ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

Alternative Approaches to the Theory of Choice in Risk-Taking Situations
Author(s): Kenneth J. Arrow
Source: *Econometrica*, Vol. 19, No. 4 (Oct., 1951), pp. 404-437
Published by: The Econometric Society
Stable URL: https://www.jstor.org/stable/1907465

ALTERNATIVE APPROACHES TO THE THEORY OF CHOICE IN RISK-TAKING SITUATIONS[1]

By Kenneth J. Arrow

This paper seeks to survey the literature in economics, philosophy, mathematics, and statistics on the subject of choice among alternatives the consequences of which are not certain. Attention is centered on the suggested modes of describing uncertainty and on the theories of rational and actual behavior of individuals in making choices.

1. INTRODUCTION

There is no need to enlarge upon the importance of a realistic theory explaining how individuals choose among alternate courses of action when the consequences of their actions are incompletely known to them. It is no exaggeration to say that every choice made by human beings would meet this description if attention were paid to the ultimate implications. Risk and the human reactions to it have been called upon to explain everything from the purchase of chances in a "numbers" game to the capitalist structure of our economy; according to Professor Frank Knight, even human consciousness itself would disappear in the absence of uncertainty.

I seek to survey here the present status of this theory, particularly in relation to choices in economic situations, as we ordinarily understand that term. The point of view will be that of a theory of choice, as it is usually conceived of. The general picture of such a theory is the following: There is a set of conceivable actions which an individual could take, each of which leads to certain consequences. The individual has in mind an ordering of all possible consequences of actions, saying, for each pair of consequences, either that he prefers one or that he is indifferent between them; these relations of preference and indifference have the property (known as transitivity) that if consequence A is preferred to consequence B and B to C, then A is preferred to C, and similarly with indifference. In a given situation, the range of actions open to an individual is limited in some way; thus, in the theory of consumers' demand under perfect competition, the actions possible are the purchases of bundles of consumers' goods whose cost does not exceed the available income. Among the actions actually available, then, that action is chosen whose consequences are preferred to those of any other available action.[2] In the

[1] I wish to express my gratitude to J. Marschak, of the Cowles Commission for Research in Economics and the University of Chicago, and L. J. Savage, University of Chicago, for many helpful comments. This paper will be reprinted as Cowles Commission Paper, New Series, No. 51.

[2] It is assumed that there exists a unique maximum; the more general case of none or many could be included by an obvious reformulation.

theory of consumer's choice it is not customary to differentiate between actions and their consequences since the two stand in one-to-one correspondence; but in the static theory of the firm we do distinguish between the actions—input-output decisions—and the consequences—varying levels of money profit. In the theory of choice under risky conditions, one of the chief problems is the description of consequences which are not certain and therefore certainly not uniquely related to the actions; the distinction between the two will be carefully maintained.

The range of actions available is, in a sense, no different from that in the theory of choice under certainty, particularly if the latter includes planning for the future, as in the theory of capital. However, as Professor A. G. Hart [**2**][3] has shown, certain distinctions among types of actions that are important for an understanding of choice in risk-taking situations are irrelevant under certainty and have therefore been ignored. Limitations of space will prevent our entering upon this subject.

There has been a steady, if slow, development in the formalization of the theory of choice under uncertainty and its relation to many business phenomena. However, three developments in recent years have represented dramatic breaks in continuity and have given hopes of a much clearer understanding of the problem: (1) the axiomatic treatment of choice among probability distributions by Professors von Neumann and Morgenstern [**1**], leading to a new understanding of the rule of maximizing the expected utility (in this, however, they were anticipated by F. P. Ramsey [**1**]); (2) the development of the modern theory of statistical inference, by Professors J. Neyman and E. S. Pearson [**1**] and Professor A. Wald [**1, 2**], which is a special form of the problem of (rational) behavior under uncertainty; and (3) Professor Shackle's new formulation of the whole problem of uncertain anticipations and actions based on them (see Shackle [**1, 2**]). This seems, therefore, to be an especially propitious time to take stock, to compare the various new developments with each other and with the whole background of the subject.

Two methodological remarks may be made at this point. (1) The uncertainty of the consequences, which is controlling for behavior, is understood to be that existing in the mind of the chooser. Of course, such subjective uncertainty or risk may very well stem from observations on the external world; thus, the behavior of an insurance company is derived from its observations on mortalities. I do not wish to face here the question whether or not there is any "objective" uncertainty in the economic universe, in the sense that a supremely intelligent mind knowing completely all the available data could know the future with

[3] Numbers in brackets refer to the list of references at the end of the paper.

certainty. The tangled web of the problem of human free will does not really have to be unraveled for our purpose; surely, in any case, our ignorance of the world is so much greater than the "true" limits to possible knowledge that we can disregard such metaphysical questions.

(2) Some of the theories discussed here purport to explain the actual behavior of individuals under conditions of uncertainty, some to give advice as to rational behavior, and some, by implication at least, to do both. In its broadest sense, rational behavior simply means behavior in accordance with some ordering of alternatives in terms of relative desirability, i.e., a theory of choice as described above. In some situations, however, there are additional conditions which appeal to the intuition as being rational. Almost all the theories discussed here seem to be rational in the first sense, but not all in the second. In view of the general tradition of economics, which tends to regard rational behavior as a first approximation to actual, I feel justified in lumping the two classes of theory together.

The plan of this study is dictated by the above outline of the structure of a theory of choice. A comparative discussion of the various ways proposed of *describing* uncertain consequences will be followed by the different theories of how they are ordered. Preceding these discussions will be a brief outline of certain facts of common experience in economic behavior which have been regarded as being related to the existence of risk, together with a sketch of the relation between scientific and statistical inference, on the one hand, and behavior under uncertainty on the other.

In the space allotted here, the coverage of the subject matter cannot be expected to be complete. Although the whole mathematical theory of probability has been, in a sense, a development of material relevant for behavior under uncertainty, the actual technical developments of recent years have not been concerned with the principles of such behavior. Indeed, the currently prevalent axiomatic treatment of probability as a branch of measure theory (see Kolmogorov [**1**]) seems designed to keep the technical development of the theory from being bogged down in the difficulties of the foundations. Also, incidental references by economists to problems of risk-bearing in the course of works on other subjects have for the most part not been discussed here; further, treatments before the beginning of the century have also largely been omitted. Finally, of course, there are doubtless many studies even within the proper bounds, especially in foreign languages, which I have not encountered.

2. THE BASIC EVIDENCE

2.1. Among economic phenomena which have in some way been tied up with the existence of uncertainty, three classes may be distinguished:

(1) those which by their very definition are concerned with uncertainty; (2) those which are not related to uncertainty by definition but nevertheless have no other conceivable explanation; (3) those whose relation to uncertainty is more remote and disputable.

2.1.1. Gambling and insurance are the obvious examples of the first type. In both cases, the consequences of the action involve a mathematical expectation of loss (as compared with nonperformance of the action), yet those actions are in fact undertaken. Gambling is exemplified by preferring the small probability of a large gain and the large probability of a small loss to the certainty of an income greater than the mathematical expectation of the gamble; insurance means preferring a certain small loss to the small chance of a large loss (see Friedman and Savage [**1**, pp. 289–291]). A theory of uncertainty must account for the presence of both.

Several writers have observed that the analysis of insurance just depicted is far from typical. Usually, the individual insured would have indirect losses in addition to those insured against; for example, an uninsured fire would probably entail a loss of credit rating (see Hardy [**1,** p. 57]; Hart [**1,** pp. 68–69]). Hence, it might well be that insurance is profitable, in the sense of mathematical expectation, to both parties.

One particular imaginary gamble, the St. Petersburg game, has played an important part in the history of uncertainty theory. A fair coin is tossed until a head appears; the individual is then paid 2^n ducats, where n is the number of tosses. How much should the individual pay for this gamble? It is easy to see that the mathematical expectation, $\sum_{n=1}^{\infty}(2^n)(1/2^n)$, is infinite, so that he "should" be willing to pay any finite stake. Introspection tells us that he will not, and this "fact" will have to be explained.

2.1.2. Among the economic phenomena of the second type are the existence of legally guaranteed incomes, variations in the rate of return on securities, and the holding of inventories beyond those demanded by pure transfer cost considerations. In a world of certainty, contractual obligations such as leases, bonds, and long-term labor contracts would have no significance, except possibly as a protection against dishonesty.

It is an obvious fact that securities and other contracts, which on their face promise the same returns, will have different market values, and these values are correlated in some sense with the risk that the contracts will not in fact be fulfilled. An extreme case is the fact that some individuals hold money, which brings no return, in preference to securities which promise some. (This phenomenon could, of course, still occur in a world of subjective certainty where individuals had differing price expectations.)

The holding of inventories is a more complicated matter since in a given firm it is usually impossible to distinguish among inventories due to indivisibilities in the cost of acquisition, unintended inventories, and inventories held for genuinely speculative motives. Only in transactions on the commodity markets do we have a relatively clear separation.

In summary, the observations just described have in common the following: They would not arise in the absence of risk, and they indicate that the reactions of individuals to a given risk situation are not all the same. Beyond that, probably little can be inferred without more definitely quantitative investigations.

2.1.3. The more recondite phenomena associated with the occurrence of risk in the minds of many economists include the existence of profits (in the "pure" sense of a residual after all factor payments including contract interest and imputed interest on the capital of the entrepreneur), the limitation on the size of the firm, and the characteristic form of the free-enterprise system.

It cannot be said that the evidence on any of these questions is at all clear. It has long been vaguely contended that "profits are the reward of risk-taking," in the sense that the expectation of profits is a necessary inducement for risk-bearing. However, after a long book devoted to establishing a variant of this proposition, Knight [**1**, pp. 362–366] suggests that perhaps profits in aggregate are negative. While negative profits call for explanation as much as positive ones do, it cannot be said that most of the current theories would really seem to have this possibility in mind. In any case, of course, the statistical evidence on the magnitude of pure profits can only be described as negligible.

