

LOSS AVERSION IN RISKLESS CHOICE: A REFERENCE-DEPENDENT MODEL*

AMOS TVERSKY AND DANIEL KAHNEMAN

Much experimental evidence indicates that choice depends on the status quo or reference level: changes of reference point often lead to reversals of preference. We present a reference-dependent theory of consumer choice, which explains such effects by a deformation of indifference curves about the reference point. The central assumption of the theory is that losses and disadvantages have greater impact on preferences than gains and advantages. Implications of loss aversion for economic behavior are considered.

The standard models of decision making assume that preferences do not depend on current assets. This assumption greatly simplifies the analysis of individual choice and the prediction of trades: indifference curves are drawn without reference to current holdings, and the Coase theorem asserts that, except for transaction costs, initial entitlements do not affect final allocations. The facts of the matter are more complex. There is substantial evidence that initial entitlements do matter and that the rate of exchange between goods can be quite different depending on which is acquired and which is given up, even in the absence of transaction costs or income effects. In accord with a psychological analysis of value, reference levels play a large role in determining preferences. In the present paper we review the evidence for this proposition and offer a theory that generalizes the standard model by introducing a reference state.

The present analysis of riskless choice extends our treatment of choice under uncertainty [Kahneman and Tversky, 1979, 1984; Tversky and Kahneman, 1991], in which the outcomes of risky prospects are evaluated by a value function that has three essential characteristics. *Reference dependence*: the carriers of value are gains and losses defined relative to a reference point. *Loss aversion*: the function is steeper in the negative than in the positive domain; losses loom larger than corresponding gains. *Diminishing sensitivity*: the marginal value of both gains and losses decreases with their

*This paper has benefited from the comments of Kenneth Arrow, Peter Diamond, David Krantz, Matthew Rabin, and Richard Zeckhauser. We are especially grateful to Shmuel Sattath and Peter Wakker for their helpful suggestions. This work was supported by Grants No. 89-0064 and 88-0206 from the Air Force Office of Scientific Research, and by the Sloan Foundation.

size. These properties give rise to an asymmetric S-shaped value function, concave above the reference point and convex below it, as illustrated in Figure I.

In this article we apply reference dependence, loss aversion, and diminishing sensitivity to the analysis of riskless choice. To motivate this analysis, we begin with a review of selected experimental demonstrations.

I. EMPIRICAL EVIDENCE

The examples discussed in this section are analyzed by reference to Figure II. In every case we consider two options x and y that differ on two valued dimensions and show how the choice between them is affected by the reference point from which they are evaluated. The common reason for these reversals of preference is that the relative weight of the differences between x and y on dimensions 1 and 2 varies with the location of the reference value on these attributes. Loss aversion implies that the impact of a difference on a dimension is generally greater when that difference is evaluated as a loss than when the same difference is evaluated as a gain. Diminishing sensitivity implies that the impact of a difference is attenuated when both options are remote from the reference point for the relevant dimension. This simple scheme

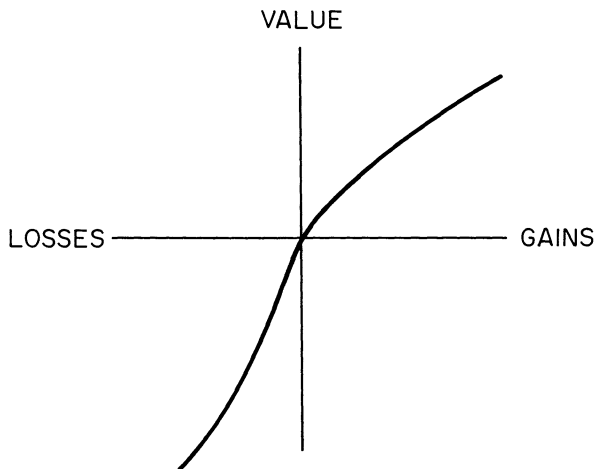


FIGURE I
An Illustration of a Value Function

serves to organize a large set of observations. Although isolated findings may be subject to alternative interpretations, the entire body of evidence provides strong support for the phenomenon of loss aversion.

