Behavior under Uncertainty and Its Implications for Policy

K. J. Arrow

February 83

Organizational Efficiency

Stanford University Center for Research on Organizational Efficiency

UNCLASSIFIED N00014-79-C-0685
BEHAVIOR UNDER UNCERTAINTY AND ITS IMPLICATIONS FOR POLICY
by
Kenneth J. Arrow

Technical Report No. 399
February 1983

A REPORT OF THE
CENTER FOR RESEARCH ON ORGANIZATIONAL EFFICIENCY
STANFORD UNIVERSITY
Contract ONR-N00014-79-C-0685, United States Office of Naval Research

THE ECONOMICS SERIES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Fourth Floor, Encina Hall
Stanford University
Stanford, California
94305

This document has been approved for public release and sale; its distribution is unlimited.
A key tool in the modern analysis of policy is benefit-cost analysis. Though its origin goes back to the remarkably prescient paper of Dupuit, 1884, its theoretical development came much later, after the "marginal revolution" of the 1870's, and its practical application really dates only from the period after 1950. The underlying theory is that of notion of economic surplus, to which, after Dupuit, such major figures as Marshall, Pareto, Hotelling, Allais, and Debreu have contributed: for a remarkable synthesis, see Allais, 1981.

Without going into technical details, the essential steps in the actual calculation of a surplus depend on using choices made in one context to infer choices that might be made in different contexts. If we find how much individuals are willing to pay to reduce time spent in going to work by one method, e.g., buying automobiles or moving closer to work, we infer that another method of achieving the same saving of time, e.g., mass transit or wider roads, will be worth the same amount. Frequently, indeed, we extrapolate, or interpolate; if it can be shown that the average individual will pay $1,000 a year more in rent to reduce his or her transit time by 30 minutes, we infer that a reduction of 15 minutes is worth $500. This is all very much according to Dupuit's reasoning; he would value an aqueduct by the amount that

*This research was supported by the Office of Naval Research Grant ONR-N00014-79-C-0685 at the Center for Research on Organizational Efficiency, Institute for Mathematical Studies in the Social Sciences, Stanford University.
individuals would be willing to pay for the water to be transported in it (and vice versa, if the opposite inference is useful).

The assumption that choices made under different conditions are consistent with each other is then essential to the practice of benefit-cost analysis. The elaboration of these consistency conditions leads to the rationality postulates of standard economic theory. In the usual formulation, we postulate that all choices are consistent with an ordering, a transitive and complete relation, and both our theory and our practice are based on this assumption. We know, of course, that even with these assumptions there are ambiguities in the inferences that can be drawn from empirical observations to policy choices, mostly because of the so-called income effects, a point on which Walras already criticized Dupuit. But in this paper, I will not be concerned with this last set of issues.

We now have a new kind of benefit-cost analysis, namely benefit-risk analysis. The risk of a disutility is itself a cost and a proper subject for measurement along with the direct costs of the usual resource-using type. Similarly, a reduction in risk is to be counted as a benefit. This is true even if individuals are risk-neutral, since we would still want to count the expected value of the risk, in the presence of risk aversion, there is still an additional cost or benefit, as the case may be.

Our current interest in risk-benefit analysis has been largely stimulated by concern with health hazards. In terms of public attention, though not in actuality, it is the risks associated with the
operation of nuclear power plants that have appeared at the forefront. At a fundamental level, the issues in benefit-risk analysis are not different from those in more familiar welfare comparisons. Again, willingness to pay either for benefits (reduced risks) or to avoid costs (increased risks) is a crucial element. Again, it can in principle be measured by seeking out comparable situations. Thus, if air pollution control results in decreased probability of death, then one way of measuring the benefit is to find out what individuals are willing to pay to decrease this probability in other contexts. A standard method is to compare wages in industries with different occupational risks. After controlling for other factors, we find that in general hazardous occupations have higher wage rates. In equilibrium, this means that workers are indifferent at the margin between the two occupations at the given wage levels. Hence, the difference in wages can be regarded as compensation for the difference in probability of death. (Needless to say, I am ignoring many obvious complications, for example, the risks of non-fatal injury.) This provides a measure of willingness to pay for reduced probability of death, to be used in evaluating the reduction in risk due to air pollution (or automobile safety measures or anything similar).