The problem of limitation of the size of the firm arises because of the argument that under perfect competition in all markets there would be no limit to the size of the firm. The restriction is found in the unwillingness of the entrepreneur to borrow so much as to risk the wiping-out of his equity and in the imperfection of the capital market, due to the unwillingness of individuals to lend more than a limited amount to any one firm, again because of risk-feelings (see Kalecki [**1**]). The whole subject is wrapped up in the abstruse mysteries of capital theory and the coordinating function of the entrepreneur, and there is probably little point in pursuing the matter here. It is interesting to note, however, that as early as 1896, E. A. Ross argued that one of the chief motives for the *increase* in the size of the firms was the reduction in risk due to the operation of the law of large numbers (quoted by Hardy [**1**, pp. 20–21]); and Irving Fisher [**1**, pp. 408–409] and Knight [**1**, p. 257] express similar views.

Most authors agree that many of our characteristic institutions are shaped by the existence of risk; Knight [**1**, Chapter IX] goes so far as to maintain that the free-enterprise system as such arises as a reaction to the existence of uncertainty. Since the phenomenon of uncertainty must certainly have preceded the capitalist era, it would still have to be shown what differences between the present and the past explain the different social organization for meeting risk.

The phenomena discussed in this section are not likely, therefore, to be immediately useful in discriminating among various theories of uncertainty.

2.2. The businessman may be compared with two other types of individuals who are essentially concerned with behavior under uncertainty—the scientist and the statistician. The scientist must choose, on the basis of limited information, among the innumerable logically conceivable laws of nature, a limited number. He cannot know whether his decisions are right or wrong, and, indeed, it is none too clear what is meant by those terms. There is a long history of attempts to reduce scientific method to system, including many which introduce probability theory, but it cannot be said that any great formal success has attended these efforts. If we were to compare the businessman to the scientist, we would be forced to the melancholy conclusion that little of a systematic nature can be said about the former's decision-making processes.

The statistician typically finds himself in situations more similar to that of the businessman. The problem of statistics can be formulated roughly as follows: It is known that one out of a number of hypotheses about a given situation is true. The statistician has the choice of one of a number of different experiments (a series of experiments can be regarded as a single experiment, so that drawing a sample of any size can be included in this schema), the outcome of any one of which is a random variable with a probability distribution depending on which of the unknown hypotheses is correct. On the basis of that outcome, the statistician must take some action (accept or reject a hypothesis, estimate the mean of a distribution to be some particular value, accept or reject a lot of goods, recommend a change in production methods, etc.) the consequences of which depend on the action taken and on the hypothesis which is actually true.

Equivalently, we could describe the statistician as choosing, before performing the experiment, what experiment to perform and what action he would take for each possible outcome of the experiment. The actual action taken would be a random variable, depending on the outcome of the experiment. The consequences, therefore, of the initial decision would have a double uncertainty, depending on what outcome is observed and on which hypothesis is true. The statistician's problem is of the same general type as the businessman's, and even the information-getting aspects have their economic counterparts. The various theories which have been proposed from time to time as foundations for statistical inference are therefore closely related to theories of economic behavior under uncertainty.

3. THE DESCRIPTION OF UNCERTAIN CONSEQUENCES

With some inaccuracy, descriptions of uncertain consequences can be classified into two major categories, those which use exclusively the language of probability distributions and those which call for some other principle, either to replace or to supplement. The difference in viewpoints is related, though not perfectly, to the dispute between those who interpret probability as a measure of degree of belief (e.g., I. Fisher or Lord Keynes; see Sections 3.1.1, 3.2.1) and those who regard probability as a measure (objective) of relative frequency. [The most careful formulations of this view stem from the work of R. von Mises (see, e.g., von Mises [**1**]).] The latter concept clearly cannot encompass certain types of ignorance. For example, in the problem of statistical behavior (see Section 2.2), in any given situation the unknown hypothesis cannot be regarded as having a probability distribution in the frequency sense; just one hypothesis is true and all others false. Frequency theorists are therefore compelled to accept the view that probability statements cannot describe all kinds of ignorance (see Keynes [**1**, pp. 95–99]). This is indeed the viewpoint that has predominated among statisticians (see Section 3.2.3).

It is understood that in economic situations the consequences about which uncertainty exists are commodity bundles or money payments over future dates. Thus, in a probability description, the consequences of an action would be described as the *joint* probability distribution of commodities received or delivered over various points in the future. For many of the general principles, the fact that the actual consequences of an action form a multidimensional random variable is irrelevant.

The principal importance of the difference between the two types of description lies in the usefulness of the data on gambling and insurance (see Section 2.1.1) for a theory of uncertainty; accepting a pure probability description means that the more easily accessible information on such genuine probability questions is relevant for the theory of business risks.

3.1. Among those who describe uncertainty by means of probability, several groups may be distinguished: (1) those who treat the probability distributions of the consequences of alternative actions as subjectively given to the individual and who do not attempt further analysis; (2) those who derive all probability judgments from a limited number of a priori probabilities; (3) those who attempt to relate the degree-of-belief and frequency theories to each other through the law of large numbers.

3.1.1. Irving Fisher's theory [**1**, Chapter XVI and Appendix; **2**, Chapters IX and XIV] is typical of those who regard uncertainty as expressible by probability distributions. His typical example is that of a bond promising a series of interest payments and a final repayment of principal. For each time period there is a certain probability that the payment will be defaulted; in his more general analysis these probabilities are not independent, since a default of one payment usually entails the default of all subsequent payments.

In Fisher's view [**2**, p. 221], probability is simply an expression of ignorance; "risk varies inversely with knowledge." In the sense that the probability of an event is always measured relative to the available evidence, this proposition would be acceptable to virtually all schools of thought. Fisher further implies, as Laplace did before him, that with sufficient knowledge there would be no probability distributions, only certainty; the truth or falsity of this statement is certainly hardly relevant to the actual economic world.

Marshall, as usual, was not very explicit on the formulation of risk situations; however, it seems clear that he fully accepted the probability description in the same manner as Fisher did (see Marshall [**1**, pp. 135 fn., 398–400, 843]).

Among more recent writers, Professors Hicks [**1**], Friedman and Savage [**1**], and Lange [**1**] make use of the simple probability description. Their motives seem to be not so much a philosophical position as a view that the relevant phenomena can be explained by theories using only probability language, such theories being preferred on grounds of simplicity. Lange seeks somewhat more generality by supposing the individual to be able to form judgments only about the ordering of probabilities of different outcomes (this view bears some relation to that of Keynes; see Section 3.2.1 below) but not about the actual values of the probabilities. In application, however, cardinal probabilities reappear, so that Lange's attempt must be judged a failure (see Section 4.1.2).[4]

3.1.2. In the whole calculus of probabilities, there is a process of evaluating the probabilities of complex events on the basis of a knowledge of the probabilities of simpler ones. This process cannot go on indefinitely; there must be a beginning somewhere. Hence, in the study of games of chance, an a priori judgment is usually made as to certain probabilities. But in the usual types of events which occur in insurance or business affairs, there is no natural way of making these judgments; instead, the appeal is to past observations, if anything. The first sys-

[4] As a matter of fact, if it is assumed that probabilities of various events can be ordered, it follows easily that they can be measured by comparison with a random machine capable of producing an event B with any given known probability. I am indebted to Professor A. G. Hart, Columbia University, for an interesting discussion of this point several years ago, and to J. Marschak for this simplification of my earlier proof.

tematic study of the inference of probabilities from empirical evidence was the justly famous contribution of Thomas Bayes [**1**]. Understand by the symbol $P(A|B)$ the probability that A is true given that B is true. Suppose that it is known that one of the mutually exclusive hypotheses $B_1, \cdots, B_n$ is true, and suppose that an event A has occurred. Then the probability of the hypothesis B_i on the evidence A is given by

$$P(B_i \mid A) = P(A \mid B_i)P(B_i)/\left[\sum_{i=1}^{n} P(A \mid B_i)P(B_i)\right].$$

The a posteriori probabilities $P(B_i|A)$ then depend on the a priori probabilities $P(B_i)$ of the various hypotheses and on the probabilities $P(A|B_i)$ of the occurrence of the observed event A under the various hypotheses (The latter probabilities are sometimes referred to as the *likelihoods* of the various hypotheses given A.)[5]

Bayes' theorem shows clearly how a new piece of information, such as a price quotation or sales figures for a month, will modify the previous judgments as to the uncertainties of a situation. In a given context, the a priori probabilities are the judgments of the relative uncertainty of various hypotheses made on the basis of all past information; the a posteriori probabilities, the judgments made with the aid of new information.

The general problem of behavior under uncertainty can be formulated in this language as follows (see Tintner [**1–4**]: Suppose there are m possible mutually exclusive events $C_1, \cdots, C_m$, one of which will occur in the future. (An event might be an entire income stream over several future time points.) The probabilities of these events depend in a known way upon which of the hypotheses $B_1, \cdots, B_n$ is true; i.e., $P(C_j|B_i)$ varies with i for each given j. Finally, we have an a priori probability distribution for the hypotheses, $P(B_i)$ $(i = 1, \cdots, n)$. This is the formulation used by Tintner and Hart [**2**].[6]

3.1.3. How are the a priori probabilities appropriate to the above formulation obtained? In any particular context, of course, the a priori probabilities are the a posteriori probabilities of the preceding time period. But this process must have a start somewhere. At this point, the Principle of Insufficient Reason has been called into play. First formulated by Jacob Bernoulli in the seventeenth century, it states that if there is no evidence leading us to believe that one of an exhaustive set of mutually exclusive events is more likely to occur than another, then the events

[5] Here, and in most of the discussion, it will be assumed that the number of alternative possible events or hypotheses is finite. Most of the principles can be illustrated without considering the case of infinite sets of alternative events.

[6] However, the basic point of Hart's work is largely independent of this particular description.

should be judged equally probable. Thus, Bayes, in applying his theorem, postulated that, in the absence of further knowledge, the hypotheses B_i $(i = 1, \cdots, n)$ could be judged equally probable, and therefore $P(B_i) = 1/n$ for each i.[7] It follows then that the a posteriori probabilities are proportional to the likelihoods of the various hypotheses.

