

Regret in Decision Making under Uncertainty

Author(s): David E. Bell

Source: *Operations Research*, Vol. 30, No. 5 (Sep. - Oct., 1982), pp. 961-981

Published by: INFORMS

Stable URL: <https://www.jstor.org/stable/170353>

Accessed: 03-02-2020 12:26 UTC

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/170353?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Operations Research*

Regret in Decision Making under Uncertainty

DAVID E. BELL

Harvard University, Cambridge, Massachusetts

(Received October 1981; accepted April 1982)

Evidence exists that people do not always make decisions involving uncertain monetary rewards as if they were maximizing expected utility of final assets. Explanations for this behavior postulate that the cognitive demands of consistency to such a theory are too great. However, situations exist in which more than mental shortcuts are involved and these anomalies raise questions about expected utility theory as a guide to behavior. This paper explores the possibility that expected utility theory appears to fail because the single outcome descriptor—money—is not sufficient. After making a decision under uncertainty, a person may discover, on learning the relevant outcomes, that another alternative would have been preferable. This knowledge may impart a sense of loss, or regret. The decision maker who is prepared to tradeoff financial return in order to avoid regret will exhibit some of the behavioral paradoxes of decision theory. By explicitly incorporating regret, expected utility theory not only becomes a better descriptive predictor but also may become a more convincing guide for prescribing behavior to decision makers.

VON NEUMANN and Morgenstern [1947] showed a particular set of behavioral principles for decision making under uncertainty to be equivalent to choice by maximization of expected utility of final assets. Many decision analysts feel that the axioms are so compelling and reasonable that the situations in which decision makers violate the axioms consistently are referred to as “paradoxes.” Because the role of decision analysis is to aid decision makers rather than to dictate to them, instances of behavior in which the principles of von Neumann and Morgenstern are knowingly broken are worthy of study for the insight they may give in improving decision analysis.

This paper hypothesizes that many of the axiom violations may be attributed to a desire by decision makers to avoid consequences in which the individuals will appear, after the fact, to have made the wrong decision even if in advance, the decision appeared correct with the information available at the time. This consequence of decision-making under uncertainty will be termed decision regret. This term includes also the euphoria and self-congratulation associated with appearing to have

Subject classification: 94 decision regret, 851 utility theory paradoxes, 852 final assets and regret as multiple attributes.

made a good decision. Associated with each decision consequence illustrated in this paper will be two measures of satisfaction—monetary assets and quantity of regret. A multiattribute utility function that incorporates these two considerations will be applied to decision situations to study the implications of regret for behavior.

The most celebrated and long-standing paradox is due to Allais [1953]. He showed that although most people select a prize of [\$1 million for sure] rather than a gamble giving [a 10% chance at \$5 million, an 89% chance at \$1 million and a 1% chance of getting nothing], a majority of people prefer a gamble offering [a 10% chance at \$5 million and a 90% chance at nothing] rather than a gamble of [an 11% chance at \$1 million and an 89% chance at nothing].

As is well known, this behavior contradicts the use of a utility function over assets alone. The first choice implies, with figures in millions of dollars, that

$$u(1) > 0.10 u(5) + 0.89 u(1) + 0.01 u(0)$$

whereas, the second choice implies that

$$0.10 u(5) + 0.90 u(0) > 0.11 u(1) + 0.89 u(0)$$

and these two inequalities are contradictory.

There are a number of explanations for this behavior. Many ascribe the inconsistency as being brought on by mental shortcuts, such as dismissing the difference between 10% and 11% as negligible. However, there are people who, even after seeing the expected utility analysis, still wish to act in accordance with their original preferences.

This behavior can be explained by decision regret. A decision maker who takes the gamble over the sure \$1 million may feel absolutely devastated if the 1% chance of getting nothing results. It would be considered normal, if not economically rational, for such a decision maker to feel angry and perhaps depressed that he or she could have been so stupid or greedy as to pass up a sure \$1 million. In the second choice between gambles there is no equivalent endpoint. If the chosen lottery results in no reward, the decision maker may feel that nothing was also the likely result of the other lottery. Thus, regret plays little or no role in the second choice of Allais' paradox but serves to downweight severely the gamble in the first choice.

Hogarth ([1980], p. 72) makes the same analysis of the paradox and terms the problem psychological regret. The effect may arise without a psychological cause since the evaluation of others, one's boss for example, may be an important consideration and hence the term decision regret is more encompassing. It is also important to emphasize that the regret arises from having made a decision. The role of regret as a criterion for decision was widely explored at one time, but was largely dismissed

because its use had undesirable properties, such as intransitivity (see Luce and Raiffa [1956], p. 280). The approach in this paper differs from these early uses of regret in two ways. One is that regret is measured here as the difference in value between the assets actually received and the highest level of assets produced by *other* alternatives. This definition allows both positive and negative values of regret. The second distinction is that minimizing the maximum regret is not the sole criterion for decision: rather, the disadvantages of regret are traded off against the value of assets received.

To illustrate the kind of thinking employed in this paper, consider the following two discussions.

Regret Symmetry

A farmer has a crop whose value at harvest time, two months from now, is uncertain. Should he accept a recent fixed-price offer for his crop in which the buyer guarantees a specific price per bushel for the grain, to be paid at harvest time? To simplify the example, suppose that the price at harvest time will be either \$3 or \$7 per bushel with equal probability. The current offer, the only one the farmer expects to receive, is for \$5 per bushel. A traditional decision analysis shows that the fixed price offer has no risk because it is \$5 for sure, whereas waiting does have risk.

How might the farmer think about the risks involved? If he sells now for \$5 and the price at harvest time proves to be \$7, he will feel that he lost \$2 by his action. If the price declines to \$3, he will feel pleased at being up by \$2. Compare that alternative with choosing to wait. If the price declines to \$3 he will regret not selling it for the \$5 and feel like his decision cost him \$2. If the price is \$7 at harvest time he will feel happy that his acumen improved his position by \$2 per bushel. With either option he is equally likely to regard himself as up or down by \$2. The alternatives are symmetric in terms of their potential for regret. In this sense the alternatives have equal risk.

