.966	.966	1	.897
11	1	1	.978	.950	.949	.949	.854	.949	1	1
Av.	.988	.995	.959	.960	.951	.955	.922	.928	.995	.960

Out of the 33 data matrices analyzed for additivity without additional constraints (that is, under the SEU and the RU models), 25 were perfectly additive and the average probability of an inversion for all these data was less than .005. Naturally, this probability increases as additional constraints are imposed on the solutions by the stronger models. Although the SEV and the EU models are equally constrained, the former was slightly better than the latter.

STRICT ADDITIVITY

According to the theory, utilities combine additively, whereas utility and subjective probability combine multiplicatively. Hence, in order to test strict additivity (1.2), a logarithmic transformation was first applied to all the risky bids from Sets I and II. The transformed risky bids as well as the riskless ones (from Set III) were then submitted to a series of individual two-factor analyses of variance. This analysis provides a statistical model for the test of the hypothesis that the two factors of each

data matrix are strictly additive in the sense that there is no significant interaction between them. Unlike the previous analysis which employed only median bids, the present analysis was based on all the data. Since all main effects were highly significant for all subjects, only the F-ratios for the interaction terms are reported. Table 2 presents the F-ratios for all subjects and sets, where a star denotes statistical significance beyone the $p = .1$ level.

TABLE 2

F-RATIOS FOR THE INTERACTION TERMS[a]

Subjects	Sets		
	I	II	III
1	.545	.899	.762
2	1.231	.979	1.087
3	.730	.221	.722
4	.180	1.245	1.592
5	1.755	3.391*	1.168
6	.624	.706	.774
7	3.91	1.271	1.424
8	.335	1.366	.690
9	.591	.983	.790
10	.032	.678	1.020
11	1.241	.337	.542

[a] Statistical significance beyond the .1 level is indicated by a single star. All F-ratios are based on 9 and 32 degrees of freedom except for Subjects 2 and 10 whose F-ratios are based on 9 and 16 degrees of freedom.

Out of the 33 bidding matrices, only a single one revealed a significant interaction. Hence the data show that, to the accuracy allowed by the variability of the bids, the subjects' bids for the risky options are expressable as multiplicative combinations of their probability and value components, and that the subjects' bids for the riskless options are expressible as additive combinations of their two value components.

UTILITY AND SUBJECTIVE PROBABILITY

Since, in general, strict additivity is preserved only by linear transformations, the strict additivity of the riskless data implies that the utility function for money is practically identical to the actual money value. [Indeed, the obtained approximate least-square solution (1.10) was almost indistinguishable from the actual bids.] Strict additivity, however, imposes no constraints on the utilities for cigarettes and candy.

Although the solution is uniquely determined, the scale values of the rows and the columns are determined only up to a common additive constant. Two methods based on two different variants of the classical SEU model, called Models 1 and 2, were used to determine the additive constant. According to Model 1, subjective probabilities of complementary events sum to one. Model 2, on the other hand, requires the same utilities for risky and riskless choices; that is, utility is assumed to be risk-invariant and no utility for gambling is allowed. Utilities are denoted risky or riskless depending on whether they were derived from risky or riskless bids. Subjective probabilities are denoted type 1 or type 2 depending on whether they were derived from Model 1 or from Model 2.

The construction of the riskless utilities was based on the observation that the two commodities were equal in monetary value, and that the observed riskless bidding matrices were indeed symmetric. Consequently, the additive constant was chosen so as to equate the means of the two scales. It is important to realize that once the additive constant has been chosen for a pair of scales, the scales are uniquely determined.

The problem of identifying events in choice experiments is subtle because every event is, in a sense, unique. Since no biases associated with the color or the position of the winning spots were found, events were identified in terms of the number of winning spots on the wheel. Model 1 regards events having complementary ratios of winning to losing spots as complementary events. Since in the present study there is a unique objective probability associated with each of the experimentally identified events, subjective probability may be viewed as a function of objective probability.

