

The Invention of the Independence Condition for Preferences

Peter Fishburn • Peter Wakker

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

Medical Decision Making Unit, University of Leiden (AZL), Leiden; The Netherlands

This paper discusses the history and interrelations of three central ideas in preference theory: the independence condition in decision under risk, the sure-thing principle in decision under uncertainty, and conjoint independence for multiattribute decisions and consumer theory. Independence was recognized as an important component of decision under risk in the late 1940s by Jacob Marschak, John Nash, Herman Rubin, and Norman Dalkey, and first appeared in publication in Marschak (1950) and Nash (1950). The sure-thing principle can be credited to Savage (1953, 1954). Conjoint independence for consumer theory was introduced by Sono (1943) and Leontief (1947a, b); a form of it can also be recognized in Samuelson (1947), presented earlier in Samuelson (1940).

Independence and the sure-thing principle are equivalent for decision under risk, but in a less elementary way than has sometimes been thought. The sure-thing principle for decision under uncertainty and conjoint independence are identical in a mathematical sense.

The mathematics underlying our three preference conditions has an older history. The independence condition for decision under risk can be recognized in the characterization of "associative means," and conjoint independence for multiattribute decisions in solutions to the "generalized associativity functional equation."

(*Expected Utility; Independence Condition; Conjoint Independence; History of Utility Theory*)

1. Introduction

The most important idea in the theory of decision under risk and uncertainty is the "independence condition" that requires separability of preferences across disjoint events. Its basic idea is illustrated by the following informal example for decision under uncertainty suggested by Savage (1954, p. 21). Suppose that a businessman would prefer buying a piece of property if the Republican candidate wins the next election, and also if the Democratic candidate wins. Then independence asserts that the businessman should buy the property even though he does not know who will win the election. At first sight, the condition seems compelling as a normative principle. Nevertheless, the famous Allais and Ellsberg paradoxes have called it into question. Independence has been staunchly defended by some and

severely criticized and modified by others, but it is seldom ignored.

Our discussion of independence begins in the context of decision under risk, then expands into decision under uncertainty and consumer preferences. We reconstruct the history of their introductions and explore relationships among them that will facilitate a deeper understanding of their origins.

For decision under risk, independence originated as a supplement to the utility theory in von Neumann and Morgenstern (1944, 1947, 1953; henceforth vNM¹). The condition was not mentioned explicitly in vNM, and its absence caused great confusion. This confusion

¹ All citations are taken from the third edition of 1953; this edition still did not mention independence explicitly.

was resolved by Malinvaud (1952). He observed that independence is automatically satisfied in the vNM formulation because von Neumann and Morgenstern applied probabilistic mixing to indifference classes rather than to single lotteries. Independence then follows from a standard mathematical principle of transferring operations from elements (lotteries) to equivalence classes (indifference classes of lotteries). This had been pointed out earlier, in Dalkey (1949), but his contribution went largely unnoticed.

Independence, in all its simplicity and subtlety, has been the most central idea in risk theory since mid-century. It underlies expected utility and the decision analysis technologies based on expected utility (Keeney and Raiffa 1976, Howard and Matheson 1984, Lau et al. 1983, von Winterfeldt and Edwards 1986, Clemen 1991). These technologies can be justified only if independence can be justified.

Opponents of expected utility often direct their criticism toward independence and base alternative models on weakenings of the independence condition. For example, the betweenness models in nonexpected utility, developed by Chew (1983) and Fishburn (1983), impose independence only within each indifference class; the rank-dependent models initiated by Quiggin (1982) and Schmeidler (1989, 1982) impose independence only within rank-ordered subsets. Variations on independence have also led to new approaches for dynamic decisions. Luce and Narens (1985), Machina (1989), McClennen (1990), Luce (1990), Segal (1990), Epstein (1992), Karni and Safra (1990), and Sarin and Wakker (1993a, b) split up independence into a few dynamic principles and then discard the principle that seems least tenable. The alternatives to expected utility have been developed primarily for descriptive rather than normative purposes because of widespread demonstrations of the fallibility of independence as a behavioral predictor. General surveys of nonexpected utility include Machina (1987), Fishburn (1988), and Camerer and Weber (1992). The normative status of nonexpected utility models that violate independence is the central theme in Edwards (1992). We will not describe in further detail the recent developments and variations of independence but will concentrate on its origins.

Our exposition of the history of independence begins in §2 with its absence from vNM. Sections 3 and 4 out-

line the history of independence for decision under risk in so far as we have been able to reconstruct it from published and unpublished materials, correspondence between principals at the time it arose, and conversations with those who are still with us. These sections concentrate on 1948–1952, a period whose history with regard to independence is not well known. Section 3 focuses on 1948–1950, when the independence condition for decision under risk came into being. Section 4 describes events in 1950–1952, when Paul Samuelson and L. J. (Jimmie) Savage entered the scene, and Savage developed his sure-thing principle that adapts independence to decision under uncertainty.

Because the independence condition for risk overlaps similar ideas from other contexts, including a separability condition from consumer demand theory, our account of the origin of independence requires consideration of those ideas and *their* origins. The discussion of interrelationships begins in §5, which describes the sure-thing principle for decision under uncertainty. It is shown that this principle does not imply expected-utility maximization unless additional conditions are assumed. Section 6 applies the sure-thing principle to decision under risk and demonstrates there that it is tantamount to independence and thus to expected-utility maximization.

Section 7 describes conjoint independence in consumer demand theory and multicriteria decisions. This version of independence is best known from the contributions of Debreu (1960) and Luce and Tukey (1964) but had important precursors in Samuelson (1940, 1947), Sono (1943), and Leontief (1947a, b). It has been called “preferential independence” (Keeney and Raiffa 1976) and “strong separability” and “coordinate independence” (Wakker 1989), among other things. We refer to this type of independence as “conjoint independence” because it concerns choice alternatives that consist of several coordinates, attributes, or components. Conjoint independence says that preference between two choice alternatives that have common values on some coordinates should not depend on the levels of those common values. We relate conjoint independence to independence conditions of previous sections and sketch its history. As we shall see, conjoint independence is the oldest of our three versions and encompasses the others as special cases.

The mathematics underlying the above results had been developed earlier in characterizations of the "generalized associative functional equation" and of "associative generalized means." For instance, the expected-utility result of von Neumann and Morgenstern (1944) can already be recognized in Nagumo (1930) and Kolmogorov (1930). We discuss the mathematical background in the appendix.

2. The Absence of Independence in vNM's Utility Analysis

Decision science terminology had not yet been standardized when von Neumann and Morgenstern presented their utility theory. In the terminology of vNM (3.3.3) "events . . . were used . . . as the substratum of preferences," i.e., events are what we call consequences or lotteries. Utilities in the vNM sense are not real numbers but have a one-to-one correspondence with indifference classes of lotteries. When utilities in the vNM sense are identified with indifference classes, the vNM theory can be readily understood from the modern perspective. Several authors have discussed whether a vNM numerical utility should be interpreted as "the" cardinal utility, long searched for by economists, that measures strength of preference, or whether it should be considered a new type of utility. We avoid this issue here but note recent discussions in Fishburn (1989) and Wakker (1994, §2).