Dominance

A portfolio manager, having decided that a forthcoming election will have an effect on asset values, calculates that should candidate *A* win, the value of her portfolio would appreciate by 10% overnight. Should candidate *B* win, her position would depreciate by 6%. She believes that each candidate has a 50-50 chance of victory. She assigns her research team to find alternative investment postures in light of the election.

Consider two scenarios. In the first, the research team reports on an alternative that produces a gain of 11% if *A* wins with a decline of only 5% if *B* wins. Clearly this is a superior alternative, which she would assuredly take. In the second scenario the research team reports the

availability of a mutually exclusive investment offering an 11% return if *B* wins, but a decline of 5% if *A* wins. Now the handsome improvement in performance should *B* win is very tempting, but what if *A* wins? She will have lost 15% of the portfolio's value by making the switch, a prospect that may cause her to have serious reservations. The two scenarios are summarized in Table I. Notice that in scenario 1 the *decision* carries no risk of regret for it results in a sure gain of 1% of asset value. In scenario 2 the *decision* will either gain her 17% or lose her 15%. Viewed from this perspective, the switch seems much less desirable in the second scenario, yet a traditional decision analysis shows the two new options to be equivalent given the probabilities involved.

People who think in the fashion of these examples may be grateful for the discipline that expected utility theory over assets could bring to their judgments—or they may find the theory somewhat lacking.

TABLE I
DOMINANCE EXAMPLE

Election Winner Candidate	Assumed Probability of Victory	Scenario 1		Scenario 2	
		Status Quo	New Option	Status Quo	New Option
A	0.50	10%	11%	10%	-5%
B	0.50	-6%	-5%	-6%	11%

This paper will treat regret as a second attribute of concern to the decision maker and incorporate the appropriate tradeoffs between regret and final assets into a two-attribute utility function which will then be used to examine comparisons between lotteries over assets.

A number of simplifications are made in the domain of lottery comparisons to be examined.

1. *Objective probabilities.* Smith [1966], commenting on Ellsberg's paradox (Ellsberg et al. [1961]), has suggested that people may have different feelings about risk aversion based on whether the underlying probabilities are objectively or subjectively known. While regret does offer an explanation for Ellsberg's paradox (Bell [1981]), this kind of complication will be omitted here.
2. *Two alternatives.* With three or more alternatives, would regret be a function of the best outcome of those alternatives not chosen or of the outcome of the second best alternative? This issue is avoided by restricting consideration to decisions with only two alternatives.
3. *Fully resolved outcomes.* In Allais' paradox the first pair of alternatives involves a consequence where zero would be received with a known loss of \$1 million. However, the problem statement never

makes clear whether, in the second pair of gambles, the gamble that is not chosen will be resolved. Nor is it known, if it is resolved, whether the two gambles would be resolved independently. This information is key to predicting how a decision maker, taking account of regret, would choose. In the explanation offered above it was assumed that lotteries not selected would not be resolved. Such arbitrary judgments will be avoided by confining attention to those situations in which, after a decision has been made, all uncertainties will be resolved in a predetermined fashion. The treatment of regret where unselected lotteries are not resolved is the subject of a future paper.

DEFINITION. *A choice between two lotteries, each defined by objective probabilities, will be termed a simple comparison if, after a decision, all uncertainties will be resolved in accordance with a predetermined joint probability distribution.*

In the next four sections, a utility function incorporating regret is formulated. Its structure is simplified by a number of assumptions that also serve to give insight about the ramifications to decision analysis that the incorporation of regret introduces. Section 2 routinely applies the resulting function to four paradoxes to provide a regret explanation of them. Section 3 discusses the problem of assessing a utility function that incorporates regret and Section 4 is a summary indicating extensions of the work.

1. INCORPORATING REGRET INTO A UTILITY FUNCTION

The concept of multiattribute utility theory is well established through considerable theoretical development and by numerous applications (Keeney and Raiffa [1976]). For those decision makers who do experience regret and are prepared to make tradeoffs to reduce it, including regret explicitly in an analysis makes sense. We will assume that appropriate attributes to capture the happiness a decision maker feels at any given endpoint are final assets, X , and the asset level perceived to have been given up as a result of the decision made, Y . The attribute Y will be called *foregone assets*. It is more convenient to use foregone assets rather than a measure of regret as a second attribute. It also has the advantage of being an objectively measurable quantity in the problem formulation. We will deduce the role of regret from the assumptions of this section.

It will be assumed that utility is an increasing function of x and a decreasing function of y . It will also be assumed that when outcomes are described with these two attributes, the decision maker obeys the axioms of von Neumann and Morgenstern so that decisions are made on the basis of expected values of a utility function $u(x, y)$. To be specific, if an

uncertain event E which will occur with probability p leads to final assets of x_1 under alternative 1 and x_3 under alternative 2, with asset values x_2 and x_4 respectively if E does not occur, then alternative 1 is preferred if and only if

$$pu(x_1, x_3) + (1 - p)u(x_2, x_4) > pu(x_3, x_1) + (1 - p)u(x_4, x_2). \quad (1)$$

This criterion is in accordance with the definition of foregone assets above.

Without further specification the function $u(x, y)$ is too general for satisfactory behavioral interpretations to be made. The remainder of this section is concerned with establishing behavioral assumptions to simplify the utility function.

It is reasonable to suppose that regret is related to the difference between assets received and those given up. However, an endpoint of $x = \$1,000$, $y = \$2,000$ may produce a much greater feeling of regret than an endpoint of $x = \$1,000,000$, $y = \$1,001,000$. The reason is that the incremental value of the thousand dollars is greater in the first case than in the second. We will suppose that there exists a value function $v(x)$ which captures the concept of incremental value in the following way: an increment from $\$a$ to $\$b$ has greater value than one from $\$c$ to $\$d$ if $v(b) - v(a) > v(d) - v(c)$. Although the notion of incremental value has great intuitive appeal, its relationship to other concepts of value is imprecise (see Fishburn [1970] or Bell and Raiffa [1979]). Since we will ultimately identify the quantity $v(x) - v(y)$ as a measure for regret, the function v could be *defined* by its relation to regret. That is, instead of taking incremental value as a primitive and deducing a measure of regret, one could regard regret as the primitive and deduce from it the concept of incremental value.