The scales based on Model 1 were constructed as follows:

From strict additivity and (1.3) we obtain

$$D(a, p) = f(a) + g(p) = u(a) s(p).$$

Hence

$$\sum_p D(a, p) = u(a) [s(.2) + s(.4) + s(.6) + s(.8)] = u(a) (1 + 1)$$

and

$$u(a) = \frac{1}{2} \sum_p D(a, p)$$

since the expression inside the brackets is the sum of the subjective probabilities of two pairs of complementary events. Once the utility scale for each subject had been constructed, we solved for the subjective probabilities.

Since

$$D(a, p) = u(a) s(p),$$

$$\sum_a D(a, p) = \sum_a u(a) s(p) = s(p) \sum_a u(a)$$

and

$$s(p) = \frac{\sum\limits_{a} D(a, p)}{\sum\limits_{a} u(a)}.$$

The subjective probabilities for Model 2 were derived from the same equation, except that the riskless utilities replaced the risky ones in the denominator of the above equation. Note that according to Model 2 the subjective probabilities of two complementary events need not sum to unity.
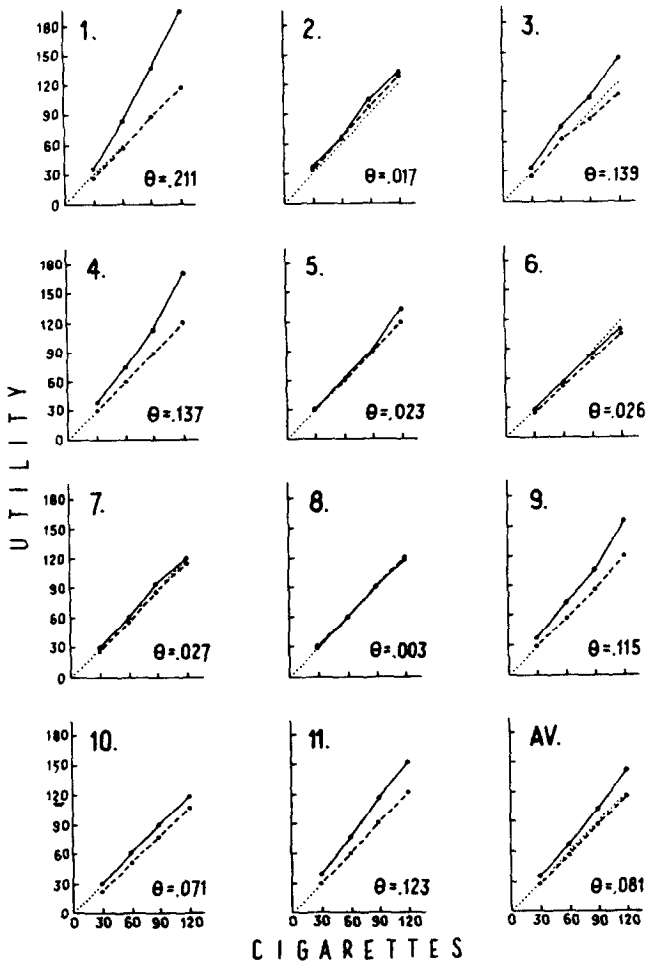

FIG. 8.   Risky (solid line) and riskless (dashed line) utility functions and θ's for cigarettes for each subject.