a. Instant Endowment. An immediate consequence of loss aversion is that the loss of utility associated with giving up a valued good is greater than the utility gain associated with receiving it. Thaler [1980] labeled this discrepancy the endowment effect, because value appears to change when a good is incorporated into one's endowment. Kahneman, Knetsch, and Thaler [1990] tested the endowment effect in a series of experiments, conducted in a classroom setting. In one of these experiments a decorated mug (retail value of about \$5) was placed in front of one third of the seats after students had chosen their places. All participants received a questionnaire. The form given to the recipients of a mug (the "sellers") indicated that "You now own the object in your possession. You have the option of selling it if a price, which will be determined later, is acceptable to you. For each of the possible prices below indicate whether you wish to (x) Sell your object and receive this price; (y) Keep your object and take it home with you. . . ." The subjects indicated their decision for prices ranging from \$0.50 to \$9.50 in steps of 50 cents. Some of the students who had not received a mug (the "choosers") were given a similar questionnaire, informing them that they would have the option of receiving either a mug or a sum of money to be determined later. They indicated their preferences between a mug and sums of money ranging from \$0.50 to \$9.50.

The choosers and the sellers face precisely the same decision problem, but their reference states differ. As shown in Figure II, the choosers' reference state is t , and they face a positive choice between two options that dominate t ; receiving a mug or receiving a sum in cash. The sellers evaluate the same options from y ; they must choose between retaining the status quo (the mug) or giving up the mug in exchange for money. Thus, the mug is evaluated as a gain by the choosers, and as a loss by the sellers. Loss aversion entails that the rate of exchange of the mug against money will be different in the two cases. Indeed, the median value of the mug was \$7.12 for the sellers and \$3.12 for the choosers in one experiment, \$7.00 and \$3.50 in another. The difference between these values reflects an endowment effect which is produced, apparently instan-

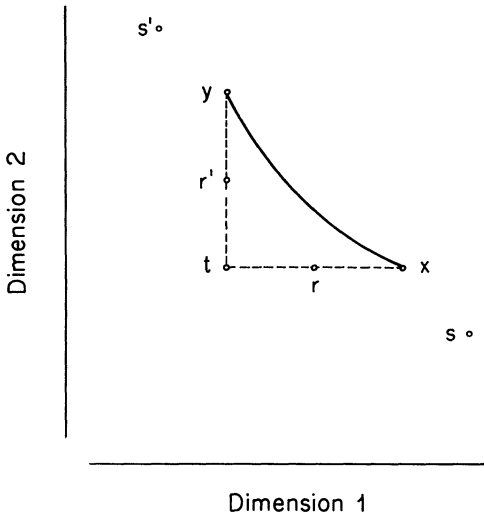


FIGURE II
Multiple Reference Points for the Choice Between x and y

taneously, by giving an individual property rights over a consumption good.

The interpretation of the endowment effect may be illuminated by the following thought experiment.

Imagine that as a chooser you prefer \$4 over a mug. You learn that most sellers prefer the mug to \$6, and you believe that if you had the mug you would do the same. In light of this knowledge, would you now prefer the mug over \$5?

If you do, it is presumably because you have changed your assessment of the pleasure associated with owning the mug. If you still prefer \$4 over the mug—which we regard as a more likely response—this indicates that you interpret the effect of endowment as an aversion to giving up your mug rather than as an unanticipated increase in the pleasure of owning it.

b. Status Quo Bias. The retention of the status quo is an option in many decision problems. As illustrated by the analysis of the sellers' problem in the example of the mugs, loss aversion induces a bias that favors the retention of the status quo over other options. In Figure II, a decision maker who is indifferent between x and y from t will prefer x over y from x , and y over x from y . Samuelson and Zeckhauser [1988] introduced the term "status quo bias" for this effect of reference position.

Knetsch and Sinden [1984] and Knetsch [1989] have offered compelling experimental demonstrations of the status quo bias. In the latter study two undergraduate classes were required to answer a brief questionnaire. Students in one of the classes were immediately given a decorated mug as compensation; students in another class received a large bar of Swiss chocolate. At the end of the session students in both classes were shown the alternative gift and were allowed the option of trading the gift they had received for the other, by raising a card with the word "Trade" written on it. Although the transaction cost associated with the change was surely slight, approximately 90 percent of the participants retained the gift they had received.