Part of the intuition of independence is found in vNM, p. 18, lines 3 and 4, where "alternatives" designate lotteries P and Q in a probabilistic mixture $\lambda P + (1 - \lambda)Q$:

We stress that the two alternatives are mutually exclusive, so that no possibility of complementarity and the like exists

Independence can indeed be interpreted as the exclusion of any effect of complementarity, or its opposite, substitutability.² The crucial lines are at the beginning of vNM §3.4.2, where independence is introduced implicitly by transferring the probability mixture operation

from lotteries ("events") to indifference classes of lotteries ("utilities"):

. . . We have assumed only one thing—and for this there is good empirical evidence—namely that imagined events can be combined with probabilities. And therefore the same must be assumed for the utilities attached to them,—whatever they may be

Von Neumann and Morgenstern compare this to well-established physical quantities. For example, mass can be endowed with an addition operation corresponding to a physical joining of objects. A necessary prerequisite for the meaningfulness of addition for masses is that $\text{mass}(A\text{-joined-with-}B)$, which is to be equal to $\text{mass}(A) + \text{mass}(B)$, does not change if object A is replaced by an object A' that has the same mass as A . This is not trivial or a priori true, but has been established empirically and is beyond dispute. However, things are quite different for the mixture operation on utilities. To explain this, we denote indifference classes of lotteries by square brackets so that $[P]$ is the class of all lotteries indifferent to lottery P . For the mixture operation on indifference classes of lotteries to be meaningful, it is necessary that $[\lambda P + (1 - \lambda)Q]$, which is to be equal to $\lambda[P] + (1 - \lambda)[Q]$, does not change if lottery P is replaced by a lottery P' that is from the same indifference class. In other words, for all lotteries P, P', Q , and $0 < \lambda < 1$,

$$P \sim P' \Rightarrow \lambda P + (1 - \lambda)Q \sim \lambda P' + (1 - \lambda)Q, \quad (2.1)$$

where \sim denotes indifference. This substitution principle is one version of the independence condition. It lacks the empirical status of the analogous condition for addition of masses as current literature in the field of risk amply demonstrates. Related versions of independence are, for $0 < \lambda < 1$,

$$\begin{aligned} P \succcurlyeq P' &\Rightarrow \lambda P + (1 - \lambda)Q \\ &\succcurlyeq \lambda P' + (1 - \lambda)Q, \quad \text{and} \end{aligned} \quad (2.2)$$

$$P \succ P' \Rightarrow \lambda P + (1 - \lambda)Q \succ \lambda P' + (1 - \lambda)Q, \quad (2.3)$$

where \succcurlyeq denotes preference-or-indifference and \succ is the strict preference relation. Our use of *independence* for decision under risk refers to any or all of (2.1)–(2.3), which are equivalent to one another in the presence of suitable ordering and continuity conditions.

² Complementarity occurs, for instance, if the value of a number of left shoes increases (up to a point) as the number of right shoes increases; substitutability occurs if the value of a number of apples decreases as the number of pears increases

However, in the present section and the next, "independence" refers primarily to (2.1).

The substitution condition for mass and condition (2.1) for indifference are special instances of a general mathematical principle. The principle refers to a set of objects endowed with an operation and an equivalence relation. The operation can be transferred from objects to equivalence classes only if a compatibility condition of the following sort is satisfied: if the operation is applied to a number of elements, then replacing any element by an equivalent one should not affect the equivalence class resulting from the operation. In risk theory the objects are lotteries, the operation is probabilistic mixing, and the equivalence relation is indifference. Independence condition (2.1) is the described mathematical principle for our particular application.

In the preface to the third edition of 1953, vNM points out on page viii:

In particular our discussion and selection of "natural operations" in those sections covers what seems to us the relevant substrate of the Samuelson-Malinvaud "independence axiom."

Part of the idea of independence reappears in Conditions 3:B:a* and 3:B:b* of vNM where "quasi-concavity" and "quasi-convexity" of preferences with respect to probability mixtures are imposed. These conditions are weaker than independence but convey some of its intuition. vNM justifies the two conditions by reference to the exclusion of complementarity and substitutability, as cited above.

We emphasize that the analysis of vNM is mathematically correct. However, as recognized in Marschak (1950, end of 44.1), the choice of indifference classes ("utilities") as primitive in their axiomatic system is unfortunate because lotteries, and not indifference classes of lotteries, are the natural empirical primitives.

3. Independence for Expected Utility: The Early Years

The independence condition for preference in decision under risk emerged at the end of the 1940s, when people became aware that independence was a necessary consequence of expected-utility maximization. Arrow (1991a) writes:

My understanding was derived from various discussions in the corridors of Rand (including Dalkey) and also with Herman Rubin and Jimmie Savage.

Samuelson (1952, footnote on p. 673) mentions "Marschak, Nash, Dalkey, and others" in connection with independence, and Arrow (1991b) assigns a considerable role to Rubin and Chernoff. We describe contributions of Marschak, Nash, Rubin, Dalkey, and Chernoff in this section, and of Samuelson, Savage, and Malinvaud in the next.

The first *published* statements of the independence condition for lotteries appeared in Marschak (1950, Postulate IV) and Nash (1950) in the same issue of *Econometrica*. Both used (2.1), and Nash extended this to include general substitution of indifferent lotteries in preference statements. Marschak's paper neither references another source for the independence idea nor lays claim to its originality. This suggests that the condition's necessity was common knowledge among Marschak and his colleagues. Moreover, it could hardly have been foreseen that this supplement to the theory in vNM would play such a central role in the future of decision under risk. Marschak (1950, Part VII) came close to understanding how vNM implicitly subsumed independence, but did not pin it down precisely. Marschak (1951) also discussed independence and showed its necessity for expected-utility maximization.

Nash (1950), a classic in game theory, presented the independence condition (item 5, p. 156) as a general substitution principle that includes (2.1). Like Marschak, Nash does not refer to another source for independence and does not claim novelty. He announces without proof the use of a numerical utility "of the type developed in *Theory of Games*" and does not mention the absence of independence in vNM.

Rubin (1949a, p. 1051, 1. 6/7; 1949b, Axiom IV) introduced independence condition (2.2) in the stronger "if and only if" form. These two papers had a different and more premature status than the papers of Marschak and Nash in that Rubin (1949a) was an abstract of a lecture and Rubin (1949b) was a working paper. It should be noted that the first version of Marschak's paper appeared as a working paper in 1948 and that Marschak (1949) is an abstract of his work; see Footnote 1 in Marschak (1950). The published version of Marschak's paper refers to Rubin (1949b) as an alternative approach. After nearly 40 years, Rubin published an elaborated and generalized version of his earlier ideas (Rubin 1987). In this paper the independence axiom is

reformulated for choice functions rather than for preference relations as in his earlier two papers. Remarkably, Rubin not only considered probability distributions over consequences, but also incorporated "states of nature," as in the basic model of §5 below. Rubin's approach considers "acts" as functions that assign to each state of nature a probability distribution over consequences and derives a predecessor of the well-known theorem of Anscombe and Aumann (1963).

Dalkey (1949) may have been the first to understand how independence was subsumed in the analysis of vNM. We give two citations from Dalkey's page 3:

Hence, the utilities of TG³ must be abstract entities. A possible interpretation—adopted in the following—is that utilities are the indifference classes of objects or commodities

Strictly speaking, this interpretation of utilities would require further axiomatization, in particular, it would require an axiom to the effect that if two different objects A , B have the same utility, then the options $[A, C, \alpha]$, $[B, C, \alpha]$ with a third object C , also have the same utility (belong to the same indifference class).

These citations clearly explain "utility" as used in vNM, make independence explicit, and show how it is involved behind the scenes in the vNM formulation. On page 4 Dalkey cites the argument of vNM that two "possibilities" (lotteries) in an option (mixed lottery) are exclusive and that there could be no complementarity between them. This is the argument for independence advanced later by Samuelson (1952). Marschak (1951, §3.5) also invoked the argument.

We mention here the work of Chernoff although it deals with decision under uncertainty, the topic of §5. Chernoff (1954, first version 1949; see also Chernoff 1950) assumed that utility, to be maximized according to the expectation principle, was given in advance. He then derived maximization of expected utility for equally likely states for the decision-under-uncertainty context. The verbal description of his Postulate 8 (for which he credits Rubin) is close to independence. Its formal description, however, mixes consequences instead of probabilities and concerns maximization of expected value. Chernoff's strong assumption that utilities are

known in advance obliterates the difference between the verbal and formal description of his Postulate 8.

4. Independence for Expected Utility: Samuelson and Savage

Samuelson and Savage made important contributions to the idea of independence but entered the scene a bit later. Samuelson (1991) credited Marschak for being the first to point out to him the independence axiom. This occurred in 1950. Samuelson himself proposed the phrase "strong independence," probably in analogy to a similar condition that leads to additive utilities in consumer demand theory (Samuelson 1947, p. 174–180). The latter condition was stated in terms of a representing function rather than in terms of preferences. A preference-based version will be discussed in §7.