The similarity in principle can be accepted to justify similarity in practice provided we accept some theory of rationality in individual behavior under uncertainty, which is precisely what is frequently questioned. It is this theme, the implications of current research on decision-making in the presence of risk for benefit-risk analysis, that I want to pursue today.
It is necessary to call attention to one important matter on which the analogy to the case of certainty is necessarily loose, that is, the establishment of probability judgments. Benefit-cost analysis under certainty of course requires not only measures of willingness to pay but also measures of the costs of alternative policies. The analogue under uncertainty is measurement of the probabilities of different possible outcomes for each possible policy. Thus, we need the probability of death associated with each possible level of atmospheric pollution or for each possible siting and design of nuclear reactors.

Now no probability can, strictly speaking, be known from a finite sample. In many cases, the evidence is very limited indeed, so that this condition is a practical as well as a theoretical limitation. In many of the most difficult situations, those with high risks but very low probabilities, such as the possibility of a nuclear core meltdown, the evidence on the relevant uncertainty is extremely small. It may all be indirect.

This raises a new and philosophically difficult problem of rationality. Just as we need some kind of rationality hypotheses for measuring willingness to pay, under uncertainty as under certainty, we need rationality hypotheses about probability judgments. These are usually supplied in theory by the theory of subjective probabilities and applications of Bayes's Theorem. But what if individuals do not make their probability judgments in this manner?

One last orientation is needed. Who is doing the benefit-risk analysis? I want to ignore all the additional complications due to the
difficulties of social choice, so I will suppose that there is a representative individual. Equally, however, I do not want to reduce the solution to the uninteresting proposition that the individual should do what he or she wants. Instead, I will suppose that the analyst is a professional adviser to the individual. Both the client and the economist expect that the latter will have something to contribute by way of clarification, even though ultimately it is the client's interests that are to be served--but the client's interests as properly interpreted, not mere expression of first thoughts.

A little intellectual history will be helpful. In 1952, at a conference on the foundations of risk-bearing held by the Centre National de la Recherche Scientifique of France, Allais, 1953, and I (Arrow, 1953) presented independently formulated models incorporating risk-bearing into the theory of general equilibrium (Allais had also presented his theory a year earlier at a meeting of the Econometric Society). As a good Paretian, Allais followed the lead of his earlier work on welfare economics, Allais, 1945, and perceived and expressed the welfare optimality that necessarily underlies any general competitive equilibrium. This is an essential first step in a benefit-risk analysis. Properly applied, it can be used to derive the shadow prices which yield the first approximations to the appropriate measures of surplus.

There are several differences between Allais' model and mine, most not very relevant to the present discussion. One that was much discussed at the conference later turned out to be irrelevant. In my paper, I assumed that individuals maximize expected utility. I accepted
the Bernoulli, 1738, theory as it had been updated by von Neumann and Morgenstern, 1947, and by Savage, 1954. Allais, as is well known, subjected that theory to very strong attack. His own view amounted to a general ordinalist position; there was an ordering on probability distributions, not necessarily linear in probabilities. This position had been advanced earlier by Hicks, 1931, though he had done little with it.1/

A few years later, Gerard Debreu, 1959, Chapter 7, showed that the two models could be synthesized. A theory of general equilibrium in contingent contracts did not require the Bernoulli hypothesis; it was consistent with any utility function over the outcomes. Debreu also extended the theory to paths over time, in which the uncertainties are realized successively.

Policy analysis with regard to risk, as in the case of certainty, is necessitated by failures of the competitive mechanism, that is, externalities and public goods. An individual living near a nuclear plant cannot buy safety for himself or herself alone; only collective safety can be obtained. Similarly, air pollution cannot for well-known reason be obtained without collective action; no assignment of property rights will permit the market to achieve an optimal allocation. We use the general equilibrium model to simulate a market; what would individuals be willing to pay at the margin for changes in the externalities if they could be implemented as commodities?