Both risky and riskless utilities for each subject, together with his average utility for gambling (1.12), are presented in Figs. 8 (for cigarettes) and 9 (for candy). The derived utilities may be described as follows:
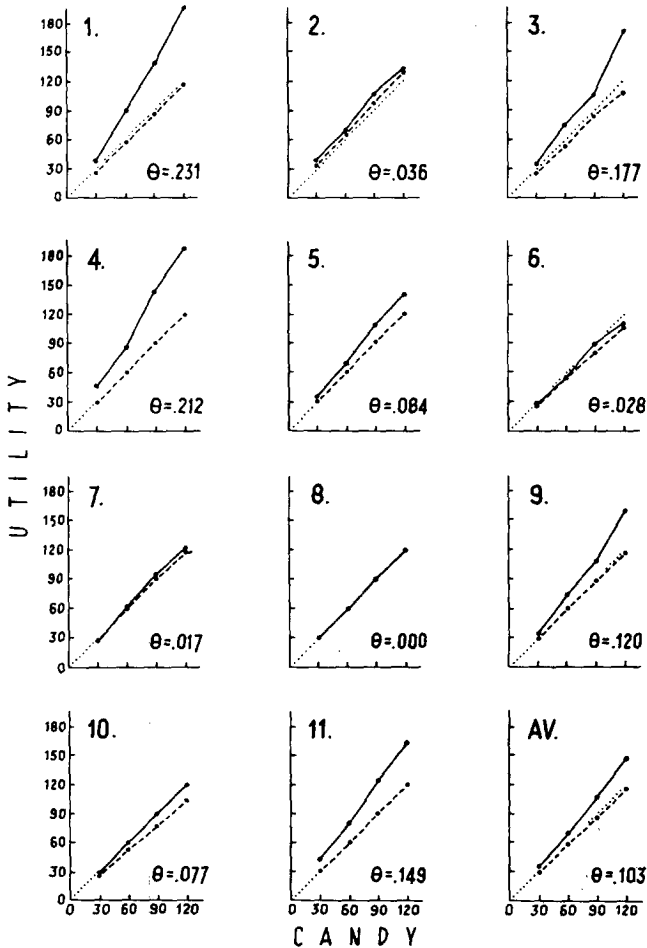

FIG. 9. Risky (solid line) and riskless (dashed line) utility functions and θ's for candy for each subject.

(a)   The utility functions for cigarettes and candy were nearly identical for both risky and riskless bids.

(b)   The riskless utilities were very similar for all subjects: linear, and slightly below the $x = y$ line.

(c)   The risky utilities exceeded the $x = y$ line almost everywhere, yielding a positive index of utility for gambling for 10 out of 11 subjects.

(d)   The product-moment correlation between the two indices of utility for gambling obtained from Sets I and II was .95, indicating invariance of the index with respect to commodity.

The subjective probabilities derived from Sets I and II were nearly identical for both models. Hence, the average subjective probabilities for each subject under the two models are presented in Fig. 10. The value of s at the lower right of each graph is the average sum of the subjective probabilities of two complementary events under Model 2. The derived subjective probabilities may be described as follows:
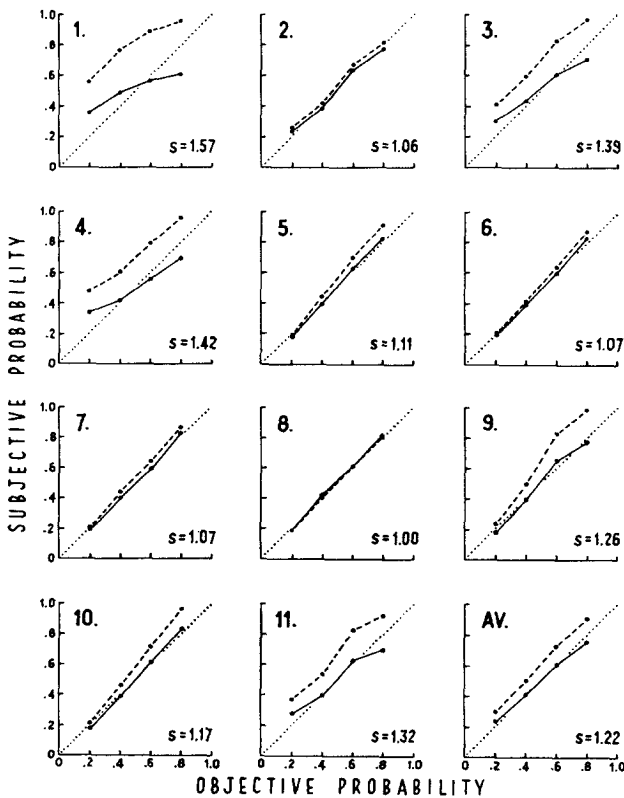

FIG. 10.   Average type 1 (solid line) and type 2 (dashed line) subjective probability functions and s's for each subject.