In an ensuing correspondence with Savage in 1950, Samuelson convinced Savage that the independence condition was needed in full strength to imply expected-utility maximization and that, contrary to a claim in Friedman and Savage (1948), quasi-concavity and quasi-convexity of preferences alone was insufficient. Savage (1951, footnote 3 on p. 57) credits Samuelson for this. In return, Savage convinced Samuelson, at first by means of a book-making argument, that expected-utility maximization and the independence axiom were normatively appealing. Samuelson credits Savage for this in several places, including the *Postscript* to Samuelson (1950) as added in Stiglitz (1966, Chapter 12). Samuelson subsequently found an argument for independence based on mutual exclusiveness of disjoint events that he considered more convincing than the book-making argument. This took place in the summer of 1950. We cite from Samuelson (1952, p. 672):⁵

Prior to 1950, I hesitated to go much farther. But much brooding over the magic words "mutually-exclusive" convinced me that there was much to be said for a further "strong independence axiom" . . . This is simply a version of what Dr. Savage calls the "sure-thing principle."

As mentioned above, the "mutually-exclusive" qualifier had been used previously in vNM. Shortly thereafter, and certainly by 1952, Savage transformed the inde-

³ That is, our vNM: TG = "Theory of Games."

⁴ $[A, C, \alpha]$ denotes " A with probability α , C with probability $1 - \alpha$."

⁵ See also Samuelson (1953, p. 130).

pendence condition into its purest and most salient form as the "sure-thing principle."

The contribution of Samuelson first appeared in Samuelson (1952, 1953), and that of Savage⁶ in Weyl and Bell (1952, Condition 2 on p. 3), Savage (1953, Axiom 5), and partly in Friedman and Savage (1952, p. 469). Both Weyl and Bell (1952) and Savage (1953) are based on Savage's presentation at the international Colloquium on the Foundations and Applications of the Theory of Risk, held from May 12 to May 17, 1952, in Paris.

It appears that Savage was not influenced by the logically equivalent condition of conjoint independence or separability in consumer demand theory, a condition with an older history as discussed below in §7. Savage credits de Finetti as influential; Postulate 4 of §13 of de Finetti (1931a) introduces an additivity axiom for qualitative probability that is identical to the sure-thing principle when there are only two consequences.

The paper of those days that has drawn the most attention is Samuelson's (1952) contribution in the *Econometrica* discussion on independence, which was actually written after Samuelson (1953). Several authors have subsequently and incorrectly ascribed the invention of independence to Samuelson (1952). One reason may have been that the editorial note in *Econometrica* (1952, p. 661) attributed the condition to Samuelson, and Samuelson (1952) himself described independence as arising from discussions with others without explicitly referring to the 1950 papers of Marschak and Nash. This applies also to Samuelson (1953), which along with Samuelson (1952) named Dalkey, Marschak, and Nash. Perhaps their papers were not referenced in Samuelson's *Econometrica* discussion because Manne (1952) had already cited Marschak (1950). Savage (1954) also did not refer to Marschak or Nash, but Arrow (1951, p. 424) cited Marschak (1950) as a "simplified treatment" of the vNM utility theory. Samuelson (1977, p. 48) writes:

Savage and Marschak, better than von Neumann-Morgenstern or Menger, address themselves to these questions along lines

⁶ Savage (1951) says nothing about independence or the sure-thing principle. On p. 58 he describes "monotonicity." This is further discussed at the end of our §7.

implicit in Ramsey [. . . , 1931] and involving what are called various "independence axioms" in Samuelson [. . . , 1952].

Ellsberg (1954, p. 544) says that the independence axiom is Samuelson's invention, although on page 538 he refers also to Marschak (1950).

While Marschak, Samuelson, and others suspected that von Neumann and Morgenstern avoided the independence condition because it was implicit in their formulation, it seems that this became generally understood only after the presentation of Samuelson (1953) at the 1952 Paris colloquium. As we noted above, independence was implied when vNM transferred the mixture operation from lotteries to indifference classes of lotteries. While already recognized by Dalkey (1949), Malinvaud (1952) is the first published work to make this explicit.

The history after 1952 is well known. Herstein and Milnor (1953) give an efficient axiomatization of expected utility which assumes independence condition (2.1) only for $\lambda = \frac{1}{2}$. Savage (1954, first five chapters) used the sure-thing principle to obtain his path-breaking axiomatization of subjective expected utility maximization. In doing this, Savage generalized the results of von Neumann and Morgenstern and of Marschak by not assuming that probabilities were given primitives. Jensen (1967) presented the version of vNM's theorem that is most popular today. It uses the strict-preference independence condition (2.3) along with an appealing continuity axiom. Later refinements and generalizations are described in Fishburn (1982).

The 1952 conference in Paris also gave birth to Maurice Allais' famous examples that criticize the independence condition. Although this criticism received little attention at first, it was pursued in Allais and Hagen (1979), Kahneman and Tversky (1979), and Machina (1982) and is now a main theme in risk theory. One version of an Allais paradox for preferences is

$$M \succ (0.98, 5M) \quad \text{and} \quad (0.0098, 5M) \succ (0.01, M),$$

where M denotes 100 million dollars and (λ, X) is the lottery that pays X with probability λ and \$0 otherwise. The given preferences violate (2.3) by the substitutions $P = M, P' = (0.98, 5M), Q = \0 , and $\lambda = 0.01$. According to Allais, most people express the given preferences.

5. Savage's Sure-thing Principle for Decision Under Uncertainty

This and subsequent sections discuss conditions related to independence for risk as described by (2.1)–(2.3). This section studies Savage's sure-thing principle in the general context of decision under uncertainty, where it is not sufficiently strong to imply expected-utility maximization. The next section describes an equivalence between independence and the sure-thing principle for decision under risk, and §7 brings conjoint independence into the picture.

In *decision under uncertainty (DUU)*, uncertainty is not described initially by probabilities, but by a *state space* S . For example, if the uncertainty concerns which of three participating horses will win a race and s_i denotes the state that horse i wins, then $S = \{s_1, s_2, s_3\}$. Exactly one state is true, the other states are not true, and the decision maker is uncertain about which state is true. Subsets of S are called *events*. In the example, $\{s_1, s_2\}$ is the event that either the first or the second horse wins.

Let \mathcal{F} denote the set of all decision alternatives, called *acts*, that the decision maker could face, and let X be the set of *consequences*, i.e., possible outcomes of decisions after the uncertainty about the true state has been resolved. Formally, an act f in \mathcal{F} is a function from S to X , specifying for each state s the consequence $f(s)$ that results if act f is chosen and state s is the true state. For simplicity we assume that consequences in X are monetary amounts in an interval $[0, M^*]$ and refer to acts as *gambles*. The term *lottery* will continue to be used as a probability distribution over consequences.

We denote by \succsim the decision maker's binary *is preferred or indifferent to* relation on \mathcal{F} . Throughout we assume that \succsim is complete ($f \succsim g$ or $g \succsim f$ for all gambles f, g) and transitive. Standard continuity conditions for \succsim are also presumed; they need not be made explicit in this analysis. We say that a function $V : \mathcal{F} \rightarrow \mathbb{R}$ *represents* the preference relation \succsim if for all f and g in \mathcal{F} ,

$$f \succsim g \iff V(f) \geq V(g).$$

Hence more-preferred gambles have greater V values.

The idea behind Savage's sure-thing principle is that a choice between gambles f and g should depend only on states for which $f(s) \neq g(s)$: states for which the

gambles yield identical consequences should not play a role in the decision. Suppose in the horse-race example that the decision maker must choose between gambles f and g , where for the first horse both gambles yield \$3; see Figure 1. Because it does not matter which choice is made if the first horse wins, the sure-thing principle says that the choice should depend only on the payments for the second and third horses. Suppose the decision-maker prefers f over g . If the \$3 payment for s_1 is changed to \$4 for both f and g , but the s_2 and s_3 payments are unchanged, then the choice should not change. This means that $f' \succsim g'$ if $f \succsim g$ in Figure 1.

Formally, the *sure-thing principle* says that a preference between two gambles is independent of the states in which the two yield identical consequences. So it requires $f \succsim g \iff f' \succsim g'$ whenever S can be partitioned into two parts— I (the “irrelevant” event) and R (the “relevant” event)—such that, on I , $f = g$ and $f' = g'$, and on R , $f = f'$ and $g = g'$. When probabilities are assigned to states, expected-utility maximization implies the sure-thing principle because the parts of the expected utilities of f and g or of f' and g' related to event I cancel.