I will consider four doubts about rational behavior in the presence of uncertainty which have arisen from recent empirical research and ask about their implications for the practice of benefit-risk analysis:
(1) questions about the expected-utility theory;
(2) miscalculations of probabilities;
(3) preference reversals; and
(4) framing.

1. The Expected Utility Hypothesis

To be concrete, let us consider the expected utility hypothesis applied to policies aimed at affecting mortality. To bring out the essence, I consider only the simplest possible case. A living individual receives a utility from consumption if alive but zero utility if dead. Let,

\[ p = \text{probability of survival}, \]
\[ U = \text{utility}, \]
\[ c = \text{consumption}. \]

Then expected utility is,

(1.1) \[ pU(c). \]

We may think of some policy which increases \( p \) but requires resources and therefore reduces \( c \). The willingness-to-pay (WTP) is then the slope of the curves on which (1) is constant, that is, the amount of consumption that an individual is willing to give up per unit probability of survival while keeping expected utility constant. We see easily that,

(1.2) \[ \text{WTP} = \frac{U(c)}{pU'(c)}. \]
This expression has the dimensions of consumption per unit life and therefore can be and is frequently referred to as the "value of life." However, it is not what an individual would pay for the certainty of life as against the certainty of death; it is a marginal evaluation of a small change in the probability of life. Since what an individual would pay for the certainty of life is limited by initial wealth, the WTP is apt to be a good deal larger.

Suppose the Bernoullli hypothesis is false, but individuals are rational in the weaker sense of Allais and Hicks; there is an ordering of probability distributions. In the present simple context, this means that there is a utility function which depends on \( c \) and \( p \) (and is defined only up to monotone transformations):

\[(1.3) \quad U(c,p) \quad .\]

Again, there is a trade-off between the two variables,

\[(1.4) \quad \frac{U_p}{U_c} \quad .\]

When a real benefit-risk analysis is done, what data are used? Suppose, as suggested above, that willingness to pay is estimated from the wage differentials to be found in riskier jobs, as in Thaler and Rosen, 1976, or Viscusi, 1979. Now, the probability of survival for one more year (which is what is to be compared with annual wages) depends on many factors, of which occupational safety is only one and not a major one. The observed risk differential is therefore small, so that the observed ratio of wage differential to risk differential is really a
measure of (1.4). In fact, when one looks closely, the empirical material made no real use of the Bernoulli hypothesis as embodied in (1.1) and (1.2).

What in fact is gained by the stronger expected-utility hypothesis? It is really the ability to extrapolate over large changes in p. But in practice, any feasible policy, whether in nuclear power safety, biomedical research, or occupational safety, will have only a relatively slight effect on the probability of survival over a year or other relevant period. Therefore, the strength of the Bernoulli hypothesis is never employed in practice.

The argument amounts to saying that every general utility functional for probability distributions, if differentiable, can be regarded as approximately linear in the probabilities, if we are not considering large changes in them. This is precisely the idea so beautifully and richly developed by Machina, 1982.

I cannot however, leave this subject without another restatement of the Allais problem. It is hard to believe that the paradox will occur when the alternatives are laid out in a sufficiently transparent fashion. Let us introduce a more specific temporal structure.

Suppose there are 3 time periods. At time 0, there is a chance move; it yields a payment of 1 monetary unit with probability .89. If the complementary event occurs, then at time 1 the subject is offered a choice. He or she can take 1 unit (with certainty), to be paid at time 2, or go on to time 2 and face a gamble yielding 2.5 with probability 10/11 and 0 with probability 1/11. At time 1, the possibility of an
immediate payment of 1 with termination is now in the past, and there can be no question that the individual faces and considers only the second gamble, as against certainty. Presumably, the individual will usually choose the gamble. Now suppose at time 0, the individual is asked: if the complementary event were to occur at the chance move coming up, would you choose the certainty or the gamble? This is clearly the same decision as in the first story; it requires only a certain imagination. Yet in this form, a hypothetical choice of certainty amounts to choosing 1 with certainly as against a distribution of 2.5 with probability .10, 1 with probability .89, and 0 with probability .01.