According to Keynes [**1**, Chapter IV], the contradictions flowing from this principle were first pointed out by J. von Kries [**1**] in 1886. Suppose we have a coin with completely unknown bias, and two tosses are made. In our assumed ignorance, we have no evidence for or against the hypothesis that the event of two heads will occur, and we must therefore assign it probability one-half, since the probability of the hypothesis equals the probability of its contradictory. But similarly, the probabilities of a head followed by a tail, a tail followed by a head, and two successive tails would each be one-half. This is impossible since the events are exclusive and their probabilities cannot therefore add up to more than one.

Keynes proposed to resolve these contradictions by restricting the application of the principle to the case where each of the events considered is incapable of further analysis. More precisely, an event A is said to be divisible if there are two exclusive events, A_1 and A_2, of the same "form" as A such that A occurs if and only if one of A_1, A_2 occurs; then the Principle of Insufficient Reason is applied only to a set of indivisible events. By the term "the same form" is probably meant "describable using the same language."[8]

It is hard to judge how satisfactory this theory is. It does avoid the paradox previously mentioned. However, it makes probability judgments depend on the form of the language, which may or may not be regarded as an objection. From a technical point of view, a greater deficiency is its failure to avoid the corresponding paradoxes which arise when there is a continuum of events rather than a finite number to consider.[9, 10]

[7] Bayes himself seems to have had his doubts about this postulate and, indeed, the paper was not published until after his death.

[8] Professor Carnap [**1**] has recently advanced a new formulation of the theory of probability as rational degree of belief. As in Keynes' work the assignment of a priori probabilities depends on consideration of indivisible events; however, the method preferred by Carnap [**1**, pp. 562–577] does not give them equal probability. An excellent account of Carnap's theories, with applications to statistical problems, is given by Tintner [**5**].

[9] Keynes' full theory of the description of uncertainty is, I believe, to be more properly classed with those which hold that probability statements alone are insufficient (see Section 3.2.1).

[10] Keynes' form of the Principle of Insufficient Reason leads to a different judgment of a priori probabilities in applications of Bayes' theorem than that usually made. Consider, for example, a finite population of n elements and let

3.1.4. Brief mention may be made here of another development of the concept of a priori probabilities. In this version the fundamental description of uncertainty is couched in terms involving other than probability statements (actually, along lines similar to those used by Neyman, Pearson, and Wald; see Section 3.2.3), but certain postulates are placed on the behavior of individuals under those conditions, and it is shown that they then act as if there were an a priori distribution of probabilities. Discussion of this viewpoint properly belongs with theories of ordering uncertain consequences and will be deferred until these are taken up; see Section 4.2.3.

3.1.5. Finally, mention must be made of a group which seeks to relate empirical observations to probability judgments via the law of large numbers. In its simplest form, due to Jacob Bernoulli, the law states that in a sequence of *independent* trials in each of which a given event E may occur with a constant probability p, the probability that the relative frequency of occurrence of E in n trials differs from p by more than any assigned positive quantity can be made as small as desired by making n sufficiently large. It is still true, of course, that in any finite number of trials, no matter how large, we cannot with certainty identify relative frequency with true probability. However, Buffon (quoted in Menger [**1,** p. 471 fn.]) and Cournot (see Fréchet [**1,** p. 62]) have suggested as a general principle that events whose probability is sufficiently small are to be regarded as morally impossible. Hence, in a sufficiently long series of trials, relative frequencies could be equated to probabilities. This theory therefore tends to lead to the same conclusions as the frequency theory.

The principle of neglect of small probabilities was used by Buffon to resolve the St. Petersburg problem (see Section 2.1.1). The probability that a head will not appear until the nth toss becomes very small for n sufficiently large; if the occurrence of that event is regarded as impossible for all n beyond a certain value, then the mathematical expectation of return becomes finite, and the paradox is resolved.

This principle seems extremely arbitrary in its specification of a particular critical probability and also runs into much the same difficulties of classification as does the Principle of Insufficient Reason. In an extreme case, suppose that we find an exhaustive set of mutually

m be the unknown number of them with some quality A; m may be any value from 0 to n. The usual treatment of this problem by Bayes' theorem would be to set the a priori probability of each value of m equal to $1/(n+1)$. Keynes' assumptions, however, would imply that m has a binomial distribution based on $p = \frac{1}{2}$. Carnap [**1,** p. 565] points out that Keynes' rule has a very undesirable consequence: after drawing any sample without replacement from a finite population (short of the whole population), the a posteriori probability that the next observation has the quality A is $\frac{1}{2}$ regardless of the sample observed. Carnap's rule avoids this difficulty.

exclusive indivisible events, each of which has a probability less than the critical value. It would be contradictory to say that all of them were impossible, since one must occur. This case actually occurs when a continuous random variable is considered.

The application of the neglect of small probabilities to Bernoulli's theorem also requires a knowledge that successive trials are independent, itself an a priori probability judgment. Hence, the principle does not supply a sufficient basis for probability reasoning.

3.1.6. The neglect of small probabilities is related to another question which has agitated the more philosophically-minded. Even if we describe uncertainty by probabilities, what significance do they have for conduct if the event occurs only once? Both Knight [**1,** p. 234] and Shackle [**2,** p. 71] argue that if the individual does not have the ability to repeat the experiment indefinitely often, the probabilities, being essentially long-run frequency ratios, are irrelevant to his conduct.

This argument would obviously have no validity in the degree-of-belief theory of probability. It is more serious in the frequency theory. It has sometimes been argued that even if an individual does not have the opportunity to repeat that particular experiment, he can nevertheless expect stability through the operation of a generalized Bernoulli's theorem in which the chance fluctuations in one trial are offset by those in others. This last argument is not entirely valid. In any one person's life the number of trials he can perform is finite; unless the neglect of small probabilities is invoked, the effect of a finite number of trials is, in principle, the same as one trial, the only difference being the quantitative one that the probability distributions are apt to be much more concentrated. Further, as Shackle has argued, in such a situation as a large investment, the number of comparable events in an individual's life is likely to be very small by anybody's standards.

While it may seem hard to give a justification for using probability statements when the event occurs only once, except on the interpretation of probability as degree-of-belief, the contrary position also seems difficult to defend. If an individual were told to predict whether or not two heads would come up in successive throws of a fair coin and further informed that he would lose his life if he guessed wrong, I find it very hard to believe that he would disregard the evidence of the calculus of probability. As is seen in Section 4.1.3, an extension of this reasoning suggests that in almost any reasonable view of probability theory the probability of a single event must still be the basis of action where there are genuine probabilities.

3.2. A number of writers maintain either that not all risky situations can properly be described by probability statements or that the latter do not provide a complete description.

3.2.1. Keynes, although an advocate of a degree-of-belief theory of probability (see Section 3.1.3), differs in certain ways from the various schools discussed in Section 3.1. Probability is for him, as for most others, a relation between the evidence and the event considered; but it is not necessarily measurable. He does not even consider that in general it is possible to order the probabilities of different events. It is true that every probability is considered to belong to some ordered sequence of probabilities, but they do not all belong to the same sequence.[11] The view that probabilities are not necessarily measurable seems to have some relation to Knight's distinction between measurable risks and unmeasurable uncertainties (see Section 3.2.2), and indeed they arise because Keynes is anxious to describe all uncertainties as probabilities. The actual effect of this extension of the probability concept is not clear since no applications are made.

Keynes, however, evidently does not regard the probability description as complete. He suggests, though tentatively, that in addition we must consider the *weight* of the evidence relative to which the probability is formed. Weight has the property that it increases with the evidence available, though it is not necessarily measurable. In Keynes' theory, then, uncertainty has the two dimensions, probability and weight, neither of which need be measurable. It is to be noted that weight is not necessarily related to the dispersion of a distribution; thus, in applications of Bayes' theorem, the a posteriori distribution of a numerical variable may have a larger standard deviation than the a priori, though the former must certainly have a greater weight. In ordinary problems of statistical inference from random samples, however, increase in weight and decrease in variance generally accompany each other (see Keynes [**1,** Chapter VI]).

3.2.2. Knight [**1,** Chapter VII] denies that all types of risks can be described by probability statements. His views stem from a functional view of uncertainty in relation to the problem-solving methods of the human mind. Because of the difficulties of reasoning, it is necessary for the mind to simplify matters by grouping cases. The relations between events derived from such an analysis will usually be uncertain rather than invariant. Some of these, of course, can be described by probability statements, which Knight classifies in the customary way as a priori and statistical; the two differ in the degree of homogeneity of the grouping. The classification upon which a statistical probability is based is partly empirical. Although Knight's analysis is so lacking in formal clarity that it is difficult to be sure, he appears to derive statistical

[11] Keynes [**1,** Chapter III]. Since the particular ordered sequences of probabilities need not be continuous, the argument given in footnote 4 cannot be applied here to establish a measurable probability.

probabilities from experience with the aid of the a priori in the manner of Bayes' theorem.[12]

Sharply distinguished from both a priori and statistical probabilities are true uncertainties or estimates. These arise when there is no valid basis for classification and are usually applied to situations which are in some sense unique. In such cases, according to Knight, the habit of the mind is not only to form an estimate of the situation but also to make a judgment as to its reliability, also on an intuitive basis.

No formal method of describing uncertainties or estimates is given by Knight, but some of their properties are discussed. He states that the estimates are generally based on some classification, either of objective circumstances or, more often, of ability of individuals to judge or of subjective confidence in one's own judgment. Later, he refers to "probable" or "improbable" outcomes of events about which uncertainty judgments have been made and argues that there is some tendency for the degree of uncertainty to be reducible by consolidation of many cases, analogously to the law of large numbers (see Knight [**1**, Chapter VIII]). In brief, Knight's uncertainties seem to have surprisingly many of the properties of ordinary probabilities, and it is not clear how much is gained by the distinction.

Knight appears to be worried about the seemingly mechanical nature of the probability calculus and its consequent failure to reflect the tentative, creative nature of the human mind in the face of the unknown. In a fundamental sense, this is probably correct, though it seems to lead only to the conclusion that no theory can be formulated for this case, at least not of the type of theory of choice discussed here (see the discussion of the scientist in Section 2.2; also see Section 4.3). The remark that estimates differ from statistical probabilities in that no valid principles of classification exist for the former, may lead in a more definite direction; it might be interpreted to mean that the a priori probabilities needed for Bayes' theorem are unknown. If this is correct, Knight's concept of uncertainty leads into the statistical theories discussed in the next section; and it should be noted that these theories seek to represent at least part of the knowledge-seeking propensities of human beings.