ASSUMPTION 1. *If all potential final asset positions of all decision alternatives are increased (decreased) by amounts of equal incremental value, then the preferred alternative is unchanged.*

Suppose alternative 1 with possible outcomes x_i , $i = 1, \dots, n$ is preferred to alternative 2 with possible outcomes y_j , $j = 1, \dots, m$. Then if each x_i is modified to x'_i and each y_j to y'_j in such a way that $v(x'_i) - v(x_i) = v(y'_j) - v(y_j)$ for all i and j then Assumption 1 asserts that alternative 1 should remain preferred.

The assumption is based upon the premise that $v(x)$ measures satisfaction with the final asset position and $v(x) - v(y)$ measures regret. If the value of each outcome is increased by a constant, the regret values are unaltered and the average value of final assets for each alternative is modified by the same amount.

For example, if a decision maker is just indifferent between \$100 for sure and a 50-50 lottery between \$0 and \$500, he has determined that the

potential benefit of improving \$100 to \$500 is just balanced by the possibility of decreasing \$100 to \$0. Now if $v(1,250) - v(1,000) = v(100) - v(0)$, and $v(2,000) - v(1,250) = v(500) - v(100)$, then by substitution of equivalent increments the improvement of \$1,250 to \$2,000 "should" balance the possible loss from \$1,250 to \$1,000. With this logic we deduce that \$1,250 would be indifferent to a 50-50 lottery between \$1,000 and \$2,000. When this assumption is applied to a utility function $u(x)$ over assets alone, it forces either $u(x)$ to equal $v(x)$ or an exponential transformation of $v(x)$. The decision maker has constant risk aversion (Pratt [1964]) relative to the value scale (see Dyer and Sarin [1982]).

LEMMA 1. *Assumption 1 implies that*

$$u(x, y) - u(y, x) = g(v(x) - v(y))\exp(-cv(x))$$

for some constant c and function g . In the special case when $v(x)$ is linear in x the conclusion is that

$$u(x, y) - u(y, x) = [u(0, y - x) - u(y - x, 0)]\exp(-cx)$$

for some constant c .

Proof. The proof will be given for the simpler case where $v(x) = x$ and the general case deduced from it. It can be seen from inequality (1) that, for simple comparisons, only the difference function $u(x, y) - u(y, x)$ can be identified since if a quantity $u(x_1, y_1)$ appears on one side of an equation, $u(y_1, x_1)$ must appear on the other. Define $w(x, y) = u(x, y) - u(y, x)$. Then inequality (1) may be written (let $y_1 = x_3$ and $y_2 = x_4$) as

$$pw(x_1, y_1) + (1 - p)w(x_2, y_2) > 0$$

and Assumption 1 implies that for all h , it follows from the above inequality that

$$pw(x_1 + h, y_1 + h) + (1 - p)w(x_2 + h, y_2 + h) > 0.$$

We will show that this implies that $w(x + h, y + h) = k(h)w(x, y)$ for some function $k(h) > 0$. Define $k(h)$ by $w(x_1 + h, y_1 + h)/w(x_1, y_1)$ for some particular values x_1 and y_1 . Now suppose that x_2, y_2 and q are values such that

$$qw(x_1, y_1) + (1 - q)w(x_2, y_2) = 0.$$

Unless $w(x_2 + h, y_2 + h) = k(h)w(x_2, y_2)$, we will have

$$qw(x_1 + h, y_2 + h) + (1 - q)w(x_2 + h, y_2 + h) \neq 0,$$

which would contradict Assumption 1. Therefore $w(x + h, y + h) = k(h)w(x, y)$ in general. From this expression we may deduce (replace h by x , x by 0 and y by $y - x$) that $w(x, y) = k(x)w(0, y - x)$ and (replace h by $x + h$, x by 0 and y by $y - x$) that $w(x + h, y + h) = k(x + h)w(0,$

$y - x$). Comparing these last three equations we may conclude that $k(x + h) = k(x)k(h)$ so that $k(x) = \exp(-cx)$ for some constant c . Recalling the definition of $u(x, y)$, we have $u(x, y) - u(y, x) = (u(0, y - x) - u(y - x, 0))\exp(-cx)$. In the case where $v(x)$ is not linear, we may assume that v is strictly monotonic increasing and define a new function u^* by $u^*(v(x), v(y)) = u(x, y)$. The proof is now as before but uses u^* in place of u . The result is that $u^*(v(x), v(y)) - u^*(v(y), v(x)) = (u^*(0, v(y) - v(x)) - u^*(v(y) - v(x), 0))\exp(-cv(x))$ or $u(x, y) - u(y, x) = g(v(x) - v(y))\exp(-cv(x))$ for the appropriate function g .

Two further assumptions will be made to simplify the functional forms being used. Assumption 2 will imply that the constant c in Lemma 1 is zero. Assumption 3 will permit a separation of $u(x, y)$ from $u(y, x)$ by use of preferences over choice sets.

ASSUMPTION 2. *If the incremental value from x_1 to x_2 equals that from x_2 to x_3 then the decision maker is indifferent between x_2 for sure and a 50-50 lottery between x_1 and x_3 .*

This assumption relates to the farmer's problem in Section 1, where he found a price of \$5 for sure to be just as risky as a 50-50 gamble between \$3 and \$7. In more detail the argument is as follows. We begin with the assertion that

$$[x_1 \text{ to } x_2] \text{ is indifferent to } [x_2 \text{ to } x_3]. \quad (2)$$

Because we have assumed the existence of a value function this also means that

$$[x_2 \text{ to } x_1] \text{ is indifferent to } [x_3 \text{ to } x_2]. \quad (3)$$

Now consider having x_2 as the status quo and giving it up for the lottery. This is a 50-50 gamble between the increments $[x_2 \text{ to } x_1]$ and $[x_2 \text{ to } x_3]$. Now consider the lottery as the status quo and giving it up for the sure x_2 . This is a 50-50 gamble between the increments $[x_1 \text{ to } x_2]$ and $[x_3 \text{ to } x_2]$. But these two gambles are equivalent by substitution, using (2) and (3). Because giving up the sure thing for the lottery is considered equal to giving up the lottery for the sure thing, then the two must be equally preferred. This argument, if applied to a utility function over assets alone, would force the utility function to be exactly $v(x)$ (Sarin [1982]).