In other words, all that is required is understanding a hypothetical choice and calculating probabilities correctly.

Now it may be that rendering the decision tree transparent may be all that is involved, in which case it suggests that the real issue is one of framing, a point to which I will return below.

2. Miscalculation of Probabilities

A more serious problem than the nonlinearity of the utility function is the calculation of the probabilities to be used in estimating the risk. There are two issues here. One is not deep conceptually; it is simply that an individual will in general simply not possess all the information available to society as a whole. The probabilities used should of course be conditional on all the information available, if the information can be assembled at a cost which is negligible compared with
measure of (1.4). In fact, when one looks closely, the empirical material made no real use of the Bernoulli hypothesis as embodied in (1.1) and (1.2).

What in fact is gained by the stronger expected-utility hypothesis? It is really the ability to extrapolate over large changes in $p$. But in practice, any feasible policy, whether in nuclear power safety, biomedical research, or occupational safety, will have only a relatively slight effect on the probability of survival over a year or other relevant period. Therefore, the strength of the Bernoulli hypothesis is never employed in practice.

The argument amounts to saying that even a general utility functional for probability distributions, if differentiable, can be regarded as approximately linear in the probabilities, if we are not considering large changes in them. This is precisely the idea so beautifully and richly developed by Machina, 1982.

I cannot however, leave this subject without another restatement of the Allais problem. It is hard to believe that the paradox will occur when the alternatives are laid out in a sufficiently transparent fashion. Let us introduce a more specific temporal structure.

Suppose there are 3 time periods. At time 0, there is a chance move; it yields a payment of 1 monetary unit with probability .89. If the complementary event occurs, then at time 1 the subject is offered a choice. He or she can take 1 unit (with certainty), to be paid at time 2, or go on to time 2 and face a gamble yielding 2.5 with probability 10/11 and 0 with probability 1/11. At time 1, the possibility of an
the improved expected benefits. Hence, there is an externality with regard to information-gathering. Therefore, if the information is assembled the expert opinion should be used to form probabilities. Presumably, any rational individual would agree to this and would voluntarily defer, as he or she does to a physician or other professional.

This observation does create some problems, not at the normative level but at that of interpreting observed choice behavior as a measure of willingness to pay to avoid risks. What is relevant is the ratio of wage difference to difference in probability of death or injury as perceived by the individuals involved. If they act not on the probabilities as estimated by a national collection of statistics but on those estimated by themselves from much more limited data, it is the latter probabilities that should be used as a divisor. There is considerable theoretical and empirical evidence in the case of occupational hazards that individuals are influenced by their own experiences (Viscusi, 1979, Chapters 4 and 13). This is consistent with the view that they have little knowledge of general injury rates and therefore condition their probabilities heavily on their own experiences.

A deeper question is raised by the well known observations mostly by psychologists that even in situations where Bayes' Theorem is clearly applicable, individuals do not use it correctly; for reports on such studies, see Tversky and Kahneman, 1974. In most of their experiments, too little weight was given by subjects to the prior information; individuals were overly influenced by the current data. This result is consistent with other studies in different fields. Eddy, 1980, showed
that physicians in relying on diagnostic tests did not take adequate account of the underlying prevalence of the disease in forming their judgments. Thus, if a test gave a probability .9 of detecting cancer if it were there and a probability .9 of a negative response if there were no cancer, it would be regarded as highly reliable. Yet the prevalence of cancer is only about .1, which is thus the prior probability of cancer in a random choice from the population. A simple application of Bayes' Theorem shows that the probability of cancer given a positive response on the test is .5, far less than most physicians would believe.