3.2.3. Professor Hardy [**1**, pp. 46, 53–55] suggested that Knight's case of uncertainty really belongs with statistical methods, though, of course, actual business judgments are apt to be biased. He therefore held that Knight's distinction was not a sharp one and, further, that the possibility of a more definite theory than Knight allowed existed.

[12] "It must be emphasized that any high degree of confidence that the proportions found in the past will hold in the future is still based on an a priori judgment of indeterminateness" (Knight [**1**, p. 225]).

At the time Hardy wrote, the foundations of statistical inference had been discussed chiefly in terms of inverse probability (Bayes' theorem), so it is not clear what he had in mind. However, contemporaneously with Hardy, Professor R. A. Fisher [**1**] was arguing that statistical inference required new principles not contained in the classical theory of probability. More recently there has been a formulation of the statistical problem due to Neyman and Pearson and modified by Wald (see Neyman and Pearson [**1**]; Wald [**1**; **2**, Section 1.1]). In the description of the statistical problem given in Section 2.2, the set of possible hypotheses is assumed known, but there are no judgments of a probability nature or otherwise which discriminate among the possible hypotheses in the range as to their truth. Thus, in an economic context, in the simplest case, there may be two possible hypotheses, of which it is known that one is true. For each action, then, we can associate an *income function* specifying for each possible true hypothesis what the consequences will be of that action. The problem is then to choose among the income functions. In a more general case, the income arising from a given action under a given hypothesis may itself be a random variable with a probability distribution specified by the hypothesis. The problem is then to choose among these possibly random income functions.

In this description there are two types of uncertainty: one as to the hypothesis, which is expressed by saying that the hypothesis is known to belong to a certain class or *model*, and one as to the future events or observations given the hypothesis, which is expressed by a probability distribution. If the model contains just one hypothesis, the uncertainty would be completely expressed by a probability distribution; if each hypothesis is a certainty, in the sense that given the hypothesis the future event is no longer uncertain, then the uncertainty is expressed by saying that the future event is one of a certain class without further information. It is to be noted that the Neyman-Pearson description does not eliminate all a priori elements since the model must still be chosen.[13] However, it does involve the elimination of a priori probabilities. It is also to be noted that, unlike Knight, Neyman and Pearson do not consider the uncertainties of the nonprobability type to be eliminable in any way by consolidation. If one model says that any income can be received from 0 to a and another, concerning an independent event, says that any income from 0 to b is possible, then the two models together say only that any income from 0 to $a + b$ is possible.[14]

[13] It is possible to use as a model the class of *all* logically possible hypotheses; but only meager statistical inferences could be drawn in that case.

[14] The importance of the Neyman-Pearson-Wald description for economic behavior has been chiefly stressed by Marschak [**1**, pp. 323–324; **2**, pp. 108–110; **3**, pp. 183–184, 192–195].

3.2.4. Shackle's formulation of uncertainty starts from a rejection of all probability elements, at least as applied to "large" decisions (see Shackle [**1**, pp. 6–7, Chapter VII]). He rejects the degree-of-belief theory of probability, since a distribution can never be verified even *ex post*, while the frequency theory is inapplicable where indefinite repetition is impossible (see Section 3.1.6). His own theory is a generalization of the Neyman-Pearson formulation in the case where the hypotheses determine the future events uniquely rather than via a probability distribution (see Shackle [**1**, pp. 10–17, 43–58, 128–133]). Neyman and Pearson divide hypotheses into two categories, possible and impossible; Shackle permits a continuous gradation from one case to another. To each possible future event he assigns a degree of *potential surprise*, the extent to which the individual would be surprised if the event occurs. Potential surprise is an ordinal concept;[15] the alternative possibilities are arranged on a scale from a minimum possible degree of potential surprise, such as might attach to a tautology, to a maximum, which is assigned to events deemed impossible.

Analogously to probability theory, rules are established for combining potential surprises; but the rules are of a nonadditive nature. Let $\eta(E)$ be the potential surprise of event E and $E_1 \cup E_2$ be the occurrence of one or another of E_1 and E_2 . Then, if E_1 and E_2 are mutually exclusive events, $\eta(E_1 \cup E_2)$ is equal to the smaller of $\eta(E_1)$ and $\eta(E_2)$; this proposition corresponds to the addition rule of probabilities. Further, let $\eta(E \mid F)$ be the potential surprise of E given that F has occurred, and let $E \cap F$ be the occurrence of both E and F. Then, analogous to the multiplication rule for compound probabilities, it is assumed that $\eta(E \cap F)$ is the larger of $\eta(E \mid F)$ and $\eta(F)$. It would be natural to say that E is independent of F if $\eta(E \mid F) = \eta(E)$; for such events, $\eta(E \cap F)$ is the larger of $\eta(E)$, $\eta(F)$. This shows that in Shackle's system, as in the Neyman-Pearson theory, there is no law of large numbers or elimination of risks by consolidation of independent events, for the potential surprise attached to a sequence of unfavorable outcomes is as large as to any one.

This last point shows that the elimination of all probability elements from Shackle's theory cannot be regarded as satisfactory. Since, in a single event, a probability distribution is irrelevant according to Shackle (see Section 3.1.6), the situation is described by a potential surprise function. But then no amount of repetition in independent trials would lead to reduction of risks, whereas even Shackle concedes that in long

[15] At some points, Shackle implies that degrees of potential surprise are measurable and even interpersonally comparable, but his whole theory can be developed without any such assumption.

runs probability rules would be applicable. In other words, Shackle's theory does not lead, even in the limit, to the probability theory.[16]

This argument can be given another form by means of an example. Consider the event of a fair coin's coming up heads twice. It certainly seems reasonable to assert that the potential surprise of this event must be greater than that of its not occurring. On the other hand, the potential surprise of a head on any given toss would be the minimum possible.[17] Under the assumption of independence of successive tosses, the potential surprise of two or any number of successive heads would be the same as of one, i.e., the minimum possible.

It would seem more reasonable, though this is not Shackle's view, to consider his concept of potential surprise as a straightforward generalization of the Neyman-Pearson theory; i.e., to attach the potential surprise to the hypotheses which define the probability distributions of future events. In that case, the Neyman-Pearson theory would appear as the special case where all potential surprises are either the maximum possible or the minimum possible.

4. THE ORDERING OF THE CONSEQUENCES

After having described uncertain consequences in some manner or another, it is necessary to set up a theory which will discuss the way in which an individual orders these consequences. Some theories say little more than that the consequences are orderable in some fashion; others make more definite assertions. The theory of preference among possible uncertainty situations will, of course, depend on the mode of description adopted.

4.1. Among those who describe uncertainties in terms of probability distributions, several theories of conduct have successively held the stage: the concept of maximizing the expected utility, the maximization of more general functionals, the Ramsey-von Neumann-Morgenstern revival and reinterpretation of expected utility, and a new stage, also originating with Ramsey, in which a priori probabilities are derived from behavior postulates. Except for the last group, all writers consider the case where the probabilities themselves are known directly.

4.1.1. As a solution of the St. Petersburg game (see Section 2.1.1), Daniel Bernoulli [**1**] proposed the proposition that the utility of a sum of money was proportional to its logarithm or at least that the marginal

[16] Some of the observations here and in Section 4.2.4 concerning Shackle's views have already been made by Graaf and Baumol in their excellent review article [**1**].

[17] Let A be the event of a head, $\bar{A}$ a tail, and, following Shackle, let 0 be the minimum possible degree of potential surprise. Since $A \cup \bar{A}$ is a tautology, $\eta(A \cup \bar{A}) = 0$; since A and $\bar{A}$ are mutually exclusive, min $[\eta(A), \eta(\bar{A})] = 0$. Since there are no grounds for preferring one of them to another, $\eta(A) = \eta(\bar{A}) = 0$.

utility of money was decreasing. It followed that the mathematical expectation of utility (rather than of money) in the game was finite, so that the individual would be willing to pay only a finite stake.

With the development of the utility theory of value in the 1870's, Bernoulli's proposal was found to fit in very well, especially in view of the common assumption of a diminishing marginal utility of income. Marshall [**1**, pp. 398–400] ascribes the risk-aversion he observes in business to this cause.

However, Marshall also notes [**1**, pp. 135 fn., 843] that under this hypothesis, gambling must be regarded as an economic loss. The practice of gambling then could only be explained by other considerations, such as the excitement derived therefrom. Professor Menger [**1**, pp. 479–480] argued that this showed that the Bernoulli theory was an insufficient resolution of the paradox.

Menger gave a further argument against the Bernoulli theory. Let $U(x)$ be the utility derived from a given amount of money x, and suppose that $U(x)$ increases indefinitely as x increases [e.g., if $U(x) = \log x$, as in Bernoulli's own theory]. Then, for every integer n, there is an amount of money x_n such that $U(x_n) = 2^n$. Consider the modified St. Petersburg game in which the payment when a head occurs for the first time on the nth toss is x_n. Then, clearly, the expected utility will be the same as the expected money payment in the original formulation and is therefore infinite. Hence, the player would be willing to pay any finite amount of money for the privilege of playing the game.[18] This is rejected as contrary to intuition.

4.1.2. Though the preceding argument would suggest that requiring the utility function to be bounded would overcome the paradox, Menger argues that in fact the discrepancy between actual behavior and that based on the mathematical expectation of return is explained not solely by undervaluation of large returns (through diminishing marginal utility) but also through undervaluation of small probabilities (a generalization of the neglect of small probabilities discussed in Section 3.1.5). He finally concludes that the ordering of different probability distributions (in the sense of the amount that an individual would be willing to pay for the various chances) depends on the diminishing marginal utility of money, the ratio of the initial size of his fortune to the minimum subsistence level in comparison with the chances of gain, and the systematic undervaluation of both very large and very small probabilities (and consequent overvaluation of medium probabilities).