LEMMA 2. *Assumptions 1 and 2 imply that*

$$u(x, y) - u(y, x) = g(v(x) - v(y))$$

for some function g . When $v(x)$ is linear the conclusion is that

$$u(x, y) - u(y, x) = u(0, y - x) - u(y - x, 0).$$

Proof. Simple substitution into the result of Lemma 1 provides a proof.

We have

$$\frac{1}{2}u(x_2, x_1) + \frac{1}{2}u(x_2, x_3) = \frac{1}{2}u(x_1, x_2) + \frac{1}{2}u(x_3, x_2)$$

or
$$\frac{1}{2}g(v(x_2) - v(x_1))e^{-cv(x_2)} = \frac{1}{2}g(v(x_3) - v(x_2))e^{-cv(x_3)}.$$

Since $v(x_2) - v(x_1) = v(x_3) - v(x_2)$ and $v(x_2) \neq v(x_3)$ we must have $c = 0$. With these assumptions a 50-50 lottery between x_1 and x_2 would be preferred to one between y_1 and y_2 if and only if

$$\frac{1}{2}g(v(x_1) - v(y_1)) + \frac{1}{2}g(v(x_2) - v(y_2)) > 0.$$

This manner of comparison was studied by Tversky [1969]. He showed that it may lead to intransitivities in lottery orderings unless g is linear.

Intuition about g is difficult because it is defined in an implicit manner from $u(x, y) - u(y, x)$. Separation of $u(x, y)$ from $u(y, x)$ is impossible so long as we restrict attention to simple comparisons. It is possible if we note that regret implies the decision maker may have a preference for one choice set over another. At an extreme, a decision maker who has severe problems with regret may sometimes prefer to have only a single alternative offered than a choice among two or more. In the dominance example of Selection 1, the portfolio manager preferred the choice set in scenario 1 to the choice set in scenario 2.

To make the point clearer, let S represent the decision maker's status quo. Suppose that the decision maker is required to undertake either lottery L_1 or lottery L_2 . If L_1 is selected then regret is measured (according to the criteria of this paper) against the foregone opportunity L_2 . However, had L_2 not been offered at all and the choice was to take L_1 in preference to staying at S then the regret would be measured against the foregone opportunity S . In this way the expected utility of receiving L_1 may differ depending upon the choice set from which it was selected. The principle of irrelevant alternatives is violated.

The following assumption is an extension of Assumption 1 to include preferences over choice sets.

ASSUMPTION 3. *Suppose the decision maker prefers to select an alternative L_1 over an alternative L_2 than to select an alternative L_3 over an alternative L_4 . Then this statement is invariant under a modification in which all outcomes of all alternatives are increased (decreased) by an equal incremental value.*

To see that Assumption 1 is a special case of Assumption 3, replace L_3 by L_2 and L_4 by L_1 in the statement of Assumption 3.

THEOREM 1. *Assumptions 2 and 3 imply that*

$$u(x, y) = v(x) + f(v(x) - v(y)) \quad (4)$$

for some function f .

Proof. We may begin with the result of Lemma 2, using first the special case $v(x) = x$ to simplify the notation. We then have $u(x, y) - u(y, x) = g(x - y)$. Suppose the 50-50 lotteries $L_1 = (a_1, a_2)$, $L_2 = (b_1, b_2)$, $L_3 = (b_1, c_1)$, $L_4 = (a_1, c_2)$ are such that choosing L_1 over L_2 is indifferent to choosing L_3 over L_4 . Then

$$\frac{1}{2}u(a_1, b_1) + \frac{1}{2}u(a_2, b_2) = \frac{1}{2}u(b_1, a_1) + \frac{1}{2}u(c_1, c_2)$$

or
$$g(a_1 - b_1) + u(a_2, b_2) = u(c_1, c_2).$$

Adding an increment h to all prizes gives

$$g(a_1 - b_1) + u(a_2 + h, b_2 + h) = u(c_1 + h, c_2 + h).$$

Comparing the last two equations gives

$$u(a_2 + h, b_2 + h) - u(a_2, b_2) = u(c_1 + h, c_2 + h) - u(c_1, c_2).$$

Since a_1 and b_1 are free variables, the quantities a_2, b_2, c_1 and c_2 may be considered arbitrary, so we may deduce that $u(x + h, y + h) - u(x, y) = j(h)$ for some function j . From this we may derive (replace h by x , x by 0 and y by $y - x$) that $u(x, y) = j(x) + u(0, y - x)$ and (replace h by $x + h$, x by 0 and y by $y - x$) that $u(x + h, y + h) = j(x + h) + u(0, y - x)$. From these last three equations we may conclude that $j(x + h) = j(x) + j(h)$ which implies j is linear. Hence $u(x, y) = x + f(x - y)$ where $f(x - y) = u(0, y - x)$. When $v(x)$ is nonlinear but monotonically increasing, the proof may be repeated using the function $u^*(v(x), v(y)) = u(x, y)$ as in the proof of Lemma 1.

With Assumption 2 omitted the result is that $u(x, y) = 1 - \exp(-cv(x) + f(v(x) - v(y))\exp(-cv(x)))$. The difference rests on whether the attributes "asset value," $v(x)$, and "regret," $v(x) - v(y)$, are additive independent or only mutually utility independent (Keeney and Raiffa). Since we seek the simplest representation that retains some behavioral implications from regret, the additive form has been assumed.