An Allais-type example that challenges the sure-thing principle is illustrated by Figure 2. We have taken the liberty here of associating states with probabilities in a particular alignment and recasting an Allais paradox (preceding section) in Savage's format. This mimics the way that Savage himself (Savage 1954, p. 103) viewed Allais' paradox. However, this is justified only if one adopts an axiom like the DUR-Assumption described in the next section. In the example, M denotes 100 million dollars, and it is assumed that state probabilities are known. Many people prefer f to g , but when the

Figure 1 The Sure-thing Principle. If $f \succsim g$ then $f' \succsim g'$.

	I	R	R
	s_1	s_2	s_3
f	3	0	7
g	3	1	5
f'	4	0	7
g'	4	1	5

Figure 2 Allais Paradox

state probabilities:	0.89	0.10	0.01
	I	R	R
	s_1	s_2	s_3
f	0	5M	0
g	0	1M	1M
f'	1M	5M	0
g'	1M	1M	1M

common consequence under s_1 is changed from 0 to 1M, most of them prefer g' to f' , because g' gives 1M with certainty. Thus many people have preferences that violate the sure-thing principle.

We observe next that the sure-thing principle does not imply expected-utility maximization, so it certainly does not imply the independence condition when probabilities are given. To illustrate this, and to characterize the strength of the sure-thing principle, consider a partition of S into three or more mutually exclusive and exhaustive events E_1, \dots, E_n . We denote partition $\{E_1, \dots, E_n\}$ by π . Suppose that all E_j are possible in the sense that each can influence preferences. We restrict attention to the subset of gambles that are constant on each E_j . We denote this subset by \mathcal{F}^* and for gamble f in \mathcal{F}^* let f_j be the consequence that results if E_j contains the true state. Under standard continuity assumptions, the sure-thing principle is equivalent to the existence of an additive representation, so there exist functions V_1, \dots, V_n on consequences such that \succsim is represented by $V(f) = V_1(f_1) + \dots + V_n(f_n)$. This was first proved by Debreu (1960, Theorem 3). The principle's necessity is easily derived from substitution, but the sufficiency derivation is deep and complicated. The expected-utility form is a special case in which $V_j = P(E_j)U$ for all j , where U denotes utility and $P(E_j)$ is the probability of event E_j . The additive model is more general because, even if we set $V_j(0) = 0$ for all j (which is always allowed), the V_j need not be proportional. The state-dependent expected-utility model (Karni 1985), in which $V_j = P(E_j)U_j$, for each j , is another example of the additive model that satisfies the sure-thing principle.

Although we have explained the matter for gambles that depend only on events E_1, \dots, E_n , the preceding results can be extended to all acts by a procedure described in Chew and Wakker (1991).

Additional conditions are clearly needed to obtain expected utility from the additive representation. The exact nature of these conditions depends on the continuity conditions that are used. Such conditions are found in Wakker (1984, 1989, Theorem IV.2.7), Chew (1989), Nakamura (1990), Fishburn (1990), and Gul (1992) as well as in Savage (1954). We summarize the above findings for DUU.

OBSERVATION 5.1. *Under standard ordering and continuity assumptions, the sure-thing principle is equivalent to the existence of an additive representation that does not imply expected-utility maximization. Further conditions must be added to characterize expected utility.* \square

As an aside on terminology, we remark that the definition of the sure-thing principle in our paper is the version used most often today. It amounts to Savage's P2. Savage (1954, §2.7) also included his P3 as part of his own designation of the sure-thing principle.⁷

6. Risk and Uncertainty: An Equivalence Theorem

This section applies the sure-thing principle to *decision under risk* (DUR) and shows that it is equivalent to independence in that context. We note first how DUR can be considered to be a special case of DUU.

DUR concerns the special case in which a probability measure μ on the state space S is given a priori. Given μ on S , each gamble generates a lottery over the consequence set X . An example appears in Figure 2 of the previous section, where probabilities 0.89, 0.10, and 0.01 are given for s_1, s_2 , and s_3 , respectively. The lotteries thus generated are, in probability / payoff form,

$$\begin{aligned} & (0.10, 5M; 0.90, 0) \quad \text{for } f; \\ & (0.11, 1M; 0.89, 0) \quad \text{for } g; \\ & (0.10, 5M; 0.89, 1M; 0.01, 0) \quad \text{for } f'; \\ & \quad \quad \quad 1M \quad \text{for } g'. \end{aligned}$$

⁷ See the first footnote of §2.7 in the 1972 edition of Savage (1954) for a minor mathematical correction of the original definition.

It is assumed in DUR that the only preference-relevant aspect of a gamble is the lottery it generates on consequences. We shall assume also that all conceivable lotteries over consequences can be obtained from gambles. This assumption is customary in DUR and is guaranteed by the absence of "atoms," where an *atom* is a positive-probability event that includes no subevent with positive but strictly smaller probability.⁸

DUR-ASSUMPTION 6.1. *If gambles f and g generate the same lottery over X , then $f \sim g$. Moreover, S contains no atoms.* \square

Under the DUR-Assumption, a preference relation over lotteries is obtained in the obvious way from preferences between gambles. Because the only relevant aspect of a gamble is the lottery it generates, gambles are usually described as lotteries in DUR, and the underlying state space S is suppressed. Under the DUR-Assumption plus typical ordering and continuity conditions, independence implies expected-utility maximization. Because the subjective expected-utility model for DUU implies the sure-thing principle, it follows that independence in the DUR reduction implies the sure-thing principle.

We have noted a difference between the assumption of expected-utility maximization and the sure-thing principle in DUU because the latter does not imply the former. However, this difference is obliterated under DUR. To see this, consider the S -partition $\pi = \{E_1, \dots, E_n\}$ and suppose that $\mu(E_1) = \dots = \mu(E_n) = 1/n$. The DUR-Assumption then implies that a gamble from \mathcal{F}^* remains in the same indifference class if its consequences for the E_i 's are permuted. Given an additive representation as implied by the sure-thing principle, it follows that all V_j 's must be identical up to shifts in origins. Hence the expected-utility model results if one sets $U = nV_i$ for an arbitrary i along with $\mu(E_j) = 1/n$ for all j , so $V_j = U/n$ for all j . This yields an expected-utility representation for each \mathcal{F}^* with equally probable events. The usual continuity conditions then imply that an expected-utility representation must exist on the entire set

\mathcal{F} of gambles, so independence holds for all lotteries. We summarize as follows.

OBSERVATION 6.2. *Under the DUR-Assumption, independence and the sure-thing principle are equivalent.* \square

Because our derivation of [sure-thing principle \Rightarrow independence] for DUR invoked equal-probability partitions and continuity, it is not as elementary as has sometimes been thought. For historical completeness and without elaboration, we mention that the heuristic derivation of independence from a sure-thing-like principle in Friedman and Savage (1952, p. 469) is not entirely rigorous.⁹

7. Independence for Consumer Theory and Its History

This section compares the independence condition for DUR with the independence condition of consumer demand theory, multiattribute utility theory, and conjoint measurement that we refer to as conjoint independence. We observe that conjoint independence can be identified with the sure-thing principle.

It should be recalled that independence in our paper designates a property of preference relations that is formally different from the older independence condition for utilities that excludes complementarity and substitutability effects. Debreu (1960), however, showed that independence for preferences is equivalent to independence for utilities when a few other assumptions are granted.

We define conjoint independence for preferences over n -tuples (x_1, \dots, x_n) in the product set $\times_{j=1}^n X_j$. *Conjoint independence* requires that, for all $x, y \in \times_{j=1}^n X_j$, $x \succsim y$ is independent of common coordinate values of x and y . In other words, it requires that $x \succsim y \Leftrightarrow x' \succsim y'$ whenever $\{1, \dots, n\}$ can be partitioned into two parts I and R such that, on I , $x = y$ and $x' = y'$, and on R , $x = x'$ and $y = y'$. An example for \succsim on \mathbb{R}^4 is given in Figure 3. Here $I = \{2, 4\}$ and $R = \{1, 3\}$.