[18] Menger [**1**, pp. 464–469]. The resolution of the St. Petersburg paradox by means of a bounded utility function was first proposed by the eighteenth-century mathematician Cramer in a letter printed in D. Bernoulli's original paper (see Bernoulli [**1**, pp. 57–58]).

A natural generalization of Menger's viewpoint is to consider simply that there is some ordering of probability distributions. If desired, this can be accomplished by assigning to each probability distribution a number representing its desirability. Such an assignment of numbers may be termed a *risk preference functional*, to use a term of Tintner's;[19] it is analogous to the assignment of numbers to commodity bundles in order of their desirability, which we call a utility function.

Tintner seems to have been the first to use this formal description, but the general idea goes back at least to Irving Fisher.[20] He argues that the value of a security will tend to depend on the probabilities of repayment. Usually, though not always, an uncertain income will be valued at less than its mathematical expectation. This risk aversion leads to devices for reducing the burden. The existence of guaranteed incomes such as bonds, as contrasted with stocks, is a device for reducing the risks for those guaranteed. Such devices do not reduce total risks but redistribute them; they come into being because different individuals have different risk aversions. Measures which serve to reduce total risk are inventory-holding, increase of knowledge, and insurance and speculation. Insurance is applicable only when the risks can be reduced to a statistical basis; otherwise it is the function of speculators to assume the risk. This specialization is a social gain, both because speculators normally have greater knowledge of the uncertainties than the average individual and because of the operation of the law of large numbers. However, if the speculators' forecasts tend to be interdependent, this law, which depends upon the independence of the random events consolidated, may be inoperative and the situation may actually be worse than without speculation.

In considering incomes which may assume a wide range of values, such as dividends, Fisher suggested that the purchaser would consider the expected value and the standard deviation of possible incomes as derived from his subjective probability distributions. Fisher himself seems to have had in mind the case where the probability distributions are normal, in which case, of course, knowing the mean and standard deviation amounts to knowing the entire distribution. However, the idea that the desirability of a probability distribution may depend solely upon a limited number of parameters and not upon the whole probability distribution became widely held, at least as a first approxima-

[19] Tintner [**1–4**]. The term "functional" is used because the measures of desirability are assigned to whole probability distributions rather than to variables in the ordinary sense.

[20] I. Fisher [**1**, Chapter XVI and Appendix to Chapter XVI]. The discussion of Pigou [**1**, Appendix I] is similar, though he puts it by saying that uncertainty-bearing is a factor of production, a form of wording that is, at best, a confusing figure of speech.

tion (see Marschak [**4**, pp. 118–119]). Thus Hicks [**2**] uses the mean and standard deviation, though his earlier discussion of uncertainty used only a generalized ordering of the probability distributions (see Hicks [**1**]); the utility of a distribution increased with its mean and decreased as the standard deviation increased. In connection with the theory of insurance, Professor Cramér suggested the probability that income will fall below some critical level as the criterion for ordering (see Cramér [**1**, Part III], as cited by Marschak [**4**, p. 120]); the higher this probability, the less desirable the distribution. Professors Domar and Musgrave [**1**] suggested that the utility of a distribution depends upon its *risk* and its *yield*, defined as follows: if $x_1, \cdots, x_n$ are the possible incomes with probabilities $p_1, \cdots, p_n$, respectively, with $x_1, \cdots, x_m$ negative and $x_{m+1}, \cdots, x_n$ positive, then the risk is $\sum_{i=1}^{n} p_i x_i$, while the yield is $\sum_{i=m+1}^{n} p_i x_i$. Lange [**1**, Chapter VI], in accordance with his general principle of avoiding measurable probabilities (see Section 3.1.1), suggests the mode and the range, utility increasing with the first and decreasing with the second. However, he shortly agrees that it is the "practical" range that is really relevant, and this may be defined as the interval between the fifth and ninety-fifth percentiles of the distribution (or any other similar pair of percentiles). Such a definition, of course, depends on measurable probability. Marschak [**1**, pp. 320–325] also suggested the use of the mean and variance as a first approximation, though other parameters, such as the skewness, are also considered to be relevant. Finally, Keynes [**1**, Chapter XXV] also thought that conduct should be guided not only by the expectation of utility but also by some measure of risk; in accordance with his views on the description of uncertainties (see Section 3.2.1), he also regards the weight of the evidence on which the probability distribution is based as still another variable in determining the utility of a given uncertainty situation.

4.1.3. The underlying intuitions which led to the rejection of the Bernoulli theory of maximizing expected utility stemmed from regarding utility as some objectively measurable quality in goods. It therefore seemed reasonable to say that not only the mean but also the dispersion of the possible utilities was relevant. This argument, however, was undermined by the rise of the indifference-curve view of utility, due to Pareto, where utility ceased to have any objective significance, and in particular diminishing marginal utility had lost its meaning (see Friedman and Savage [**1**, pp. 280–281]). The immediate effect of this development on the theory of risk was, as is natural, to actually hasten the trend towards a general ordinal theory of risk-bearing. But it was first observed by Ramsey [**1**][21] that, by starting from an ordinal theory of risk-bearing and adding a few reasonable assumptions, one could derive

[21] I am indebted for this reference to N. C. Dalkey, The RAND Corporation.

a utility function such that the individual could be said to behave in such a way as to maximize the expected value of his utility.

Ramsey's work was none too clear and attracted little attention. It remained for von Neumann and Morgenstern [**1**, Appendix]) to enunciate a clear set of axioms on choice among probability distributions which led to the assumption of maximizing expected utility and which were convincing. Recently Marschak [**4**] has given a simplified treatment, which will be followed here.

The key point turns out to be that no matter how complicated the structure of a game of chance is, we can always describe it by a single probability distribution of the final outcomes. Thus, suppose there are two gambling machines, one of which yields outcomes x_1 and x_2 with probabilities p and $1 - p$, respectively, while the second yields outcomes x_3 and x_4 with probabilities q and $1 - q$, respectively. Suppose now the choice of which machine is to be played is made by a preliminary gamble, in which machine 1 will be selected with probability r and machine 2 with probability $1 - r$. Then, of course, we have a complicated game with four possible outcomes x_1, x_2, x_3, x_4, with probabilities rp, $r(1 - p)$, $(1 - r)q$, $(1 - r)(1 - q)$, respectively. It obviously should make no difference for conduct, at least for the rational individual, whether he is told that he is playing the two-stage game or a single-stage game with the probabilities just described.

Suppose, for convenience, that the total number of possible outcomes is finite, say n in number. Then, denote by p $(p_1, \cdots, p_n)$ a probability distribution where p_i is the probability of the ith possible outcome; $p_i \geqslant 0$ for each i, $\sum_{i=1}^{n} p_i = 1$. If p and q are two probability distributions and a a real number with $0 \leqslant a \leqslant 1$, then by $ap + (1 - a)q$ we shall mean the distribution in which the probability of the ith outcome is $ap_i + (1 - a)q_i$. From the preceding paragraph it is easily seen that, from the viewpoint of behavior, the risk situation described by the distribution $ap + (1 - a)q$ is the same as that of the two-stage game in which a choice is made between the distributions p and q by means of a random device which chooses the first with probability a and the second with probability $1 - a$. We will refer to such a distribution as a *mixture* of p and q.

Von Neumann and Morgenstern start with the general principle that probability distributions are ordered. However, it is argued that the above considerations on the structure of probability distributions suggest certain further reasonable postulates on the choice among probability distributions. For example, since a mixture of p and q may be regarded as a game in which one of those two distributions is chosen by chance, the desirability of such a mixture should lie between that of p and that of q. It is natural, therefore, to assume, as Marschak does, that if p is preferred to q and q to r, then there is some mixture of p and r which is just indifferent to q. Further, suppose that p is indifferent to q. Then it is reasonable to assume that a mixture $ap + (1 - a)r$ with any other distribution r is indifferent to the same mixture of q with r, $aq + (1 - a)r$. For, in the two-stage game, we have in one case a probability a of getting p and $1 - a$ of getting r; in the other, a probability a of getting q and $1 - a$ of getting r. Since it makes no difference to the individual whether he receives p or q, by assumption, it seems reasonable to hold that the individual would be indifferent to the two gambles.

From these very reasonable assumptions,[22] the following remarkable theorem follows: There is a method of assigning utilities to the individual outcomes so that the utility assigned to any probability distribution is the expected value under that distribution of the utility of the outcome. The numbers assigned are said to be utilities because they are in the same order as the desirability of the distribution or outcome (desirability itself being merely an ordinal concept). The method of assigning utilities, furthermore, is unique up to a linear transformation.

The following remarks may be made about this result. First, the utilities assigned are not in any sense to be interpreted as some intrinsic amount of good in the outcome (which is a meaningless concept in any case). Therefore, all the intuitive feelings which led to the assumption of diminishing marginal utility are irrelevant, and we are free to assume that marginal utility is increasing, so that the existence of gambling can be explained within the theory. Second, the behavior postulates appear to be reasonable even under the frequency or any other definition of probability. The von Neumann-Morgenstern theorem then leads to the conclusion that the probability distribution is relevant even when only one event is to be observed; i.e., any definition of probability leads to a degree-of-belief interpretation. The objections, therefore, to the use of the probability concept in the absence of indefinite repetition seem to fall to the ground (see Section 3.1.6).

Some of the literature on the ordinal theory of risk-bearing can be interpreted within the present framework. Menger's objection to the Bernoulli theory in the case of unbounded utility functions (see Section 4.1.1) can be reformulated as follows: If the utility function is unbounded, then there is a probability distribution with infinite utility. If, as seems natural, we demand that all utilities be finite,[23] then Menger's argument leads to the conclusion that no utility should exceed some fixed upper bound.

[22] Actually, these assumptions are not quite adequate. An additional one which would enable the theorem to be proved is, e.g., that there are at least four probability distributions, no two of which are indifferent (or all are indifferent).

[23] Pascal, of course, considered the hope of heaven to be of infinite value compared with any earthly reward. But I doubt that he would have regarded any version of the St. Petersburg game as entitled to such a special place.

The characterization of probability distributions by mean and variance would be valid, at least in a special case, under the present theory if the utility function of income were quadratic; in that case the utility of a probability distribution would be the same quadratic function of the mean plus a term proportional to the variance. Finally, the rule of minimizing the probability that the income falls below some critical level would be a special case of the von Neumann-Morgenstern theory where the utility was 0 for all incomes below the critical level and 1 for all incomes above it.