The next section applies the additive form (4) to four kinds of behavior regarded as paradoxical. To simplify the presentation it will be assumed that v is approximately linear in the range of dollar values used in the examples. It will be shown that regret provides a consistent explanation for the behavior so long as f is decreasingly concave. Section 3 discusses constructive tests to determine qualitative properties of f .

2. USING REGRET TO EXPLAIN BEHAVIOR

(i) The Coexistence of Insurance and Gambling

The fact that people gamble on negative expected value lotteries has been explained by reference to a second attribute, entertainment value. It is also easy to explain such behavior in terms of regret. In a long odds

situation where the potential payoff to gambling is great, the consequence with the largest regret is that in which you choose not to bet but hear that you would have won. For example, if you think of betting on a particular horse for the next race and then decide not to, it would be awful to see it win at long odds. A similar example can occur with the state numbers games. If you have bet on the number 2368 every day for the last 6 months, how can you stop now? The thought that 2368 may finally come up on the first day you quit may be sufficiently disturbing to keep you betting. In Massachusetts, recent advertising was aimed at getting you to think of a number. If you do, how can you then not bet on it? Indeed, for those who suffer from regret perhaps the only way to get out of such a dilemma would be to actively avoid hearing that day's winning number.

TABLE II
PAYOFFS ON HORSE RACE

	Bet	Don't Bet
Horse Wins	$\$(1-p)$	\$0
Horse Loses	$\$-p$	\$0

The purpose of the analysis that follows is to show that regret provides an explanation for gambling (risk prone behavior) even when the same decision maker may simultaneously purchase insurance against other, unrelated risks (risk averse behavior).

Suppose that a decision maker is considering betting on a particular horse which she thinks has a probability p of winning. For a bet of size $\$p$ the bookmaker is prepared to pay a total of \$1 if the horse wins and zero if it loses. See Table II. Both alternatives (bet-no bet) have, by construction, the same expected value. We will suppose that p is close to zero but certainly less than one half. A regret analysis shows that choosing to bet will either bring the satisfaction of winning a net amount of $\$(1 - p)$ or the dissatisfaction of losing $\$p$. Not betting, however, can produce the satisfaction that $\$p$ was not wasted or the dissatisfaction that $\$(1 - p)$ was missed. Since $p < \frac{1}{2}$, betting minimizes maximum regret, although this is not the criterion which results from the assumptions of Section 2. Using $u(x, y) = x + f(x - y)$ the decision maker will prefer the bet if

$$p[1 - p + f(1 - p)] + (1 - p)[-p + f(-p)] \\ > p[0 + f(p - 1)] + (1 - p)[0 + f(p)]$$

$$\text{or} \quad p[f(1 - p) - f(p - 1)] > (1 - p)[f(p) - f(-p)]. \quad (5)$$

Now suppose that, on another occasion, the decision maker is thinking whether to buy insurance for her car. The car, if damaged, will cost \$1 to repair. She feels the probability of the car being damaged is p , where p is close to zero but certainly less than one half. The insurance premium is $\$p$ (see Table III). Once again the expected values of the two options are equal. In this case the maximum regret occurs if no insurance is bought and the car is damaged. Taking expected utilities, the insurance is preferred if

$$p[-p + f(1 - p)] + (1 - p)[-p + f(-p)] \\ > p[-1 + f(p - 1)] + (1 - p)[0 + f(p)]$$

or
$$p[f(1 - p) - f(p - 1)] > (1 - p)[f(p) - f(-p)]. \quad (6)$$

Note that inequalities (5) and (6) are identical. If the decision maker prefers the bet she will necessarily also buy the car insurance. A key

TABLE III
PAYOFFS FOR INSURANCE EXAMPLE

	Insure	Don't Insure
Car Damaged	-\$p	-\$1
Car Undamaged	-\$p	\$0

difference between the situations is that in the bet the gamble represents a small probability of a (relatively) high payoff. Self insurance on the other hand offers a high probability of a low payoff (the cost of the insurance).

Inequalities (5) and (6) will hold in the direction shown if the function $(f(x) - f(-x))/x$ is increasing in x . This will be the case if, but not only if, $f''(x) > f''(-x)$ for all x and, therefore, if f is decreasingly concave (for example, negative exponential).

Uniformly risk averse behavior is impossible with (4) so long as v is linear. As v is made more and more concave, the decreasing marginal value of dollars is eventually sufficient to counteract the influence of regret and produce risk averse behavior even for low probability, high payoff bets.

(ii) The Reflection Effect

Kahneman and Tversky [1979] describe the following phenomenon. If [a gain of $\$x_2$ for sure] is just indifferent to [a p -chance at $\$x_1$ and a (1

$-p$)-chance at $\$x_3$] then [a loss of $\$x_2$ for sure] is indifferent to [a p -chance at a loss of $\$x_1$ and a $(1-p)$ -chance at losing $\$x_3$].

The paradox in this behavior stems from the view that utility for final assets should not depend upon the decision maker's current asset position. In particular if a decision maker is risk averse ($x_2 < px_1 + (1-p)x_3$) for gambles that improve his position he should also be risk averse for gambles that worsen his position. Yet if $x_2 < px_1 + (1-p)x_3$ we also have $-x_2 > p(-x_1) + (1-p)(-x_3)$, so that the above behavior contradicts this view. This phenomenon is usually explained by a utility function for assets that is S -shaped with the inflection point at current asset position.

Using $u(x, y) = x + f(x - y)$ the first indifference is represented by

$$\begin{aligned} p[x_2 + f(x_2 - x_1)] + (1-p)[x_2 + f(x_2 - x_3)] \\ = p[x_1 + f(x_1 - x_2)] + (1-p)[x_3 + f(x_3 - x_2)] \end{aligned}$$

and the second by

$$\begin{aligned} p[-x_2 + f(-x_2 + x_1)] + (1-p)[-x_2 + f(-x_2 + x_3)] \\ = p[-x_1 + f(-x_1 + x_2)] + (1-p)[x_3 + f(-x_3 + x_2)]. \end{aligned}$$

These two equalities are identical. To see this, collect the terms involving f on each right-hand side and all other terms on the left. Now multiply throughout one of the equations by -1 . This shows that the reflection effect is a necessary consequence of $u(x, y) = x + f(x - y)$. This explanation does not rely on a special role for current asset position, since regret is calculated only with respect to the two available alternatives.