To relate conjoint independence to the sure-thing principle, we note that if all X_j 's are identical to the

⁸ We omit the proof that every probability distribution over X is then generated by at least one gamble. Also, to avoid technical details, we do not discuss continuity conditions and measure-theoretic aspects that can be dealt with by imposing a σ -algebra on S .

⁹ One complication in the Friedman/Savage analysis is that gambles-conditional-on-events cannot be readily identified with nonconditional gambles.

Figure 3 Conjoint Independence. If $x \geq y$ then $x' \geq y'$.

	<i>R</i>	<i>I</i>	<i>R</i>	<i>I</i>
<i>x</i>	3	1	5	7
<i>y</i>	4	1	6	7
<i>x'</i>	3	2	5	8
<i>y'</i>	4	2	6	8

same set X , then $\times_{j=1}^n X_j$ can be identified with the set of functions from $\{1, \dots, n\}$ to X . When $S = \{1, \dots, n\}$, conjoint independence then reduces to the sure-thing principle.¹⁰ Similarly, conjoint independence can be extended to infinite product sets for which $\{1, \dots, n\}$ is replaced by an infinite index set, in which case the sure-thing principle for infinite state spaces can be related to conjoint independence. We conclude that the sure-thing principle is equivalent to conjoint independence when all X_j sets are identical.

The history of conjoint independence is tied to contributions of Sono, Leontief, Samuelson, and Irving Fisher. For simplicity, we assume throughout that $n \geq 3$. Near the end of the preceding century it was customary to evaluate an n -tuple of goods (x_1, \dots, x_n) by an additive utility function, say $V_1(x_1) + \dots + V_n(x_n)$. However, it was also recognized (Fisher 1892) that effects of complementarity and substitutability could preclude additive representability. Such effects were expressed in terms of a holistic representing function $V(x_1, x_2, \dots, x_n)$, often through cross-derivatives, in which the value of x_i can depend on coordinate values x_j for $j \neq i$. Slutsky (1915), Allen (1934), and Friedman (1935) also noted restrictions on preferences that are implied by an additive representation.

Sono (1943) seems to be the first publication to identify a version of conjoint independence related to the family of conditions studied in our paper. However, this paper was written in Japanese during the second world war and became widely known only after it was translated into English as Sono (1961). The paper proposed a definition of "separability" to the effect that a group of commodities $A \subset \{1, \dots, n\}$ is *separable* if the preference relation induced over $\times_{j \in A} X_j$ by keeping the

non- A coordinates at fixed levels is independent of those levels. Obviously, conjoint independence holds if and only if every subset of $\{1, \dots, n\}$ is separable. We remark that the requirement of separability of all subsets of $\{1, \dots, n\}$ is implied by seemingly weaker separability assumptions. For example, it suffices for all two-element subsets to be separable. This induces additive representability and thus conjoint independence. An even more general result is proved in Gorman (1968), with later contributions in von Stengel (1993).

Sono did not pay much attention to the strong condition of separability of all subsets. In particular, he did not derive the additive form. Independently, Leontief (1947a, b) also introduced the idea of separability. He used the phrase *strong separability* for the condition that all subsets are separable (conjoint independence) and showed that this implies the existence of an additive representation. The papers of Sono and Leontief formulated conditions in terms of representing functions and used differentiability methods with each X_j a real interval. A typical way to formulate separability of, say, commodities 1 and 2 in this mode is to require that, for each commodity bundle (x_1, \dots, x_n) , the amount δ to be gained for commodity 2 in order to make up for a loss of ϵ of commodity 1¹¹ should be independent of x_3, \dots, x_n .

Samuelson independently arrived at the same result as Leontief—that conjoint independence implies additive representability for $n \geq 3$ (Samuelson 1947, p. 174–180). His formula (32) is equivalent to the requirement that the rate of substitution between two goods be independent of the other goods. After his formula (33), Samuelson pointed out that this condition is necessary and sufficient for additive representability.¹² It should be noted that Samuelson's work, while published in 1947, was written as a Ph.D. dissertation in 1940.

Separability of single commodities (sometimes called *weak separability*) was used previously in Fisher (1927,

¹¹ That is $(x_1 - \epsilon, x_2 + \delta, x_3, \dots, x_n) \sim (x_1, x_2, x_3, \dots, x_n)$.

¹² Samuelson's term "independence (of utility)," in this context, means additive representability, which is to be distinguished from the independence condition for preferences in our paper. The interesting footnote 95 of Allais (1953), §3.5 in Marschak (1951), as well as most writings of that time, also use the term independence to designate the existence of an additive representation.

¹⁰ Further substitutions are f, f', g, g' for x, x', y, y' .

p. 175 ff). This condition is, however, too weak to be considered a version of separability or conjoint independence. Only when separability is imposed on subsets of two or more commodities and there are other commodities not in the subset do we obtain tradeoffs between commodities that are strong enough to qualify as a version of conjoint independence. Consequently, conjoint independence obtains its full strength only when there are three or more commodities, attributes, or coordinates, as presumed above.

As explained shortly, conjoint independence can be viewed as a type of monotonicity condition, a condition that has been around for centuries. Some people have argued that monotonicity should be considered as an early form of independence, but we think this is inappropriate. The rest of this section explains how we distinguish independence from monotonicity.

Suppose that a weak order \succsim_j is given on each coordinate set X_j . Then *monotonicity* requires $x \succsim y$ when $x_j \succsim_j y_j$ for all j , with $x \succ y$ if also $x_j \succ_j y_j$ for at least one j . Weak separability (separability of single commodities) can be reformulated as the requirement that on each X_j a weak order \succsim_j exists such that \succsim satisfies monotonicity with respect to these \succsim_j 's. For example, define \succsim_1 by $x_1 \succsim_1 y_1$ if there exist z_2, \dots, z_n such that $(x_1, z_2, \dots, z_n) \succsim (y_1, z_2, \dots, z_n)$. By weak separability, this does not depend on the particular z_j values that were chosen. Other \succsim_j are defined similarly. Repeated applications of weak separability and transitivity then imply monotonicity. The thing to note for weak separability is that the \succsim_j relations are not given beforehand, but are derived from \succsim .

Conjoint independence can be reformulated as the condition that, for each partition of $\{1, \dots, n\}$ into m "blocks" A_1, \dots, A_m , there exist weak orders \succsim_{A_i} on $\times_{j \in A_i} X_j$ such that \succsim satisfies monotonicity with respect to these orders. To illustrate "block-monotonicity," suppose $n = 3$ and $X_1 = X_2 = X_3 = \mathbb{R}$, so that \succsim is defined on \mathbb{R}^3 . Assume that we observe $(7, 4, 6) \succsim (7, 3, 6)$ and $(1, 5, 1) \succsim (0, 5, 2)$. Then we can conclude from conjoint independence that $(1, 4, 1) \succsim (0, 3, 2)$: the first observed preference $(7, 4, 6) \succsim (7, 3, 6)$ and conjoint independence imply $(1, 4, 1) \succsim (1, 3, 1)$; the second observed preference $(1, 5, 1) \succsim (0, 5, 2)$ and conjoint independence imply $(1, 3, 1) \succsim (0, 3, 2)$. Transitivity then gives $(1, 4, 1) \succsim (0, 3, 2)$. The argu-

ment becomes more transparent if we recognize block monotonicity in the following manner. Divide $\{1, 2, 3\}$ up into the blocks $A_1 = \{2\}$ and $A_2 = \{1, 3\}$; then $(7, 4, 6) \succsim (7, 3, 6)$ implies $4 \succsim_{A_1} 3$, $(1, 5, 1) \succsim (0, 5, 2)$ implies $(1, 1) \succsim_{A_2} (0, 2)$, and the conclusion $(1, 4, 1) \succsim (0, 3, 2)$ is a consequence of monotonicity with $4 \succsim_{A_1} 3$ and $(1, 1) \succsim_{A_2} (0, 2)$. The differences between conjoint independence and monotonicity are first, that conjoint independence imposes monotonicity with respect to each partition of $\{1, \dots, n\}$ into blocks, and second, that the weak orders \succsim_{A_i} are not given a priori.