Friedman and Savage [**1**] have used the von Neumann-Morgenstern construction to arrive at some information about the utility function for income. Members of lower-income groups both take out insurance and gamble. Insurance is rational if the utility function has a decreasing derivative over the interval between the two incomes possible (decreasing on the average but not necessarily everywhere), while gambling is rational if the utility has a predominantly increasing derivative over the interval between the possible incomes. In view of the structure of gambles and insurance (see Section 2.1.1), this requires that the utility function have an initial segment where marginal utility is decreasing, followed by a segment where it is increasing. The fact that lotteries generally have several prizes instead of one grand prize equal to the sum indicates that the willingness to gamble does not rise indefinitely; therefore, there must be a final segment of high incomes in which marginal utility decreases again.

These results were anticipated in part by Professor L. Törnqvist [**1**]. Arguing from the fact that some lottery tickets will be unsold and therefore the lottery owner is also running a risk, it follows that the utility function must be such as to explain the behavior of both lottery-buyers and sellers. He suggests on this basis the following as a utility function for income:

$$U(x) = x\,\{1 + k[x/(K + x)]^2\},$$

where k and K are constants, and where there is diminishing marginal utility for small values of x and increasing marginal utility for other values.

4.2. We may now turn to the theories of behavior under uncertainty held by those who do use other than probability statements in their description. I will include in this group, for convenience, those who use a priori probabilities not as a fundamental assumption but as a derivation from postulates on behavior (see Sections 3.1.4 and 4.2.3).

4.2.1. As already indicated (see Section 3.2.2), Knight has not given a formal description of uncertainty situations in his sense (as opposed to risks or probability distributions) and correspondingly has given no definite set of assumptions concerning their ordering. Actually, his uncertainties produce about the same reactions in individuals as other writers ascribe to risks. For example, the list of devices which are used to reduce uncertainty is about the same as that given by Irving Fisher when starting from a strictly probability point of view (see Knight [**1**, Chapter VIII]).

There is, however, one analytic point of considerable interest in Knight's theory. He asserts as a basic theorem that if all risks were measurable, then risk-aversion would not give rise to any profit (see Knight [**1**, pp. 46–47]). The argument is that, in that case, the law of large numbers would permit the wiping-out of all risks through insurance or some other form of consolidation of cases. This proposition, if true, would appear to be of the greatest importance; yet, surprisingly enough, not a single writer, as far as I know, with the exception of Hicks [**1**], has mentioned it, and he denies its validity.

Knight's argument does not seem to be a rigorous proof by even a generous stretch of the imagination. His proposition, of course, depends not only on the behavior of individuals in risk-taking situations but also on the workings of the economic system as a whole and the way in which random events, such as technological uncertainty, enter. Knight gives no indication of how such a fundamental concept as constant returns would appear under uncertainty, though he is most meticulous in his corresponding discussions under certainty.

Some indications of what might happen are contained in a paper by C. Palm [**1**] (discussed in Feller [**1**, pp. 379–383]). Palm considers the case of a number of automatic machines which need tending only when out of commission, an event which is supposed to occur in a random manner. The time needed for repair is also a random variable. Palm then calculated the probability distribution of the length of time the machines are operating for a given number of machines and a given number of repair men. (The conditions of the problem are that when all the repair men are busy, any machine which then breaks down must remain idle until one of the men finishes repairing the machine he is working on.) It is shown that the expected amount of operating time would more than double if both men and machines are doubled. This suggests that under conditions most favorable to Knight's assertion, i.e., independent technological risks, there would be a universal tendency towards increasing returns, which would be incompatible with his basic assumption of perfect competition.

A somewhat more obvious criticism of Knight's result is the argument that in fact the risks occurring are mutually correlated, so that the law of large numbers may not operate. This point was already made by Irving Fisher (see Section 4.1.2) and was used by Hicks [**1**, pp. 175, 188] as a criticism of Knight's proposition. Hicks also observes that contractual guarantees can never constitute a complete shifting of risks, since there is always some risk of default (see Hicks [**1**, p. 178]).

Finally, it may be observed that even if Knight's proposition were true, the use he makes of it would not be justified. He argues that this proposition, in conjunction with the existence of profits, shows that there must be other types of uncertainty. But, of course, the consolidation of risks would not proceed to the limit if there were a predominance of risk-preferring attitudes,[24] and we would have on the aggregate negative profits (though by chance individual firms will have positive profits). Since Knight concludes by suggesting that aggregate profits are in fact negative (see Section 2.1.3), he argues that there must be some preference for uncertainty-bearing in his sense; but there seems to be no reason why he could not have used the simpler theory of postulating risk preference in regard to probability distributions and dispensed with the distinction between the two types of uncertainty.

4.2.2. Neyman and Pearson did not have a unique principle for choosing among actions under their formulation of uncertainty (see Section 3.2.3). We may consider as a typical illustration their discussion of behavior when there are only two possible hypotheses H_1 and H_2 in their model, and an action is considered to be a rule for choosing, for any possible observation, one of the two hypotheses. For a given observation, the action may be described as right or wrong. If H_1 is true, for a given rule of action, there is a certain probability that the observation will call for a wrong action, and similarly with H_2. Neyman and Pearson suggest the following rule: Fix some probability α; then, among all actions for which the probability of error if H_1 is true is α or less, choose an action which minimizes the probability of error if H_2 is true.

This rule is arbitrary in that there is in general no criterion for choosing α, and there will be a different best action for each α. The economist will notice immediately the analogy to Pareto's weak definition of an optimal economic state, and the above criterion has been generalized to a form even more like Pareto's by Professor Lehmann [**1**]. We will say that one action *dominates* another if the (expected) return[25] to the first action is at least as great as to the second action under each of the alternative possible hypotheses, and actually greater for at least one

[24] Consider the following example: Individuals A and B, facing independent risk situations, each have a probability p of receiving 0 and a probability $1 - p$ of receiving 1. If they consolidate and share the returns equally, they each have a probability p^2 of receiving 0, $2p(1 - p)$ of receiving $\frac{1}{2}$, and $(1 - p)$ of receiving 1. Assume that they act so as to maximize the expected utility; then consolidation will be preferred if and *only if* $U(\frac{1}{2}) > \frac{1}{2}[U(0) + U(1)]$, a situation of diminishing marginal utility of income, or risk-aversion.

[25] For a given action and a given hypothesis, the return (consequence) is a random variable with a known probability distribution. Each such distribution can be given a numerical utility in accordance with the von Neumann-Morgenstern theorem (see Section 4.1.3).

hypothesis. Then an action A is said to be *admissible* if there is no other available action which dominates A.

The rule that one should restrict oneself to admissible actions is extremely reasonable, but, like the corresponding rule in welfare economics, it rarely leads to definite decisions as to which action to take. Wald [**1**; **2**, pp. 18, 26–27] has suggested the following rule: for each action consider the minimum of the expected returns possible under the alternative hypotheses; choose that action which maximizes the minimum return.[26] This means, in effect, that that action should be chosen about which the best certain statement can be made. This rule can be interpreted as making behavior under uncertainty analogous to a zero-sum two-person game with the individual and Nature as the two players; the individual chooses the action, Nature chooses the hypothesis (each without knowing the opponent's "move"), and the individual then receives the expected return determined by the two choices.

Wald's theory is intuitively highly appealing in that it reflects fully the idea of complete ignorance about the hypotheses. However, the theory of the zero-sum two-person game is based on the idea of an opponent with a definite interest in reducing one's gains (see von Neumann and Morgenstern [**1**, pp. 98–100, 220]). No such motive can be ascribed to Nature. A difficulty arising in connection with Wald's principle naively stated, as well as a possible resolution of it, was first suggested by L. J. Savage.[27] Suppose that there are two possible actions, A_1 and A_2, and two possible hypotheses, H_1 and H_2. Suppose that if A_1 were chosen, the utility would be 100 under H_1 and 0 under H_2, while if A_2 were chosen, the utility would be 1 under either hypothesis. Wald's criterion would call for choosing A_2; yet, since it can hardly be said that Nature would choose H_2 deliberately to prevent the individual from realizing the gains he would receive with A_1 under H_1, it does not seem reasonable that the hope of a small gain under H_2 should outweigh the possibility of a large gain under H_1, especially since it can be shown that the choice would not be altered by any accumulation of observations. Savage and, independently, J. Niehans [**1**, p. 445] argued that the statistician or businessman is only responsible for doing as well as he can under the hypothesis which actually prevails. For each possible hypothesis, therefore, he should find the best he could expect if he knew the hypothesis were true, i.e., the maximum of the returns of all possible actions under that hypothesis. If he takes action

[26] Wald himself refers to losses rather than gains, and so he states the rule as that of minimizing the maximum loss. The criterion is therefore frequently referred to as the *minimax principle*.

[27] In his course in mathematical statistics (oral lectures), winter, 1948; see also Marschak [**3**, pp. 192–195]; Savage [**2**, pp. 58–62].

A and it turns out that hypothesis H is correct, the statistician should be given a penalty (which has been given the not altogether felicitous name of *regret*) equal to the difference between the reward to A under H and the maximum income possible under H if known. The regret is computed in this way for each possible action and hypothesis; the individual is then supposed to choose his action so as to minimize the maximum regret.

Like Wald's, Savage's principle expresses the idea of complete ignorance of the true hypothesis and at the same time seems to get around the assumption of a malevolent universe. But there is another difficulty, pointed out by Professor Chernoff [**1**]. Suppose the incomes accruing under various combinations of three actions and two hypotheses are given by the following table:

Action	Hypothesis 1	Hypothesis 2
1	0	10
2	3	6
3	6	0

Under hypothesis 1, the best action would yield 6; therefore, the regrets of the different actions under hypothesis 1 are the differences between the numbers in the first column and 6, and similarly with the regrets under hypothesis 2. The regret matrix is, then,

Action	Hypothesis 1	Hypothesis 2
1	6	0
2	3	4
3	0	10

Under the principle of minimaxing regret, action 2 would be chosen.[28] But now suppose that action 3 were no longer allowed. Then a new regret matrix would have to be formed, since, for example, the best available action under hypothesis 1 would now yield only 3.