By making the substitution $x_2 = p$, $x_1 = 1$ and $x_3 = 0$, it can be seen that the reflection effect is the same phenomenon as the coexistence of insurance and gambling discussed in Part (i) of this section.

(iii) Probabilistic Insurance

Kahneman and Tversky describe an interesting experiment in which a decision maker who is indifferent between self-insurance and buying insurance will nevertheless reject half price insurance, which is in force only a predetermined 50% of the time. Suppose that $-p$ for sure (insurance) is just indifferent to a q -chance at losing 1 and a $(1-q)$ -chance of losing nothing. With a traditional expected utility analysis we would have

$$u(-p) = qu(-1) + (1-q)u(0). \quad (7)$$

Half price insurance has an expected utility of

$$(1-q)u(-p/2) + q[\frac{1}{2}u(-p) + \frac{1}{2}u(-1)] \quad (8)$$

under the rules for payment laid out by Kahneman and Tversky (see

Table IV). For a risk averse decision maker (8) would be larger than each side of (7) because $u(-p/2) > \frac{1}{2}u(-p) + \frac{1}{2}u(0)$.

Using the two attribute utility function (4), indifference between insurance and no insurance gives

$$q[-p + f(1 - p)] + (1 - q)[-p + f(-p)] \\ = q[-1 + f(p - 1)] + (1 - q)[0 + f(p)]$$

or

$$-q + (1 - q)[f(p) - f(-p)] = -p + q[f(1 - p) - f(p - 1)]. \quad (9)$$

TABLE IV
PROBABILISTIC INSURANCE

	Accident	No Accident
Self Insurance	-1	0
Full Insurance	-p	-p
Probabilistic Insurance:		
Pays	-p	-p/2
Does Not Pay	-1	-p/2

Probabilistic insurance will be less preferred to full insurance if

$$q[\frac{1}{2}u(-1, -p) + \frac{1}{2}u(-p, -p)] + (1 - q)u(-p/2, -p) \\ < q[\frac{1}{2}u(-p, -1) + \frac{1}{2}u(-p, -p)] + (1 - q)u(-p, -p/2),$$

that is if

$$\frac{1}{2}q[-1 + f(p - 1)] + \frac{1}{2}q[-p + f(0)] + (1 - q)[-p/2 + f(p/2)] \\ < -p + \frac{1}{2}qf(1 - p) + \frac{1}{2}qf(0) + (1 - q)f(p/2).$$

If f is arbitrarily scaled so that $f(0) = 0$, this last inequality becomes

$$-q + 2(1 - q)[f(p/2) - f(-p/2)] \\ < -p + q[f(1 - p) - f(p - 1)]. \quad (10)$$

Subtracting (9) from (10) and dividing by $2(1 - q)$ gives

$$f(p/2) - f(-p/2) < \frac{1}{2}[f(p) - f(-p)]. \quad (11)$$

This condition will hold if $(f(x) - f(-x))/x$ is increasing in x , the same condition deduced from inequalities (5) and (6).

Now consider comparison of probabilistic insurance with self insurance. Probabilistic insurance is preferred if

$$\begin{aligned} & \frac{1}{2}q[-p + f(1 - p)] + \frac{1}{2}q[-1 + f(0)] + (1 - q)[-p/2 + f(p/2)] \\ & > \frac{1}{2}q[-1 + f(p - 1)] + \frac{1}{2}[-1 + f(0)] + (1 - q)[0 + f(p/2)]. \end{aligned}$$

Rearranging, subtracting (9) and dividing by $2(1 - q)$ gives as before, the inequality (11). Hence, if (11) is true, then probabilistic insurance is preferred to self insurance but less preferred to full insurance, even though self insurance and full insurance are indifferent. If (11) is not true, the preference relations are just reversed. Therefore, if a decision maker is offered a third choice of probabilistic insurance at least one of the two existing options will be preferable, which may explain why probabilistic insurance is dismissed.

It can be shown that intransitivities among simple comparisons cannot occur with $u(x, y)$ if and only if $u(x, y) = u^*(x) + s(x, y)$ where $u^*(x)$ is any function but $s(x, y)$ is symmetric in x and y , that is $s(x, y) = s(y, x)$. For simple comparisons the contributions from s always cancel leaving only those of $u^*(x)$. The function $u(x, y) = x + f(x - y)$ is not of this form unless f is linear (Tversky).

(iv) Preference Reversals

The probabilistic insurance example makes it clear that preference reversals (intransitivities) can occur using the function $u(x, y) = x + f(x - y)$. More straightforward instances have been investigated by Tversky, and Grether and Plott [1979]. The intransitivities occur in the following way. The certainty equivalent of a lottery L_1 is higher than the certainty equivalent of a lottery L_2 . However, in a direct comparison of the lotteries, L_2 is preferred. The most common circumstance is when L_1 is a high-probability low-payoff bet (P -bet) and L_2 is a low-probability high-payoff bet ($\$$ -bet). According to the results of Grether and Plott, reversals were much more common when the P -bet certainty equivalent was higher than when the $\$$ -bet certainty equivalent was higher.