We now relate independence for DUR to monotonicity, using the version of independence in (2.3). Our product set is $X_1 \times X_2$, where $X_1 = X_2 = X$ and X is the set of all lotteries. We interpret (P, Q) in $X_1 \times X_2$ as lottery $\lambda P + (1 - \lambda)Q$, where $0 < \lambda < 1$ is fixed and is as in (2.3). This means that $X_1 \times X_2$ is the same set of lotteries as X_1 , X_2 , and X . By taking the orderings \succsim_1 and \succsim_2 on X_1 and X_2 as identical to the preference relation \succsim over lotteries, the implication

$$P \succsim P' \Rightarrow \lambda P + (1 - \lambda)Q \succsim \lambda P' + (1 - \lambda)Q$$

of (2.3) can be interpreted as monotonicity! This may explain the exceptional intuitive appeal of independence for DUR. Its interpretation as a monotonicity condition is, however, very special and subtle. There is actually a monotonicity condition for each λ in $(0, 1)$ with many structural symmetries, and the orderings \succsim_1 and \succsim_2 are not given beforehand but are identical to the preference relation \succsim .

In view of the preceding results, we recall that some authors cite Savage (1951, p. 58) for the sure-thing principle (conjoint independence for DUU). However, only monotonicity is described there.¹³ A similar remark holds for Axiom 2a in Ramsey (1931). Further comments on the recent history of conjoint independence and related conditions appear in Wakker (1989, §II.5).

8. Conclusion

This paper has discussed three conditions: independence in DUR, the sure-thing principle in DUU, and conjoint independence in consumer theory. The sure-thing prin-

¹³ Savage (1954) does discuss the sure-thing principle quite explicitly.

ciple is mathematically identical to conjoint independence. Our basic form of independence is defined only for DUR, and there it is equivalent to the sure-thing principle.

The independence condition as it is known today is due in large part to Marschak (1950). The vNM formulation did not invoke independence because it used indifference classes of lotteries as primitives and transferred the probability mixture operation to indifference classes. Dalkey (1949) and Malinvaud (1952) observed that this automatically presumes independence. Savage developed his sure-thing principle between 1950 and 1952 as a derivative of independence. He was probably not influenced by conjoint independence in consumer demand theory that was then available in the literature through Leontief (1947a, b) and could also be recognized in Samuelson (1940, formula (32); 1947, formula (32)). Because of the relatedness of independence with conjoint independence, Samuelson proposed the term *independence* for what has proved to be such an important condition in DUR.

There are earlier results in mathematics that embody the independence idea. In particular, Nagumo (1930) and Kolmogorov (1930) may have been the first to work out the *mathematics* of vNM expected utility. We describe this in the appendix.

We conclude with a tongue-in-cheek citation from Friedman and Savage (1952, p. 468) which suggests that independence can be traced back to the ancient Greeks:

... a principle that we believe practically unique among maxims for wise action in the face of uncertainty, in the strength of its intuitive appeal. This principle is universally known and recognized; and the Greeks must surely have had a name for it, though current English seems not to.

Acknowledgements

This paper received helpful comments from Norman Dalkey, Gerard Debreu, Jacques Drèze, Ward Edwards, Wassily Leontief, Duncan Luce, Mark Machina, Tom Marschak, John Pratt, Patrick Suppes, and Karl Vind, and could not have been written without the invaluable help of Kenneth Arrow, Chew Soo Hong, Milton Friedman, Paul Samuelson, and Lloyd Shapley, who shared with us their archives and memories.

Appendix

Related Mathematical Results

It is probably impossible to attribute the invention of independence to one person at one point in time. We have observed that Samuelson,

Sono, Leontief, Marschak, and Nash discovered related ideas in preference theory, and we noted that it is unclear precisely how these discoveries were interrelated or influenced by other people such as Rubin and Dalkey. Section 2 explained that independence can be viewed as a special case of a standard mathematical principle which transfers operations from individual elements to equivalence classes. We now describe other results in mathematics that are closely related to independence in preference theory but have an older history. Unlike preference theory, they are based on real numbers rather than on the more primitive notion of a qualitative preference relation.

We consider first the generalized associativity equation for functional equations. Associativity of addition means that $(a + b) + c = a + (b + c)$, so we can write $a + b + c$ unambiguously. If we write the addition operation as a two-place function $f(a, b) = a + b$, then associativity says that $f(f(a, b), c) = f(a, f(b, c))$. For general f , this is the *associativity functional equation*. The generalized associativity functional equation replaces each occurrence of f by a different function:

$$G(H(a, b), c) = K(a, M(b, c)).$$

This equation is described in Aczél (1968, §7.2.2), where many older references are given.

Now suppose that \succsim is a preference relation on \mathbb{R}^3 which satisfies the usual monotonicity and continuity conditions. Then there is a continuous representing function $F(a, b, c)$ that strictly increases in each argument. If the coordinate set $\{1, 2\}$ for the first two arguments is separable, then a relation $\succsim_{1,2}$ can be defined independently of the level at which the third argument is kept fixed. This relation also satisfies monotonicity and continuity, so a continuous strictly increasing representing function H can be found for $\succsim_{1,2}$. Because of separability for $\{1, 2\}$, all that is needed to determine the level of $F(a, b, c)$ are the levels of $H(a, b)$ and c . We then define G to satisfy $F(a, b, c) = G(H(a, b), c)$, where all functions are continuous and strictly increasing in their arguments. We arrive at the following conclusion (whose converse implication is obvious):

$\{1, 2\}$ is separable if and only if there exist G and H such that $F(a, b, c) = G(H(a, b), c)$, where all functions are strictly increasing and continuous. Similarly, $\{2, 3\}$ is separable if and only if there exist K and M such that $F(a, b, c) = K(a, M(b, c))$ with all functions continuous and strictly increasing.

Thus, given the usual continuity and monotonicity conditions, separability for a subset of coordinates is equivalent to the long-studied possibility of decomposing a representing function. It was established in the preceding century that under certain regularity conditions the generalized associativity functional equation

$$F(a, b, c) = G(H(a, b), c) = K(a, M(b, c))$$

implies that $F(a, b, c)$ has the form $f(F_1(a) + F_2(b) + F_3(c))$. This form implies that \succsim can be represented additively by $F_1(a) + F_2(b) + F_3(c)$, and therefore conjoint independence follows. The result extends in a natural way to more than three dimensions.

We now turn to the mathematical theory of generalized mean functions and relate it to DUR. A *simple probability distribution function* F

assigns to each real number y the probability that $Y \leq y$, where Y is a random variable taking finitely many real values. Let $\mathcal{D}(A, B)$ denote the set of simple probability distribution functions on the interval $[A, B]$, and let M be a function from $\mathcal{D}(A, B)$ to \mathbb{R} . Hardy et al (1934) proved that there is a continuous strictly increasing function u on $[A, B]$ such that

$$M(F) = u^{-1} \left(\int_{-\infty}^{\infty} u(x) dF(x) \right),$$

if and only if M satisfies the following conditions:

- (i) M assigns value x to the degenerate distribution function that has value 0 for all $y < x$, and value 1 for all $y \geq x$ ("x with certainty").
- (ii) M satisfies strict stochastic dominance: if $F \leq G$ and $F \neq G$ then $M(F) > M(G)$.
- (iii) If $M(F) = M(G)$, and $0 < \alpha < 1$, then

$$M(\alpha F + (1 - \alpha)H) = M(\alpha G + (1 - \alpha)H).$$

We relate the theory of generalized means to decision under risk by interpreting mean values as certainty equivalents of lotteries. Under the usual continuity and monotonicity conditions for a preference relation over lotteries, a certainty equivalent exists for each lottery, and preferences and certainty equivalence functions (generalized means) are uniquely related to each other. Condition (iii) is equivalent to independence of the preference relation over probability distributions. The first publication of the above result appears to be de Finetti (1931b, p. 379), which refers to (iii) as an associativity condition. This paper is remarkable in other respects. For instance, p. 386 gives a two-line proof for the result, later obtained by Pratt (1964) and Arrow (1965), that one person is more risk averse than another if and only if his or her utility is a concave transform of the other's utility.