(a) Under Model 1, most subjects overestimated the low probability (.2) and underestimated the high probability (.8). The subjective probabilities of the rest of the subjects coincides with the objective probabilities.

(b) Subjective probabilities derived from Model 2 exceeded the objective probabilities everywhere for all but one subject.

(c) The two subjective probability scales appeared to be constant multiples of each other.

INDEPENDENT PREDICTIONS

Models 1 and 2, along with the classical SEU and the simple EV models, were compared in predicting the median bids of Set IV for each subject. This set contains gambles of the form $(p; a, b)$ in which one can win a packs of cigarettes and b bags of candy with probability p, or win nothing with probability $1 - p$. The expected utility of such gambles is, therefore, $s(p)[u(a) + u(b)]$. The predictions from Model 1 were obtained by multiplying the sum of the risky utilities by a type 1 subjective probability. The predictions from Model 2 were obtained by multiplying the sum of the riskless utilities by a type 2 subjective probability. The predictions from the classical SEU model were obtained by multiplying the sum of the riskless utilities by a type 1 subjective probability. The expected values were computed directly. Although the models differ in the number of free parameters, no parameters were estimated from the predicted data. The average absolute deviations (in cents) for each subject under each one of the models are given in Table 3.

The hypothesis that there is no significant difference between models (in their average absolute deviations) was tested for the following three pairs of models: (i) 2 vs. 1. (ii) 1 vs. SEU. (iii) SEU vs. EV. The variances of the absolute deviations (from all 27 data points) were computed under each model and three two-tailed t-tests for each subject were performed. The results are summarized in Table 3 where a single star between a pair of models indicates that the difference between them was significant beyond the .05 level and a pair of stars indicates that the difference was significant beyond the .005 level. Despite the usual difficulties in interpreting the results of multiple t-tests, the overall picture may be described as follows:

(a) The behavior of three subjects (7, 8, and 10) was in almost perfect agreement with the EV model. Consequently, the prediction of their data could hardly be improved by any of the other models.

(b) The SEU model was better than the EV model for six out of the remaining eight subjects, although the difference was statistically significant for only two subjects.

(c) Model 2 was better than Model 1 for eight subjects, although only for one was the difference statistically significant.

TABLE 3

AVERAGE ABSOLUTE DEVIATION (IN CENTS) FOR SET IV
UNDER EACH OF FOUR MODELS[a]

Subjects	Model 2	Model 1		SEU		EV
1	5.97	8.42	**	35.87		40.55
2	1.73	2.52		2.06	**	8.31
3	4.06	6.21	**	22.37		20.71
4	4.82	4.64	**	23.27		27.35
5	7.00	6.37	*	12.52		11.93
6	3.80	1.20		2.20	*	5.70
7	1.10	1.15	*	2.93		0.80
8	0.07	0.40		0.07		0.30
9	2.00	3.70	**	15.06		15.30
10	1.10	1.20	**	7.69	**	0.40
11	3.90	4.20	**	21.29		25.50
Av.	3.23	3.64		13.21		14.26

[a] Statistical significance beyond the .05 and the .005 levels are indicated by one and two stars, respectively.

(d) The major finding of the present analysis was the large difference between Model 1 and the SEU model which was statistically significant for eight out of the 11 subjects.

(e) Both Model 1 and Model 2 provided a very good fit to the data. The average deviations (over all subjects) for the models were 3.2¢ and 3.6¢, as compared with values of 13.2¢ and 14.3¢ for the SEU and the EV models, respectively.

For a more detailed view of the data, the median bids were averaged over all subjects and the observed averages, along with the predicted averages under all four models, are presented in Table 4.

IV. DISCUSSION

In this section the empirical findings of the study are discussed in relation to previous work and some critical comments are offered.