Let L_1 be [p -chance at \$1 and a $(1 - p)$ -chance at \$ - 2] and L_2 be [p -chance at \$ - 1 and a $(1 - p)$ -chance at \$2]. If $p = \frac{2}{3}$, then from Part (i) we know that if f is decreasingly concave and v is linear we expect the certainty equivalent of L_1 to be less than zero and that of L_2 to be greater than zero. This result is because zero is the expected value of both lotteries. Because we expect to see a shift toward L_2 in a direct comparison the choice of $p = \frac{2}{3}$ will not produce the desired intransitivity. However, if we raise the value of p , L_1 becomes relatively more desirable. Suppose we choose p such that the certainty equivalent of L_1 is zero. From Part (ii) of this section we know the certainty equivalent of L_2 is thus also

zero. The probability p is defined by

$$p[0 + f(-1)] + (1 - p)[0 + f(2)] = p[1 + f(1)] + (1 - p)[-2 + f(-2)]. \quad (12)$$

L_2 will be preferred over L_1 if

$$p[1 + f(2)] + (1 - p)[-2 + f(-4)] < p[-1 + f(-2)] + (1 - p)[2 + f(4)]. \quad (13)$$

Let $h(x) = f(x) - f(-x)$ then Equation 12 becomes $ph(1) = (1 - p)h(2)$ and inequality (13) becomes $ph(2) < (1 - p)h(4)$. Note that $h''(x) = f''(x) - f''(-x)$. Hence h is convex if f is decreasingly concave. Subtracting (12) from (13) gives the requirement $p(h(2) - h(1)) < (1 - p)(h(4) - h(2))$. Clearly this will be satisfied if h is sufficiently convex, that is, if f is sufficiently decreasingly concave. In fact, so long as f is decreasingly concave (12) and (13) will hold for sufficiently small units of money.

An example that satisfies (12) and (13) may be converted to one of strict preference reversal by a small increase in the value of p .

The explanations provided by regret for the paradoxes examined in Parts (i), (ii) and (iv) all rested on f having a positive third derivative (a sufficient, not a necessary, condition). It is worth emphasizing that this is a rather mild requirement, given that it is already presumed that f is monotonically increasing and concave. In particular if $f'''(x)$ is to have constant sign, it must be strictly positive. The next section outlines some constructive tests to determine some relevant properties of preferences for regret.

3. ASSESSING REGRET TRADEOFFS

Assume the decision maker can assess the strength of preference function v (see Fishburn [1967] or Bell and Raiffa) and finds Assumptions 1 and 2 applicable. To assess the function g of Lemma 2, first note that $g(r) = -g(-r)$. If x' and y' are asset levels such that $v(x') - v(y') = r > 0$, then $g(r)$ may be assessed by finding the probability p that causes indifference between the lotteries L_1 and L_2 of Table V. The equivalence yields

$$pg(v(x') - v(y')) + (1 - p)g(v(-1) - v(0)) = 0$$

which implies

$$g(r) = (1 - p)/p$$

if $g(v(0) - v(-1))$ is arbitrarily scaled to unity.

If it is also the case that $u(x, y) = v(x) + f(v(x) - v(y))$ then $g(r) = r + f(r) - f(-r)$. Note that three of the paradoxes of the last section were

TABLE V
ASSESSING THE DIFFERENCE FUNCTION

	L_1	L_2
Probability p	x'	y'
Probability $(1-p)$	-1	0

contingent on the behavior of the function $(f(r) - f(-r))/r$ and consistency with experimentally observed behavior required that this quantity be increasing in r . Equivalently, $g(r)/r$ should be an increasing function of r . Convexity of g could be tested by reference to lotteries L_1 and L_2 in Table VI. (We will assume for simplicity that either $v(x)$ is linear or that X and Y have already been transformed to have constant marginal value.) If L_2 is preferred to L_1 for all positive values of x and h then

$$\frac{1}{2}g(x) < \frac{1}{4}g(x+h) + \frac{1}{4}g(x-h)$$

which implies that g is convex. By the reflection property of Section 2(ii) g would be concave for negative x .

Direct assessment of f using only simple comparisons requires statements from the decision maker of preference over choice sets. For example, f may be established as concave by a generalization of the dominance example of Section 1. Table VII shows three situations, a status quo and two alternatives. Assume $x_1 \neq x_2$ and that the probability of E is one half. If the decision maker, for all h , prefers selecting alternative 1 over the status quo, rather than selecting alternative 2 over the status quo, then f is concave.

Taking alternative 1 over the status quo gives an expected utility of

$$\frac{1}{2}u(x_1 + h, x_1) + \frac{1}{2}u(x_2 + h, x_2).$$

TABLE VI
TEST FOR CONVEXITY

	L_1	L_2
Probability $1/2$	x	0
Probability $1/4$	0	$x+h$
Probability $1/4$	0	$x-h$

Taking alternative 2 over the status quo gives an expected utility of

$$\frac{1}{2}u(x_2 + h, x_1) + \frac{1}{2}u(x_1 + h, x_2).$$

Substituting for $u(x, y)$, setting the first quantity as greater than the second and canceling gives

$$f(h) > \frac{1}{2}f(h + x_2 - x_1) + \frac{1}{2}f(h + x_1 - x_2)$$

which implies concavity.

The case $f(r) = 1 - \exp(-\gamma r)$, for $\gamma > 0$, satisfies the conditions (i) that $f(r)$ is concave, (ii) that $g(r)$ is convex, and (iii) that $g(r)/r$ is increasing for positive r . These conditions are sufficient to imply that the decision maker would act in a manner consistent with the paradoxical behavior examined in Section 2. Though the negative exponential function is by

TABLE VII
DOMINANCE ASSUMPTION

	Status Quo	Alternative 1	Alternative 2
E occurs	x_1	$x_1 + h$	$x_2 + h$
E does not occur	x_2	$x_2 + h$	$x_1 + h$

no means the only function with this property, it may be useful as an approximation requiring few assessments from the decision maker.

A test that would establish that f is negative exponential is as follows. Add to the alternatives of Table VII an alternative 3 with asset levels of $x_1 + h - \gamma(h)$ if E occurs and $x_2 + h - \gamma(h)$ if E does not occur. Define $\gamma(h)$ to be that quantity that makes the decision maker indifferent between selecting alternative 3 over the status quo and selecting alternative 2 over the status quo. Then the function f is exponential or linear if and only if $\gamma(h)$ is independent of h . This follows immediately from the equation

$$f(h - \gamma(h)) = \frac{1}{2}f(h + x_2 - x_1) + \frac{1}{2}f(h + x_1 - x_2)$$

which represents the condition that the two choice sets are indifferent.