The misuse of Bayes' Theorem is at least compatible with the evidence for volatility in the securities and futures markets. Since the value of a long-lived bond, stock, or futures contract is ultimately dependent upon a great many events which will occur in the future, it should be unresponsive to any particular piece of new information. These markets, however, are notoriously volatile, with large movements in a single day. This has been argued many times for the futures markets; for a summary of the evidence, see Cagan, 1981. A very rigorous analysis for the bond and stock markets (Shiller, 1979, 1981) has shown the incompatibility of observed behavior with rational expectations models, at least in a simple form. For more extended discussion of these misperceptions, see Arrow, 1982.

The extent to which the average person exaggerates the risks in a possible accident to nuclear power plants is of course well known. (I hasten to add that I am not an unreserved admirer of nuclear power. The risks to the plant itself, as exemplified by the Brown's Ferry and Three
Mile Island accidents, plus the costs of construction, themselves increased by safety precautions, make the economics of nuclear power very doubtful. Indeed, the main case for nuclear power are the social costs of coal-fired power plants in the form of atmospheric pollution caused by combustion and carbon dioxide effects on the world's climate.)

What is the normative or policy implication of this propensity to irrational judgments about uncertainty? Here, I feel strongly we must invoke the appropriate role of the expert. I have postulated a relation of professional to client, and it is certainly in judgments of reality and probability that this professional concern is most appropriate. The normative judgment may and should respect the utility functions (linear or nonlinear) of the public being advised, but it should certainly use probability judgments based both on the maximum of information and the maximum of correct statistical and probabilistic logic.

The two problems discussed thus far, the more general ordinal theory of choice among probability distributions as against the expected-utility hypothesis and the miscalculation of probabilities, have been made much of by those, for example at the conference at which this paper is presented, who think of themselves as revolutionaries against an established (though rather recently established) orthodoxy. Beware! These arguments are those of the moderate revolutionaries, the Girondins or the Mensheviks. The cognitive psychologists have found evidence for worse traps; if the implications are as they seem, it is hard to see how any form of benefit-risk analysis can survive.
3. Preference Reversal

So far, transitivity has been unquestioned; and transitivity is essential to any type of benefit-cost analysis; the substitution of compensations for costs or risks depends essentially on the (usually unstated) transitivity of indifference. I have not checked the literature, but I believe that experiments do not even verify transitivity fully even in the case of certainty. It was for this reason that economists and psychologists developed notions of stochastic orderings (see, e.g., Davidson and Marschak, 1959). But experimental studies of choice under uncertainty has revealed what at the present appear to be a less remedial form of intransitivity.

I refer to the well-known phenomenon of preference reversal, first identified by Lichtenstein and Slovic, 1971. Suitably chosen pairs of gambles can be found with the following characteristics: When subjects are asked to choose between the two, they express a preference for one. But when asked to state the amount of money which, if given with certainty, would be indifferent to each gamble, the amounts chosen are in opposite order to the expressed preferences.

This result is so upsetting, indeed to almost any theory of choice under uncertainty, that the experiments were carefully replicated by Grether and Plott, 1979. They varied the experiments in ways designed to test various explanations (e.g., cost of information processing) which would preserve transitivity. Not only were the original results confirmed, but no simple rational explanation could be found.
This work does seem to be a major barrier to the use of risk-cost tradeoffs from one area in measuring benefits and costs in another. I can only offer some observations derived from earlier work in consumer demand theory as a partial solution.

The comparisons in the preference reversal experiments are global or long-range rather than local. The identification of certainty equivalents requires comparison of two alternatives, one risky, one certain, which are far from each other in any reasonable metric. This distinction was considered also in the theory of consumer's choice under certainty; it is the essence of what has been called the integrability problem. There are many variations in the literature. One is the proposed theory advanced by Hicks and Allen, 1934, and Georgescu-Roegen, 1936: At each point in commodity space, there is a local indifference map (hence, transitive and complete ordering), but it is an additional assumption that the local indifference maps integrate into a global indifference map permitting indifference judgments across large differences in alternatives.

The very meaning of optimality and therewith the meaning of a benefit-risk analysis as a basis for policy is in principle undermined if comparisons are only local. One might conceivably have a series of local improvements which cycle.