The additivity analysis of the bidding matrices supported the SEU model. Out of the 22 matrices analyzed under the SEU model, 16 were perfectly additive and the

TABLE

OBSERVED AND PREDICTED AVERAGE BIDS FOR SET IV

		Observed			Model 2			Model 1			SEU			EV			PROBABILITY
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
C	3	86.8	110.4	130.5	86.0	106.8	128.1	84.0	104.4	127.0	70.2	87.2	106.1	72	90	108	.6
	2	65.8	86.7	107.4	62.8	83.6	104.9	62.8	84.0	105.8	52.5	70.2	88.4	54	72	90	.6
	1	41.0	65.3	85.5	40.9	61.6	83.0	40.2	62.8	84.7	33.6	52.5	70.8	36	54	72	.6
A	3	60.3	78.2	95.9	57.8	71.8	86.1	57.5	71.5	87.0	47.2	58.6	71.3	48	60	72	.4
N	2	44.4	59.2	71.4	42.2	56.2	70.5	43.0	57.5	72.5	35.3	47.2	59.4	36	48	60	.4
D	1	28.4	43.4	59.6	27.5	41.4	55.8	27.5	43.0	58.0	22.6	35.3	47.6	24	36	48	.4
Y	3	33.0	44.9	50.7	33.8	42.0	50.4	34.5	42.9	52.2	27.6	34.3	41.8	24	30	36	.2
	2	26.4	37.1	42.7	24.7	32.9	41.3	25.8	34.5	43.5	20.6	27.6	34.8	18	24	30	.2
	1	17.1	24.9	35.6	16.1	24.2	32.6	16.5	25.8	34.8	13.2	20.6	27.8	12	18	24	.2

CIGARETTES

overall proportion of inversions was less than .005. Strict additivity was also confirmed by the data as only one out of the 33 data matrices revealed a significant interaction. Hence, the subjects' bids can be expressed as simple additive (or multiplicative) combinations of the options' components. Furthermore, since strict additivity is preserved only by linear transformations, it implies the linearity of the utility for money.

Strict additivity was not directly tested in the past. The majority of the experimentally-derived utility functions, however, were not incompatible with the linearity hypothesis. See Edwards (1955) and Tversky (in press) for examples.

In studies that found interaction between utility and subjective probability, (Irwin, 1953; Slovic, 1966) the latter was directly estimated by the subject rather than inferred from his choices. Moreover, payoffs were typically independent of the subject's estimates. It remains to be seen whether a systematic interaction effect can be demonstrated in studies where payoffs are contingent upon the subjective probabilities.

In the absence of additional constraints on the data, additivity exhausts the empirical content of the SEU model. However, once the riskless utilities are introduced and assumptions concerning complementary events are made, the invariance of the utilities and the complementarity of the subjective probabilities can be tested. Two variants of SEU model were employed to derive the scales. Model 1 is essentially equivalent to Davidson, Suppes and Siegel's (1957) "Weak finitistic rational choice structure" in which the subjective probabilities of two complementary events sum to one Since their axiomatization is limited to choices between risky options, utilities need not be risk-invariant.

Model 2 is practically identical to Edwards' (1962) "weighted SEU model," in which utilities are risk-invariant, but subjective probabilities of complementary events need not sum to one. An implication of this reasoning is that the total amount of subjective probability depends on the events from which it is composed, e.g., the total amount of pie depends on the way it is cut. Using both models, two subjective probability and utility functions were obtained for each subject. These are unique up to an identity rather than a linear transformation.

Subjective probability type 1 coincided with objective probability for some of the subjects; others overestimated the low probability and underestimated the high one. The latter effect was found in numerous studies, some of which are summarized in Luce and Suppes (1965). Subjective probability type 2 exceeded objective probability everywhere for all but one subject. Similar functions were obtained by Edwards (1955). It is interesting to note that the two most commonly found subjective probability functions correspond to those derived from Models 1 and 2.

The risky and the riskless utilities differed markedly with the former exceeding the latter everywhere for all but one subject. The existence of discrepancies between the scales derived from the two models show that the data may be accounted for by two alternative models. If subjective probabilities of complementary events add to

unity, then the risky utility function exceeds the riskless one. This is inadmissible under classical utility theory. Alternatively, if utility is risk-invariant, then the subjective probabilities of complementary events do not sum to unity which is incompatible with any expectation model.