Action	Hypothesis 1	Hypothesis 2
1	3	0
2	0	4

[28] This is not true if randomized actions are permitted, but a similar paradox will obtain.

Now, however, action 1 would be chosen. In other words, even though action 2 would be preferred when both 1 and 3 are available, action 1 would be chosen over 2 if 3 is no longer available. This result does not seem compatible with rational behavior; the difficulty is that the Savage minimax principle does not and is not intended to give rise to a genuine ordering in the usual economic sense (see also Savage [**2**, p. 64]).

4.2.3. The preceding discussion suggests the possibility of setting down axioms for rational behavior under uncertainty starting with the Neyman-Pearson description, i.e., axioms on the ordering of income functions (see Section 3.2.2). Ramsey [**1**], Professor de Finetti [**1**], and Savage [**1**] start with a particular class of actions, namely, betting on outcomes of events (see, in particular, Ramsey [**1**, pp. 179–184]; de Finetti [**1**, pp. 4–11]). Thus de Finetti says that if an individual is willing to exchange the sum of money pS but no more for the privilege of receiving S when the event E occurs, then p is the "subjective" probability of E. A similar definition can be given for conditional probability. Then, if it is postulated simply that there is no system of bets of the above type by which the individual feels sure of winning regardless of the actually observed event, then all the usual theorems of probability can be deduced. This particular formulation is unacceptable since it is implicitly postulated that p is independent of S, an assumption contradicted by the everyday experience that individuals will not make indefinitely large bets at the same odds. Ramsey's construction is similar, except that he substitutes utility (see Section 4.1.3) for money as the medium of betting. To define the utility itself requires some probability elements; it is presupposed that there exists an event which the individual would as soon bet for as against, i.e., in effect an event with probability $\frac{1}{2}$.

More recently there have been other discussions by Professors Rubin [**1, 2**] and Chernoff [**1, 2**].[29] Their starting point is the concept of a mixture of income functions, analogous to a mixture of probability distributions in the von Neumann-Morgenstern theory (see Section 4.1.3). If x and y stand for two income functions, then the mixture $ax + (1 - a)y$, where $0 \leqslant a \leqslant 1$, is the income function derived by choosing x with probability a and y with probability $1 - a$. It can also be obtained by mixing the returns for each possible hypothesis corresponding to x and y; for any given hypothesis, this mixing is of the type already described. Then it is argued that the von Neumann-Morgenstern axioms already described for probability distributions are equally reasonable when applied to income functions. That is, it is assumed (1) that income functions can be ordered;[30] (2) that if x is preferred to y and y to z, there is a mixture of x and z indifferent to y; (3)

[29] Savage has made important oral contributions to this development.

[30] Rubin has shown that this condition can be considerably weakened.

if x is indifferent to y, then any mixture of x and a third alternative z is indifferent to the same mixture of y and z.[31] If we further add the condition that an income function which is not admissible (see Section 4.2.2) will not be used, then it can be shown there exists an assignment of nonnegative numbers $p_1, \cdots, p_n$ to the various hypotheses $H_1, \cdots, H_n$ and of utilities to the probability distributions defined by the actions and hypotheses such that the desirability of an income function x is measured by the utility function $\sum_{i=1}^{n} p_i U(x, H_i)$, where $U(x, H_i)$ is the von Neumann-Morgenstern utility of the outcome specified by the income function x for the hypothesis H_i. We can so choose the p_i's that $\sum_{i=1}^{n} p_i = 1$; then it is natural to interpret them as a priori probabilities. The theorem, in effect, says that in order for an individual to act rationally he must act as if he had a set of such probabilities in mind.

Chernoff, like J. Bernoulli, argued that in the case of complete ignorance all the alternative hypotheses must be treated on a par, so that the a priori probabilities would all be equal to $1/n$. This argument supplies a justification for the Principle of Insufficient Reason (see Section 3.1.3).

It should be noted in passing that both the von Neumann-Morgenstern utility theory and the present extension to a priori probabilities are ideally capable of refutation. Both the utilities and the probabilities can be discovered by suitable formulations of choices in simple situations; then behavior in more complicated situations can be checked against that predicted by the theory with the numerical data supplied by study of simple cases. Similar considerations apply to the minimax theories.

4.2.4. Shackle's theory [**1**, Chapter II] of uncertainty-bearing stems from his description in terms of potential surprise (see Section 3.2.4). The exposition is greatly complicated by his insistence on differentiating between gains and losses. It is completely unclear to me what the meaning of the zero-point would be in a general theory; after all, costs are usually defined on an opportunity basis only.

For a given action, there is attached to each possible gain a potential surprise. The possible pairs of this type are ranked according to their power to stimulate the mind, and it is assumed that the effect of the most stimulating one will sum up the entire influence of possible gains on choice. The stimulation of an outcome is assumed to increase as the size of the gain increases and to decrease as the potential surprise increases. Similarly, the maximum stimulation associated with a loss is found. This pair of stimulations completely defines the effects of the

[31] The third assumption can be shown to eliminate the difficulty found by Savage in the Wald minimax principle.

given action; there is then an ordering among pairs of stimulations (referred to as the *gambler indifference map*) which determines which action is taken.[32] This theory is not based on considerations of rational behavior, which Shackle [**2**, p. 74] specifically rejects, but on an alleged inability of the mind to consider simultaneously mutually exclusive events.

In the same way that the concept of potential surprise is a generalization of the Neyman-Pearson formulation of the statistical problem, so Shackle's theory of behavior under uncertainty generalizes the Wald minimax principle. Suppose (1) that only losses occur or else that the gambler's indifference map depends only on the standardized focus-loss, and (2) that for any outcome, potential surprise is either 0 or else the maximum possible. Then Shackle's theory leads to the rule of minimizing the maximum loss.

Shackle has the sound impulse to base a theory of uncertainty-bearing on the necessity of the human mind to simplify a problem in order to be able to deal with it. However, his particular simplification seems to be purely arbitrary. We have already seen that the complete rejection of probability elements cannot be regarded as successful (see Section 3.2.4). As a matter of fact, description *solely* in terms of probability distributions seems to me to be at least as plausible as Shackle's hypothesis, which is, further, lacking in the virtue of simplicity. (Why do we need both a stimulation function and a gambler indifference map to describe reactions to uncertainty?)

The theory can indeed be shown to lead to an odd result. Suppose an event has two possible outcomes, A and B. Then it follows from the assumptions about potential surprise (see Section 3.2.4) that one of them, say A, must have potential surprise 0. Suppose that under action 1 the individual gets one dollar regardless of what happens, while under action 2 he gains one dollar if A occurs and nothing if B occurs. Then, in both cases, he gains one dollar with potential surprise 0 and loses nothing with the maximum possible potential surprise, so that the two actions have the same standardized focus-outcomes (viz., focus-gain 1 and focus-loss 0) and are indifferent. Yet the second action is obviously not admissible (see Section 4.2.2).

This result also shows that the theory is, at least, a genuine theory in the sense of having refutable consequences. In fact, it is possible to construct a sequence of simple experiments which will enable us to

[32] Shackle defines a *standardized focus-gain* as that outcome which, if potential surprise 0 were attached to it, would be as stimulating as the most stimulating gain, and a *standardized focus-loss* similarly. The gambler indifference map has as coordinates the pair of standardized focus-outcomes. These are obviously merely a method of attaching numbers to the maximum degrees of stimulation and have no special significance.

actually find the potential surprise function, the stimulation function, and the gambler indifference map. I will briefly sketch the procedure. For any two hypotheses, H_1 and H_2, let A_1 be an action which yields one dollar if H_1 is true and nothing otherwise, and let A_2 be an action yielding one dollar if H_2 is true and nothing otherwise. Then H_1 has a lower potential surprise than H_2 if and only if A_1 is preferred to A_2. Since potential surprise is an ordinal concept, that function is completely defined. In particular, we may find a hypothesis H which has the same potential surprise as its contradictory; then both will have potential surprise 0. Consider an action for which the individual receives a if H occurs and loses b otherwise. Then the standardized focus-outcomes are a and b; by ordering all actions obtained by varying a and b, we can fill out the gambler indifference map. Finally, for any given degree of potential surprise y and any given gain c, choose a hypothesis H_1 with potential surprise y. Let action A_1 yield c if H_1 is true and 0 otherwise; let action A_2 yield d if H is true and 0 otherwise. Vary d until A_1 and A_2 are indifferent; then, since H has potential surprise 0, d is the standardized focus-gain corresponding to the pair (y, c). The stimulation function is therefore defined for gains; a similar procedure is applicable to losses.

Though the theory is therefore not empty, it has not actually given rise to much in the way of useful consequences. The discussion of the holding of assets under uncertainty (see Shackle [**1**, Chapter IV]) is vitiated by a logical error at the beginning of the argument. As a matter of fact, since there is no law of large numbers in the calculus of potential surprise, diversification of assets cannot be explained. His discussion of the bargaining process (see Shackle [**1**, Chapter VI]) amounts to the well-known proposition that the process is determinate if the reaction curves of the participants are known, the only difference being that the reaction curves constructed by Shackle include uncertain behavior.[33]

4.3. There is one body of opinion which argues that there can be no theory of rational behavior under uncertainty by the nature of the problem. As early as 1879, Cliffe-Leslie argued that the wiping-out of profits by competition would not occur in reality since the profits of other firms were unknown (quoted by Hardy [**1**, pp. 35–37]). Irving Fisher also doubted that a satisfactory theory of behavior under uncertainty was really possible (see I. Fisher [**2**, Chapters IX, XIV]), though his conclusion was that the theory of economic behavior under certainty afforded a good approximation to actuality. Knight and Shackle (see

[33] An interesting suggestion of Shackle's is that when an event with high potential surprise has occurred, forcing a re-evaluation of the uncertainty situation, the effect will not take place all at once but only after an interval of time during which almost every possibility is assigned a low potential surprise (Shackle [**1**, pp. 73–75]). This proposition could be reformulated in terms of other theories.