In practice, though, it could be argued, though not with entire conviction, that the possibilities of paradox do not really arise. As we have already argued with relation to policies which affect the probability of death, any feasible changes are apt to be small. Hence, only
the local indifference maps are relevant. In that case, the theoretical possibility of cycling will not be realized. To put the matter another way, the willingness-to-pay data used in benefit-risk analysis does not really measure comparison of gambles with certainties. Rather, it compares gambles with small differences in probabilities and stakes. Therefore, the preference reversal phenomenon need not occur in the choices which are the basis for benefit-risk analysis.

4. Framing

The most damning criticism of risk-benefit analysis form experiments is the evidence for what Tversky and Kahneman, 1981, have called, "framing." An element of rationality, so obvious to the analyst as to pass almost unnoticed, is its extensionality, to use the language of logic. That is, if a choice is to be made from a set of alternatives, the choice should depend only on the membership in the set and not on how the set is described. If I have to choose which night of next week I will go to a play, it should not matter if each alternative is labeled by the day of the week or the numbered day of the month. If my budget permits me to divide $1,000 between housing costs and other expenditures, my alternatives can be identified indifferently in terms of either of the two kinds of costs.

Yet the lesson of framing experiments is precisely that these invariances do not hold. How the choice question is framed affects the choice made.2/ Let me draw a dramatic illustration from a paper on the choice of medical therapy by McNeil, Pauker, Sox, and Tversky, 1982. McNeil and some of her colleagues have had a program, which economists
and decision theorists should applaud, of letting the patients' values affect medical decision-making. In this study, a comparison was being made between two therapies, surgery and radiation therapy, for the treatment of certain forms of cancer. A therapy defines a probability distribution of length of survival. In general, surgery has a positive risk of mortality during the operation, for which there is no counterpart in radiation therapy, but it has a longer expected survival even when this risk is taken into account. Different groups of individuals, including a group of physicians, were presented with the probabilities of survival during treatment, for one year, and for five years for each of the two therapies. With these data, 84% of the physicians preferred surgery, 16% radiation therapy. Then another group was presented with the same data presented differently: the probabilities of dying at each stage were given instead of the probabilities of survival. At each stage, the probability of dying is 1 minus the probability of survival, so the two formulations are not merely logically equivalent but can be transformed into each other by a trivial calculation. Yet the proportion of physicians choosing surgery over radiation therapy dropped from 84% to 50%.

I leave the implications of framing for benefit-risk analysis as an open problem. Economists would tend to argue that the choices made in the market, where the stakes to the individual are high, reflect the correct choice of frame. But this is probably too complacent a view. It may well be true that the individual makes different trade-offs in contexts which, to the analyst, appear to be identical. But this is a deep topic for further study.
Footnotes

1/ In the development of his specific general equilibrium model, Allais assumed that all distributions were normal, and therefore individuals were assumed to order distributions according to their means and variances. However, I take this to be a particular application, not the underlying general principle.

2/ The rest of this paragraph is drawn, with slight modification, from Arrow, 1982.


Dupuit, J. [1844], "De la Mesure de l'Utilité des Travaux Publics," Annales des ponts et Chaussées, 2ième Série, 8.


Reports in this Series


378. "Unemployment Equilibrium with Stochastic Rationing of Supplies" by Ho-mou Wu.


380. "On the NTU Value" by Robert J. Aumann.


382. "Informational Equilibrium" by Robert Kast.

383. "Cooperative Oligopoly Equilibrium" by Mordecai Kurz.


390. "Aggregates, Activities and Overheads" by W.M. Gorman.


392. "Efficiency and Fairness in the Design of Bilateral Contracts" by S. Honkapohja.

393. "Diagonality of Cost Allocation Prices" by L.J. Mirman and A. Neyman.


396. "Dominance-Solvability and Cournot Stability" by Herve Moulin.

Reports in this Series

398 "Generalizations of the Censored and Truncated Least Absolute Deviations Estimators" by J.L. Powell.

399 "Behavior Under Uncertainty and its Implications for Policy" by K.J. Arrow.