The basic finding of overbidding for risky offers and underbidding for riskless ones may be explained by either: (i) a positive utility for gambling, or by (ii) a general overestimation of the objective probabilities. Thus, the data are explicable by either one of two incompatible additive models, each of which contradicts the classical SEU model. This conclusion does not depend on any particular numerical solution because if utility is risk-invariant, subjective probabilities have to exceed objective probabilities everywhere and hence they cannot sum to one. Conversely, if subjective probabilities sum to one, risky utilities have to exceed riskless utilities everywhere and hence utility cannot be risk-invariant.

The SEU model encompasses three fundamental assumptions:

(a)   The independence principle: utility and subjective probability contribute independently to overall worth of a gamble.

(b)   The invariance assumption: utility or subjective value are risk-invariant and no utility for gambling is allowed.

(c)   The complementarity notion: subjective probabilities of complementary events add to unity.

Additivity was employed to test Assumption (a). Assumptions (b) and (c) were tested indirectly by comparing Models 1 and 2. The data showed that although Assumption (a) was confirmed, the acceptance of (b) led to the rejection of (c) and the acceptance of (c) led to the rejection of (b). Hence the failure to satisfy simultaneously the complementarity and the invariance principles led us to reject the classical SEU model in spite of the fact that the independence principle was satisfied. This conclusion emphasizes the need for comparisons between risky and riskless offers in order to test the above assumptions.

One way to interpret the utility for gambling within the classical framework is to redefine the consequences so that winning a certain amount in a gamble is regarded as a different consequence from receiving the same amount as a sure-thing. In spite of its apparent plausibility, this approach does not yield testable predictions because consequences cannot be identified independently of gambles. Furthermore, according to this approach utility has to be defined not on monetary outcomes but on abstract consequences which depend on subjective probabilities as well. This renders the experimental identification of consequences practically unfeasible and the SEU model virtually invulnerable. Thus, although it is possible to argue that the experimental identification, rather than the SEU model, is in error, the fruitfulness of such an approach is questionable.

The discrepancy between the bids for the risky and the riskless offers may be explained in terms of the availability of the offers. The riskless offers were readily available for the subjects at the prison store and consequently their selling prices were essentially linearly related to monetary value. The risky offers, however, were not easily available and consequently had relatively higher selling prices. The availability argument can be used to explain the obtained utilities or to provide an alternative explanation which is independent of utility theory.

Both Models 1 and 2 predicted the data of Set IV quite well. Model 2 tended to surpass Model 1 in accuracy, indicating that the data could be fit better when the utility rather than subjective probability was constrained. This finding is reminiscent of the slight superiority of the SEV model over the EU model. These results agree with Edwards' (1955) finding that the SEV model predicted choices between two-outcome bets significantly better than the EU model. Thus, in spite of the formal symmetry between the models, the data seem to be better accounted for by subjective probabilities and objective values than by objective probabilities and subjective values.

After more than 15 years of experimental investigation of decisions under risk, the evidence on the descriptive validity of the SEU model is still inconclusive. In view of the extreme generality of the model on the one hand and the experimental limitations on the other, it seems that the basic question is not whether the model can be accepted or rejected as a whole. Instead, the problem is to discover which of the assumptions of the model hold or fail to hold under various experimental conditions. The present study showed that the SEU model was satisfied by each set separately but was violated when the risky and the riskless options were combined. Thus, although utilities and subjective probabilities were additive and subjective probabilities were commodity-invariant, utilities were not risk-invariant. Hence, the subjects' general preference structure could not be described by classical utility theory, although each separate data matrix is consistent with it. The usefulness of utility theory for the psychology of choice, however, depends not only on the accuracy of its predictions but also on its potential value as a general framework for the study of individual choice behavior.

REFERENCES

ADAMS, E., AND FAGOT, R. A model of riskless choice. *Behavioral Science*, 1959, **4**, 1-10.