Samuelson and Zeckhauser [1988] documented the status quo bias in a wide range of decisions, including hypothetical choices about jobs, automobile color, financial investments, and policy issues. Alternative versions of each problem were presented to different subjects: each option was designated as the status quo in one of these versions; one (neutral) version did not single out any option. The number of options presented for each problem was systematically varied. The results were analyzed by regressing the proportions of subjects choosing an option designated as status quo $P(SQ)$, or an alternative to the status quo $P(ASQ)$, on the choice proportions for the same options in the neutral version $P(N)$. The results were well described by the equations,

$$P(SQ) = 0.17 + 0.83P(N) \quad \text{and} \quad P(ASQ) = 0.83P(N).$$

The difference (0.17) between $P(SQ)$ and $P(ASQ)$ is a measure of the status quo bias in this experiment.

Samuelson and Zeckhauser [1988] also obtained evidence of status quo bias in a field study of the choice of medical plans by Harvard employees. They found that a new medical plan is generally more likely to be chosen by new employees than by employees hired before that plan became available—in spite of the yearly opportunity to review the decision and the minimal cost of changing it. Furthermore, small changes from the status quo were favored over larger changes: enrollees who did transfer from the originally most popular Blue Cross/Blue Shield plan tended to favor a new variant of that plan over other new alternatives. Samuelson and Zeckhauser also observed that the allocations of pension reserves to TIAA and CREF tend to be very stable from year to year, in spite of large variations in rate of return. They invoked the status quo bias as an explanation of brand loyalty and

pioneer firm advantage, and noted that rational models that ignore status quo effects “will present excessively radical conclusions, exaggerating individuals’ responses to changing economic variables and predicting greater instability than is observed in the world” [p. 47].

Loss aversion implies the status quo bias. As noted by Samuelson and Zeckhauser [1988], however, there are several factors, such as costs of thinking, transaction costs, and psychological commitment to prior choices that can induce a status quo bias even in the absence of loss aversion.

c. Improvements versus Tradeoffs. Consider the evaluation of the options x and y in Figure II from the reference points r and r' . When evaluated from r , option x is simply a gain (improvement) on dimension 1, whereas y combines a gain in dimension 2 with a loss in dimension 1. These relations are reversed when the same options are evaluated from r' . Considerations of loss aversion suggest that x is more likely to be preferred from r than from r' .

Ninety undergraduates took part in a study designed to test this hypothesis. They received written instructions indicating that some participants, selected at random, would receive a gift package. For half the participants (the dinner group) the gift consisted of “one free dinner at MacArthur Park Restaurant and a monthly Stanford calendar.” For the other half (the photo group) the gift was “one 8×10 professional photo portrait and a monthly Stanford calendar.” All subjects were informed that some of the winners, again selected at random, would be given an opportunity to exchange the original gift for one of the following options:

x : two free dinners at MacArthur Park Restaurant

y : one 8×10 professional photo portrait plus two 5×7 and three wallet size prints.

The subjects were asked to indicate whether they preferred to (i) keep the original gift, (ii) exchange it for x , or (iii) exchange it for y . If people are averse to giving up the reference gift, as implied by loss aversion, then the preference for a dinner-for-two (x) over multiple photos (y) should be more common among the subjects whose reference gift was a dinner-for-one (r) than among subjects whose reference gift was the single photo (r'). The results confirmed this prediction. Only ten participants chose to keep the original gift. Among the remaining subjects, option x was selected by 81 percent of the dinner group and by 52 percent of the photo group ($p < 0.01$).

d. Advantages and Disadvantages. In our next demonstration a combination of a small gain and a small loss is compared with a combination of a larger gain and a larger loss. Loss aversion implies that the same difference between two options will be given greater weight if it is viewed as a difference between two disadvantages (relative to a reference state) than if it is viewed as a difference between two advantages. In the representation of Figure II, x is more likely to be preferred over y from s than from s' , because the difference between x and y in dimension 1 involves disadvantages relative to s and advantages relative to s' . A similar argument applies to dimension 2. In a test of this prediction subjects answered one of two versions of the following question:

Imagine that as part of your professional training you were assigned to a part-time job. The training is now ending, and you must look for employment. You consider two possibilities. They are like your training job in most respects except for the amount of social contact and the convenience of commuting to and from work. To compare the two jobs to each other and to the present one, you have made up the following table:

	Social contact	Daily travel time
Present job	isolated for long stretches	10 min.
Job x	limited contact with others	20 min.
Job y	moderately sociable	60 min.