De Finetti (1931b, p. 380) mentions a variation of (iii) that is also called associativity. To describe it, let M map $\bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ to \mathbb{R} , and assign to each n -tuple (x_1, \dots, x_n) a value between the maximal and minimal values in $\{x_1, \dots, x_n\}$. Then $M(a, \dots, a) = a$ for all a . Suppose further that M is symmetric in its arguments. Then it is called a *generalized mean*. M is *associative* if, for all $r < n < \infty$,

$$\begin{aligned} M(x_1, \dots, x_r, x_{r+1}, \dots, x_n) \\ = M(M(x_1, \dots, x_r), \dots, M(x_1, \dots, x_r), x_{r+1}, \dots, x_n). \end{aligned}$$

Let us now restrict attention for a fixed $m \in \mathbb{N}$ to $\mathbb{R}^m \subset \bigcup_{n \in \mathbb{N}} \mathbb{R}^n$. On this subdomain, associativity of M entails separability of the coordinate subset $\{1, \dots, r\}$, for each $r \leq m$, and thus, by symmetry, separability of each subset of $\{1, \dots, m\}$. So it implies conjoint independence, which in decision under uncertainty and under risk comes down to the sure-thing principle. To these conditions it adds symmetry. The definition of associativity of means seems to have been given first in Bemporad (1926, p. 87) in a characterization of the arithmetic mean $(x_1 + \dots + x_n)/n$. The present approach is related to DUR by identifying the n -tuple (x_1, \dots, x_n) with the probability distribution that assigns probability $1/n$ to each x_i . De Finetti did not explicitly mention the relevance for DUR even though he published a first version of his famous coherence argument in the same year (de Finetti 1931a).

Suitable regularity conditions imply that an associative mean M on $\bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ can be written as $u^{-1}((u(x_1) + \dots + u(x_n))/n)$ for a continuous strictly increasing function u . Then, for a preference relation related to the associative mean, the certainty equivalent $a = M(x_1, \dots, x_n)$ satisfies $u(a) = (u(x_1) + \dots + u(x_n))/n$. Thus u can be considered to be a von Neumann-Morgenstern utility representing the preference relation. Although the n -tuples (x_1, \dots, x_n) only describe probability distributions with rational probabilities, minimal continuity conditions give the expected utility representation for all simple probability distributions. The first papers to axiomatize associative means on $\bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ are Nagumo (1930) and Kolmogorov (1930).

References

- Aczél, J., *Lectures on Functional Equations and Their Applications*, Academic Press, New York, 1966
- Allais, M., "Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine," *Colloq. Internat. Centre National Rech. Sci.*, 40 (1953), *Econometrie*, 257–332. Translated into English, with additions, as "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School," in M. Allais and O. Hagen (Eds.), *Expected Utility Hypotheses and the Allais Paradox*, 27–145, Reidel, Dordrecht, The Netherlands, 1979.
- and O. Hagen (Eds.), *Expected Utility Hypotheses and the Allais Paradox*, Reidel, Dordrecht, 1979
- Allen, R. G. D., "A Comparison Between Different Definitions of Complementary and Competitive Goods," *Econometrica*, 11 (1934), 168–175.
- Anscombe, F. J. and R. J. Aumann, "A Definition of Subjective Probability," *Ann. Math. Statist.*, 34 (1963), 199–205.
- Arrow, K. J., "Alternative Approaches to the Theory of Choice in Risk-Taking Situations," *Econometrica*, 19 (1951), 404–437.
- , *Aspects of the Theory of Risk-Bearing*, Academic Bookstore, Helsinki, 1965; Elaborated as Arrow, K. J., *Essays in the Theory of Risk-Bearing*, North-Holland, Amsterdam, 1971.
- , Personal communication (August 5, 1991a).
- , Letter to P. A. Samuelson (September 9, 1991b).
- Bemporad, G., "Sul Principio della Media Aritmetica," *Rend. Acad. Nazionale dei Lincei*, 3 (1926), 87–91.
- Camerer, C. and M. Weber, "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity," *J. Risk Uncertainty*, 5 (1992), 325–370.
- Chernoff, H., "Remarks on a Rational Selection of a Decision Function," hectographed, Cowles Commission Discussion Papers No. 326 and 326A, January 1949; 346 and 346A, April 1950, Yale University, New Haven, CT
- , "Remarks on a Rational Selection of a Decision Function" (abstract), *Econometrica*, 18 (1950), 183.
- , "Rational Selection of Decision Functions," *Econometrica*, 22 (1954), 422–433.
- Chew, S. H., "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision

- Theory Resolving the Allais Paradox," *Econometrica*, 51 (1983), 1065-1092.
- , "The Rank-Dependent Quasilinear Mean," Unpublished Manuscript, Department of Economics, University of California, Irvine, 1989.
- and P. P. Wakker, "Generalizing Choquet Expected Utility by Weakening Savage's Sure-Thing Principle," University of California, Irvine Research Unit in Mathematical Behavioral Sciences, MBS 91-16, Irvine, CA, 1991.
- Clemen, R., *Making Hard Decisions: An Introduction to Decision Analysis*, PWS-Kent, Boston, MA, 1991.
- Dalkey, N. C., "A Numerical Scale for Partially Ordered Utilities," RAND memo 296, Dec. 5, 1949, Santa Monica, CA.
- de Finetti, B., "Sul Significato Soggettivo della Probabilità," *Fund. Mathematicae*, 17 (1931a), 298-329.
- , "Sul Concetto di Media," *G. dell'Ist. Ital. degli Attuari*, 2 (1931b), 369-396.
- Debreu, G., "Topological Methods in Cardinal Utility Theory," in K. J. Arrow, S. Karlin and P. Suppes (Eds.), *Mathematical Methods in the Social Sciences*, 1959, 16-26, Stanford University Press, Stanford, CA, 1960.
- Edwards, W. (Ed.), *Utility Theories: Measurement and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992.
- Ellsberg, D., "Classic and Current Notions of 'Measurable Utility,'" *Economic J.*, 62 (1954), 528-556.
- Epstein, L. G., "Behavior Under Risk: Recent Developments in Theory and Applications," in J. J. Laffont (Ed.), *Advances in Economic Theory II*, 1-63, Cambridge University Press, 1992.
- Fishburn, P. C., *The Foundations of Expected Utility*, Reidel, Dordrecht, The Netherlands, 1982.
- , "Transitive Measurable Utility," *J. Econ. Theory*, 31 (1983), 293-317.
- , *Nonlinear Preference and Utility Theory*, Johns Hopkins University Press, Baltimore, MD, 1988.
- , "Retrospective on the Utility Theory of von Neumann and Morgenstern," *J. Risk Uncertainty*, 2 (1989), 127-158.
- , "Skew Symmetric Additive Utility with Finite States," *Math. Social Sci.*, 19 (1990), 103-115.
- Fisher, I., "Mathematical Investigations in the Theory of Values and Prices," *Trans. Conn. Acad. Arts and Sci.*, 9 (1892), 1-124.
- , "A Statistical Method for Measuring 'Marginal Utility' and Testing the Justice of a Progressive Income Tax," in J. H. Hollander (Ed.), *Economic Essays Contributed in Honor of John Bates Clark*, 157-193, MacMillan, New York, 1927.
- Friedman, M., "Professor Pigou's Method for Measuring Elasticities of Demand from Budgetary Data," *Quarterly J. Economics*, 49 (1935), 151-163.
- and L. J. Savage, "The Utility Analysis of Choices Involving Risk," *J. Polit. Economy*, 56 (1948), 279-304.
- and —, "The Expected Utility Hypothesis and the Measurability of Utility," *J. Polit. Economy*, 60 (1952), 463-474.
- Gorman, W. M., "The Structure of Utility Functions," *Rev. Econ. Studies*, 35 (1968), 367-390.
- Gul, F., "Savage's Theorem with a Finite Number of States," *J. Econ. Theory*, 57 (1992), 99-110.
- Hardy, G. H., J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1934. (second edition 1952, reprinted 1978.)
- Herstein, I. N. and J. Milnor, "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21 (1953), 291-297.
- Howard, R. A. and J. E. Matheson, "Influence Diagrams," in R. A. Howard and J. E. Matheson (Eds.), *The Principles and Applications of Decision Analysis*, Vol. II, 719-762, Strategic Decisions Group, Palo Alto, CA, 1984.
- Jensen, N. E., "An Introduction to Bernoullian Utility Theory, I, II," *Swedish J. Econ.*, 69 (1967), 163-183, 229-247.
- Kahneman, D. and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (1979), 263-291.
- Karni, E., *Decision-Making Under Uncertainty: The Case of State-Dependent Preferences*, Harvard University Press, Cambridge, MA, 1985.
- and Z. Safra, "Behaviorally Consistent Optimal Stopping Rules," *J. Econ. Theory*, 51 (1990), 391-402.
- Keeney, R. L. and H. Raiffa, *Decisions with Multiple Objectives*, Wiley, New York, 1976.
- Kolmogorov, A., "Sur la Notion de Moyenne," *Rend. Acad. Nazionale dei Lincei*, 12 (1930), 388-391.
- Lau, J., J. P. Kassirer, and S. G. Pauker, "Decision Maker 3.0: Improved Decision Analysis by Personal Computer," *Medical Decision Making*, 3 (1983), 39-43.
- Leontief, W. W., "A Note on the Interrelation of Subsets of Independent Variables of a Continuous Function with Continuous First Derivatives," *Bull. Amer. Math. Society*, 53 (1947a), 343-350.
- , "Introduction to a Theory of the Internal Structure of Functional Relationships," *Econometrica*, 51 (1947b), 361-373.
- Luce, R. D., "Rational versus Plausible Accounting Equivalences in Preference Judgments," *Psychological Sci.*, 1 (1990), 225-234.
- and L. Narens, "Classification of Concatenation Measurement Structures According to Scale Type," *J. Math. Psychology*, 29 (1985), 1-72.
- and J. W. Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *J. Math. Psychology*, 1 (1964), 1-27.
- Machina, M. J., "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, 50 (1982), 277-323.
- , "Choice Under Uncertainty: Problems Solved and Unsolved," *Econ. Perspectives*, 1 (1987), 121-154.
- Malinvaud, E., "Note on von Neumann-Morgenstern's Strong Independence Axiom," *Econometrica*, 20 (1952), 679.
- McClennen, E. F., *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge University Press, Cambridge, 1990.
- Manne, A. S., "The Strong Independence Assumption—Gasoline Blends and Probability Mixtures," *Econometrica*, 20 (1952), 665-669.
- Marschak, J., Cowles Commission Discussion Paper (title unknown), Economics No. 226 (holographed), July, 1948, Yale University, New Haven, CT.