BECKER, G. M., DEGROOT, M. H., AND MARSCHAK, J. Measuring utility by a single response sequential method. *Behavioral Science*, 1964, **9**, 226-233.

COOMBS, C. H., BEZEMBINDER, T. G. G., AND GOODE, F. M. Testing expectation theories without measuring utility or subjective probability. *Journal of Mathematical Psychology*, 1967, **4**, 72-103.

DAVIDSON, D., SUPPES, P., AND SIEGEL, S. *Decision making*. Stanford, California: Stanford Univer. Press, 1957.

DEBREU, G. Cardinal utility for even-chance mixture of pairs of sure prospects. *Review of Economic Studies*, 1959, **26**, 174-177.

DEBREU, G. Topological methods in cardinal utility theory. *In* K. J. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical methods in the social sciences*. Stanford, California: Stanford Univer. Press, 1960. Pp. 16-26.

EDWARDS, W. The theory of decision making. *Psychological Bulletin*, 1954, **51**, 380-417.

EDWARDS, W. The prediction of decisions among bets. *Journal of Experimental Psychology*, 1955, **50**, 201-214.

EDWARDS, W. Behavioral decision theory. *Annual Review of Psychology*, 1961, **12**, 473-498.

EDWARDS, W. Subjective probabilities inferred from decisions. *Psychological Review*, 1962, **69**, 109-135.

FISHBURN, P. C. Conjoint measurement in utility theory with incomplete product sets. *Journal of Mathematical Psychology*, 1967, 4, 104-119.

GULLIKSEN, H. Measurement of subjective values. Research Bulletin, Educational Testing Service, 1955.

IRWIN, F. W. Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, 1953, **21**, 329-335.

KRANTZ, D. H. Conjoint measurement: the Luce-Tukey axiomatization and some extensions. *Journal of Mathematical Psychology*, 1964, **1**, 1-27.

KRANTZ, D. H., AND TVERSKY, A. A critique of the applicability of cardinal utility theory. Technical report MMPP 65-4, Michigan Mathematical Psychology Program, 1965.

LUCE, R. D., AND SUPPES, P. Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology, III*. New York: Wiley, 1965. Pp. 249-410.

LUCE, R. D., AND TUKEY, J. W. Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1964, **1**, 1-27.

PFANZAGL, J. A general theory of measurement applications to utility. *Naval Research Logistics Quarterly*, 1959, **6**, 283-294.

ROYDEN, H. L., SUPPES, P., AND WALSH, K. A model for the experimental measurement of the utility of gambling. *Behavioral Science*, 1959, **4**, 11-18.

SAMUELSON, P. A. *Foundations of economic analysis*. Cambridge: Harvard Univer. Press, 1953.

SAVAGE, L. J. *The foundation of statistics*. New York: Wiley, 1954.

SCOTT, D. Measurement models and linear inequalities. *Journal of Mathematical Psychology*, 1964, **1**, 233-248.

SLOVIC, P. Value as a determiner of subjective probability. *IEEE Transactions on Human Factors in Electronics*, 1966, **7**, 22-28.

SUPPES, P., AND WINET, M. An axiomatization of utility based on the notion of utility differences. *Management Science*, 1955, **1**, 259-270.

THURSTONE, L. L., AND JONES, L. V. The rational origin for measuring subjective values. *Journal of the American Statistical Association*, 1957, **52**, 458-471.

TVERSKY, A. Additive choice structures. Unpublished doctoral dissertation, The University of Michigan, 1964.

TVERSKY, A. Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology* (in press).

TVERSKY, A. A general theory of polynomial conjoint measurement. *Journal of Mathematical Psychology*, 1967, **4**, 1-20.

TVERSKY, A., AND ZIVIAN, A. A computer program for additivity analysis. *Behavioral Science*, 1966, **11**, 78-79.

VON NEUMANN, J., AND MORGENSTERN, O. *Theory of games and economic behavior*. Princeton: Princeton Univer. Press, 1944.