Sections 4.2.1 and 4.2.4) tend to somewhat similar viewpoints. Recently, Professor Alchian [1] has argued that the absence of definite criteria for action, corresponding to profit maximization in the theory of behavior under certainty, means that we should regard behavior under uncertainty as essentially random. The process of convergence to optimal behavior by trial and error is impossible when the basic conditions are changing and unknown. However, since there will be a process of selection in the economic struggle for existence, there will be some tendency towards optimal behavior.

Stanford University and Cowles Commission for Research in Economics

REFERENCES

ALCHIAN, A. A. [1], "Uncertainty, Evolution, and Economic Theory," *Journal of Political Economy*, Vol. 58, June, 1950, pp. 211–221.

BAYES, THOMAS [1], "An Essay toward Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society of London*, Vol. 53, 1763, pp. 370–418. (Reprinted in *Facsimiles of Two Papers by Bayes*, W. D. Deming, ed., Washington, D. C.: The Graduate School, Department of Agriculture, 1940.)

BERNOULLI, D. [1], "Specimen theoriae novae de mensura sortis," *Commentarii academiae scientiarum imperiales Petropolitanae*, Vol. 5, 1738, pp. 175–192; translated by A. Pringsheim under the title, *Die Grundlage der modernen Wertlehre: Versuch einer neuen Theorie der Wertbestimmung von Glücksfallen*, Leipzig: Duncker and Humblot, 1896, 60 pp.

CARNAP, RUDOLF [1], *Logical Foundations of Probability*, Chicago: University of Chicago Press, 1950, 607 pp.

CHERNOFF, HERMAN [1], "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.

——— [2], "Remarks on a Rational Selection of a Decision Function" (abstract), ECONOMETRICA, Vol. 18, April, 1950, p. 183.

CRAMÉR, H. [1], "On the Mathematical Theory of Risk," *Försäkringsaktiebolaget Skandias Festskrift*, Stockholm: Centraltryckeriet, 1930, pp. 7–84.

DOMAR, E., AND R. A. MUSGRAVE [1], "Proportional Income Taxation and Risk-Taking," *Quarterly Journal of Economics*, Vol. 58, May, 1944, pp. 388–422.

FELLER, W. [1], *An Introduction to Probability Theory and its Applications*, Vol. I, New York: John Wiley and Sons, 1950, 419 pp.

DE FINETTI, BRUNO [1], "La Prévision: Ses Lois Logiques, Ses Sources Subjectives," *Annales de l'Institut Henri Poincaré*, Vol. 7, 1937, pp. 1–68.

FISHER, I. [1], *The Nature of Capital and Income*, New York and London: The Macmillan Co., 1906, 427 pp.

——— [2], *The Theory of Interest*, New York: The Macmillan Co., 1930, 550 pp.

FISHER, R. A. [1], "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London*, Part A, Vol. 222, 1922, pp. 309–368.

FRÉCHET, M. [1], "L'Estimation Statistique des Paramètres" (abstract), ECONOMETRICA, Vol. 16, January, 1948, pp. 60–62.

FRIEDMAN, M. [1], "Lange on Price Flexibility and Employment: A Methodological Criticism," *American Economic Review*, Vol. 36, September, 1946, pp. 613–631.

FRIEDMAN, M., AND L. J. SAVAGE [1], "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, Vol. 56, August, 1948, pp. 279–304.

GRAAF, J. DE V., AND W. J. BAUMOL [1], "Three Notes on 'Expectations in Economics' II," *Economica*, New Series, Vol. 16, November, 1949, pp. 338–342.

HARDY, C. O. [1], *Risk and Risk-Bearing*, Chicago: University of Chicago Press, 1923.

HART, A. G. [1], *Anticipations, Uncertainty, and Dynamic Planning*, New York: Augustus M. Kelley, 1951, 98 pp.

——— [2], "Risk, Uncertainty, and the Unprofitability of Compounding Probabilities," in *Studies in Mathematical Economics and Econometrics*, O. Lange, F. McIntyre, and T. O. Yntema, eds., Chicago: University of Chicago Press, 1942, pp. 110–118.

HICKS, J. R. [1], "The Theory of Uncertainty and Profit," *Economica*, Vol. 11, May, 1931, pp. 170–189.

——— [2], "Application of Mathematical Methods to the Theory of Risk" (abstract), ECONOMETRICA, Vol. 2, April, 1934, pp. 194–195.

KALECKI, M. [1], "The Principle of Increasing Risk," *Economica*, New Series, Vol. 4, November, 1937, pp. 440–447.

KEYNES, J. M. [1], *A Treatise on Probability*, London: Macmillan and Co., 1921, 466 pp.

KNIGHT, F. H. [1], *Risk, Uncertainty, and Profit*, New York: Houghton Mifflin and Co., 1921. (Reprinted, London: London School of Economics and Political Science, 1946, 381 pp.)

KOLMOGOROV, A. N. [1], "Grundbegriffe der Wahrscheinlichkeitsrechnung," *Ergebnisse der Mathematik und Ihre Grenzgebiete*, Vol. 2, 1933, translated as *Foundations of the Theory of Probability*, New York: Chelsea Publishing Co., 1950, 70 pp.

VON KRIES, J. [1], *Die Principien der Wahrscheinlichkeitsrechnung*, Freiburg, 1886, 298 pp.

LANGE, O. [1], *Price Flexibility and Employment*, Cowles Commission Monograph 8, Bloomington, Ind.: Principia Press, 1944, 114 pp.

LEHMANN, E. L. [1], "On Families of Admissible Sets," *Annals of Mathematical Statistics*, Vol. 18, March, 1947, pp. 97–104.

MARSCHAK, J. [1], "Money and the Theory of Assets," ECONOMETRICA, Vol. 6, October, 1938, pp. 311–325.

——— [2], "Von Neumann's and Morgenstern's New Approach to Static Economics," *Journal of Political Economy*, Vol. 54, April, 1946, pp. 97–115.

——— [3], "Role of Liquidity under Complete and Incomplete Information," *American Economic Review, Papers and Proceedings*, Vol. 39, 1949, pp. 182–195.

——— [4], "Rational Behavior, Uncertain Prospects, and Measurable Utility," ECONOMETRICA, Vol. 18, April, 1950, pp. 111–141.

MARSHALL, A. [1], *Principles of Economics*, eighth ed., New York: The Macmillan Co., 1948, 871 pp.

MENGER, K. [1], "Das Unsicherheitsmoment in der Wertlehre," *Zeitschrift für Nationalökonomie*, Vol. 4, No. 4, pp. 459–485.

VON MISES, R. [1], "On the Foundation of Probability and Statistics," *Annals of Mathematical Statistics*, Vol. 12, June, 1941, pp. 191–205.

VON NEUMANN, J., AND O. MORGENSTERN [1], *Theory of Games and Economic Behavior*, second ed., Princeton: Princeton University Press, 1947, 641 pp.

Neyman, J., and E. S. Pearson [1], "The Testing of Statistical Hypotheses in Relation to Probabilities A Priori," *Proceedings of the Cambridge Philosophical Society*, Vol. 29, 1933, pp. 492–510.

Niehans, J. [1], "Zur Preisbildung bei ungewissen Erwartungen," *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, Vol. 84, No. 5, 1948, pp. 433–456.

Palm, C. [1], "The Distribution of Repairmen in Servicing Automatic Machines" (in Swedish), *Industrietidningen Norden*, Vol. 75, 1947, pp. 75–80, 90–94, 119–123.

Pigou, A. C. [1], *The Economics of Welfare*, second ed., London: Macmillan and Co., 1924, 783 pp.

Ramsey, F. P. [1], "Truth and Probability," in *The Foundations of Mathematics and Other Logical Essays*, London: K. Paul, Trench, Trubner and Co., 1931.

Rubin, H. [1], "The Existence of Measurable Utility and Psychological Probability" (hectographed), Cowles Commission Discussion Paper: Statistics No. 331, July, 1949, 5 pp.

——— [2], "Postulates for the Existence of Measurable Utility and Psychological Probability" (abstract), *Bulletin of the American Mathematical Society*, Vol. 55, November, 1949, pp. 1050–1051.

Samuelson, P. A. [1], *Foundations of Economic Analysis*, Cambridge, Mass.: Harvard University Press, 1947, 447 pp.

Savage, L. J. [1], "The Role of Personal Probability in Statistics" (abstract), Econometrica, Vol. 18, April, 1950, pp. 183–184.

——— [2], "The Theory of Statistical Decision," *Journal of the American Statistical Association*, Vol. 46, March, 1951, pp. 55–67.

Shackle, G. L. S. [1], *Expectation in Economics*, Cambridge, England: Cambridge University Press, 1949, 146 pp.

——— [2], "A Non-Additive Measure of Uncertainty," *Review of Economic Studies*, Vol. 17, No. 1, 1949–50, pp. 70–74.

Törnqvist, L. [1], "On the Economic Theory of Lottery Gambles," *Skandinavisk Aktuarietidskrift*, Vol. 28, Parts 3 and 4, 1945, pp. 228–246.

Tintner, G. [1], "The Theory of Choice under Subjective Risk and Uncertainty," Econometrica, Vol. 9, July–October, 1941, pp. 298–304.

——— [2], "The Pure Theory of Production under Technological Risk and Uncertainty," Econometrica, Vol. 9, July–October, 1941, pp. 305–312.

——— [3], "A Contribution to the Non-Static Theory of Choice," *Quarterly Journal of Economics*, Vol. 56, February, 1942, pp. 274–306.

——— [4], "A Contribution to the Non-Static Theory of Production," in *Studies in Mathematical Economics and Econometrics*, O. Lange, F. McIntyre, and T. O. Yntema, eds., Chicago: University of Chicago Press, 1942, pp. 92–109.

——— [5], "Foundations of Probability and Statistical Inference," *Journal of the Royal Statistical Society*, Series A, Vol. 112, Part 3, 1949, pp. 251–279.

Wald, A. [1], "Contributions to the Theory of Statistical Estimation and Testing Hypotheses," *Annals of Mathematical Statistics*, Vol. 10, December, 1939, pp. 299–326.

——— [2], *Statistical Decision Functions*, New York: John Wiley and Sons, 1950, 179 pp.