- Marschak, J., "Measurable Utility and the Theory of Assets" (abstract), *Econometrica*, 17 (1949), 63–64.
- , "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica*, 18 (1950), 111–141.
- , "Why 'Should' Statisticians and Businessmen Maximize 'Moral Expectation'?", in J. Neyman (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, CA, 1951.
- Nagumo, M., "Über eine Klasse der Mittelwerte," *Japanese J. Math.*, 7 (1930), 71–79.
- Nakamura, Y., "Subjective Expected Utility with Non-Additive Probabilities on Finite State Spaces," *J. Econ. Theory*, 51 (1990), 346–366.
- Nash, J. F., "The Bargaining Problem," *Econometrica*, 18 (1950), 155–162.
- Pratt, J. W., "Risk Aversion in the Small and in the Large," *Econometrica*, 32 (1964), 122–136.
- Quiggin, J., "A Theory of Anticipated Utility," *J. Econ. Behaviour and Organization*, 3 (1982), 323–343.
- Ramsey, F. P., "Truth and Probability," in *The Foundations of Mathematics and Other Logical Essays*, 156–198, Routledge and Kegan Paul, London, 1931. Reprinted in H. E. Kyburg and H. E. Smokler (Eds.), *Studies in Subjective Probability*, 61–92, Wiley, New York, 1964.
- Rubin, H., "Postulates for the Existence of Measurable Utility and Psychological Probability (abstract 493)," *Bull. Amer. Math. Society*, 55 (1949a), 1050–1051.
- , "The Existence of Measurable Utility and Psychological Probability," Cowles Commission Discussion Paper: Statistics: No. 322, 1949b, Yale University, New Haven, CT.
- , "A Weak System of Axioms for 'Rational' Behavior and the Nonseparability of Utility from Prior," *Statistics and Decision*, 5 (1987), 47–58.
- Samuelson, P. A., "Foundations of Analytical Economics, the Observational Significance of Economic Theory," Ph.D. dissertation, Harvard University, Boston, MA, 1940.
- , *Foundations of Economic Analysis*, Harvard University Press, Cambridge, MA, 1947.
- , "Probability and the Attempts to Measure Utility," *Econ. Review*, 1 (1950), 117–126. Reprinted in Stiglitz (1966, Chapter 12).
- , "Probability, Utility, and the Independence Axiom," *Econometrica*, 20 (1952), 670–679.
- , "Utilité, Préférence et Probabilité" (including discussion), *Colloq. Internat. Centre National Rech. Sci.*, 40 (1953), *Econométrie*, 141–164. English translation in Stiglitz (1966, Chapter 13).
- , "St.-Petersburg Paradoxes: Defanged, Dissected and Historically Described," *J. Econ. Literature*, 15 (1977), 24–55.
- , Personal communication (1991).
- Sarin, R. K. and P. P. Wakker, "Consistency in Dynamic Choice Situations with Nonexpected Utility Models," University of California, Los Angeles, CA, 1993.
- and —, "Folding Back in Decision Tree Analysis," *Management Sci.*, 40 (1994), 625–628.
- Savage, L. J., "The Theory of Statistical Decision," *J. Amer. Statist. Association*, 46 (1951), 55–67.
- , "Une Axiomatisation du Comportement Raisonnable Face à l'Incertain," *Colloq. Internat. Centre National Rech. Sci.*, 40 (1953), *Econométrie*, 29–40.
- , *The Foundations of Statistics*, Wiley, New York, 1954. (second edition, Dover, New York, 1972.)
- Schmeidler, D., "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57 (1989), 571–587. (first working paper version, 1982.)
- Segal, U., "Two-Stage Lotteries Without the Reduction Axiom," *Econometrica*, 58 (1990), 349–377.
- Slutsky, E. E., "Sulla Teoria del Bilancio del Consumatore," *G. Economisti*, 51 (1915), 1–26. Translated into English as "On the Theory of the Budget of the Consumer," in R. Irwin (Ed.), *Readings in Price Theory*, 27–56, American Economic Association, 1952.
- Sono, M., "The Effect of Price Changes on the Demand and Supply of Separable Goods," (in Japanese), *Kokumin Keisai Zasshi*, 74 (1943), 1–51.
- , "The Effect of Price Changes on the Demand and Supply of Separable Goods," *Internat. Econ. Review*, 2 (1961), 239–271.
- Stiglitz, J. E. (Ed.), *The Collected Scientific Papers of Paul A. Samuelson*, MIT Press, London, 1966.
- von Neumann, J. and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944/1947/1953.
- von Stengel, B., "Closure Properties of Independence Concepts for Continuous Utilities," *Math. Oper. Res.*, 18 (1993), 346–389.
- von Winterfeldt, D. and W. Edwards, *Decision Analysis and Behavioral Research*, Cambridge University Press, Cambridge, 1986.
- Wakker, P. P., "Cardinal Coordinate Independence for Expected Utility," *J. Math. Psychology*, 28 (1984), 110–117.
- , *Additive Representations of Preferences*, Kluwer Academic, Dordrecht, The Netherlands, 1989.
- , "Separating Marginal Utility and Probabilistic Risk Aversion," *Theory and Decision*, 36 (1994), 1–44.
- Weyl, F. J. and M. E. Bell, "Preference Patterns in the Face of Uncertainty," Summary of contributions to the International Colloquium on the Foundations and Applications of the Theory of Risk, held May 12–17 in Paris under the sponsorship of the Centre National de la Recherche Scientifique. Tech. Report ONRL-115-52, Office of Naval Research, London, 1952.

Accepted by Robert T. Clemen; received January 12, 1993. This paper has been with the authors 10 months for 1 revision.