This section has been concerned exclusively with assessment when Assumptions 1, 2, and 3 of Section 2 were judged to hold for the decision maker. So long as the decision maker can make preference judgments over choice sets, assessment of $u(x, y)$ could be carried out in the standard manner (Keeney and Raiffa) without any structural presumptions. Note that the attributes X and Y need not be single dimensional for the theory of this paper and the assessment procedures of this section to be applied.

See Dyer and Sarin [1979] for the assessment of multiattribute strength of preference functions.

4. DISCUSSION AND EXTENSIONS

Utility theory is often criticized because it fails to predict actual behavior for some quite straightforward comparisons between alternatives with uncertain consequences. It is easy to understand why decision makers may be skeptical of expected utility analysis as a prescriptive tool when it apparently fails even for some simple comparisons. The analyst's only defense of the theory has been to offer a careful explanation of why certain of the decision maker's responses are inconsistent with the axioms of von Neumann and Morgenstern. This explanation usually leaves the decision maker unimpressed. This paper shows that some of the paradoxical behavior of decision makers is consistent with a desire to avoid post decision regret.

The assessment procedures of Section 3 and the assumptions of Section 1 do not explicitly involve tradeoffs between assets and regret. This fact allows assessment to proceed without an explicit declaration from the decision maker that regret avoidance is a concern. The resulting multiattribute utility function may offer a better chance of reflecting, descriptively, the decision maker's holistic appraisals of independently posed alternatives. Deriving a utility function that is descriptively accurate is an important first step for establishing a prescriptive framework for aiding the decision maker.

The next step is to determine, with the decision maker, whether a regret term is an appropriate component of the analysis. Even if the decision maker agrees that regret avoidance is a goal to be traded off against final assets, he may wish to consider whether the tradeoffs he is implicitly using are appropriate. A constructive analysis might then be undertaken. Of course the decision maker may wish to eliminate the regret component entirely. Just as weather forecasters accept training to improve their probability calibration so perhaps decision makers may accept training to eliminate, as appropriate, the practice of comparing uncertain alternatives by a weighted function of value differences (Tversky).

The approach of this paper is subject to the criticism that adding an extra variable, regret, to the utility function inevitably allows for a better fit to observed behavior. Of course, this is true. However, the appeal of this additional factor lies in the widespread recognition that regret is a factor in decision making and because it is the same property of regret (decreasing concavity of the function f) that accounts for the explanations of the paradoxes discussed. It is because of this "degrees of freedom" objection that attention in this paper is restricted to simple comparisons.

Future research will examine the role of regret in situations in which the joint probability distribution of alternative consequences is unspecified and choices in which foregone lotteries are never resolved. There are other potential complications to the theory. For example, the degree of regret felt may depend on the amount of thought that went into the decision, or on how close the decision maker was to going with the alternative subsequently foregone. The level of regret felt may sometimes be related to the original status quo no matter what the outcome of foregone alternatives. More generally, this question is an example of the problem of defining regret when multiple alternatives are involved. No account is taken of mental biases identified in people's understanding of probabilities (Kahnemann and Tversky).

Clearly there is a need for experimental exploration of the role of regret in decision making. This paper suggests that regret may be an important factor in resolving the apparent failure of utility theory to reflect observed behavior and moreover, that multiattribute utility theory provides a satisfactory resolution.

REFERENCES

- ALLAIS, M. 1953. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* **21**, 503-546.
- BELL, D. E. 1981. Explaining Utility Theory Paradoxes by Decision Regret. In *Organizations: Multiple Agents with Multiple Criteria*, pp. 28-39, J. N. Morse (ed.). Springer-Verlag, Berlin.
- BELL, D. E., AND H. RAIFFA. 1979. Marginal Value and Intrinsic Risk Aversion, Working Paper, 79-65, Harvard Business School.
- DYER, J. S., AND R. K. SARIN. 1979. Measurable Multiattribute Value Functions. *Opns. Res.* **27**, 810-822.
- DYER, J. S., AND R. K. SARIN. 1982. Relative Risk Aversion. *Mgmt. Sci.* (in press).
- ELLSBERG, D., W. FELLNER AND H. RAIFFA. 1961. Symposium: Decisions under Uncertainty. *Quart. J. Econ.* **75**, 643-94.
- FISHBURN, P. C. 1967. Methods of Estimating Additive Utilities. *Mgmt. Sci.* **13**, 435-453.
- FISHBURN, P. C. 1970. *Utility Theory for Decision Making*. John Wiley & Sons, New York.
- GRETHER, D. M., AND C. R. PLOTT. 1979. Economic Theory of Choice and the Preference Reversal Phenomenon. *Am. Econ. Rev.* **69**, 623-638.
- HOGARTH, R. M. 1980. *Judgment and Choice*. John Wiley & Sons, Chichester.
- KAHNEMANN, D., AND A. TVERSKY. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47**, 263-290.
- KEENEY, R. L., AND H. RAIFFA. 1976. *Decisions with Multiple Objectives*. John Wiley & Sons, New York.
- LUCE, R. D., AND H. RAIFFA. 1956. *Games and Decisions*. John Wiley & Sons, New York.

- VON NEUMANN, J., AND O. MORGENSTERN. 1947. *Theory of Games and Economic Behavior*, Ed. 2. Princeton University Press, Princeton, N.J.
- PRATT, J. W. 1964. Risk Aversion in the Small and in the Large. *Econometrica* **32**, 122–136.
- SARIN, R. K. 1982. Strength of Preference and Risky Choice. *Opns. Res.* **30**, 982–997.
- SMITH, V. L. 1966. Measuring Nonmonetary Utilities in Uncertain Choices: The Ellsberg Urn. *Quart. J. Econ.* **80**, 324–329.
- TVERSKY, A. 1969. Intransitivity of Preferences. *Psychol. Rev.* **76**, 31–48.