The second version of this problem included the same options x and y , but a different reference job (s'), described by the following attributes: "much pleasant social interaction and 80 minutes of daily commuting time."

In the first version both options are superior to the current reference job on the dimension of social contact and both are inferior in commuting time. The different amounts of social contact in jobs x and y are evaluated as advantages (gains), whereas the commuting times are evaluated as disadvantages (losses). These relations are reversed in the second version. Loss aversion implies that a given difference between two options will generally have greater impact when it is evaluated as a difference between two losses (disadvantages) than when it is viewed as a difference between two gains (or advantages). This prediction was confirmed: Job x was chosen by 70 percent of the participants in version 1 and by only 33 percent of the participants in version 2 ($N = 106$, $p < 0.01$).

II. REFERENCE DEPENDENCE

In order to interpret the reversals of preference that are induced by shifts of reference, we introduce, as a primitive concept, a preference relation indexed to a given reference state. As in the standard theory, we begin with a choice set $X = \{x, y, z, \dots\}$ and assume, for simplicity, that it is isomorphic to the positive quadrant of the real plane, including its boundaries. Each option, $x = (x_1, x_2)$ in X , $x_1, x_2 \geq 0$, is interpreted as a bundle that offers x_1 units of good 1 and x_2 units of good 2, or as an activity characterized by its levels on two dimensions of value. The extension to more than two dimensions is straightforward.

A *reference structure* is a family of indexed preference relations, where $x \geq_r y$ is interpreted as x is weakly preferred to y from reference state r . The relations $>_r$ and $=_r$ correspond to strict preference and indifference, respectively. Throughout this article we assume that each $\geq_r, r \in X$, satisfies the standard assumptions of the classical theory. Specifically, we assume that \geq_r is complete, transitive, and continuous; that is, $\{x : x \geq_r y\}$ and $\{x : y \geq_r x\}$ are closed for any y . Furthermore, each preference order is strictly monotonic in the sense that $x \geq_r y$ and $x \neq y$ imply that $x >_r y$. Under these assumptions each \geq_r can be represented by a strictly increasing continuous utility function U_r (see, e.g., Varian [1984], Ch. 3).

Because the standard theory does not recognize the special role of the reference state, it implicitly assumes *reference independence*; that is, $x \geq_r y$ iff $x \geq_s y$ for all $x, y, r, s \in X$. This property, however, was consistently violated in the preceding experiments. To accommodate these observations, we describe individual choice not by a single preference order but by a family or a book of indexed preference orders $\{\geq_r : r \in X\}$. For convenience, we use the letters r, s to denote reference states and x, y to denote options, although they are all elements of X .

A treatment of reference-dependent choice raises two questions: what is the reference state, and how does it affect preferences? The present analysis focuses on the second question. We assume that the decision maker has a definite reference state in X , and we investigate its impact on the choice between options. The question of the origin and the determinants of the reference state lies beyond the scope of the present article. Although the reference state usually corresponds to the decision maker's current position, it can also be influenced by aspirations, expectations, norms, and

social comparisons [Easterlin, 1974; van Praag, 1971; van de Stadt, Kapteyn, and van de Geer, 1985].

In the present section we first define loss aversion and diminishing sensitivity in terms of the preference orders $\geq_r, r \in X$. Next we introduce the notion of a decomposable reference function and characterize the concept of constant loss aversion. Finally, we discuss some empirical estimates of the coefficient of loss aversion.

Loss Aversion

The basic intuition concerning loss aversion is that losses (outcomes below the reference state) loom larger than corresponding gains (outcomes above the reference state). Because a shift of reference can turn gains into losses and vice versa, it can give rise to reversals of preference, as implied by the following definition.

A reference structure satisfies *loss aversion* (LA) if the following condition holds for all x, y, r, s in X . Suppose that $x_1 \geq r_1 > s_1 = y_1, y_2 > x_2$ and $r_2 = s_2$; see Figure III. Then $x =_s y$ implies that $x >_r y$; the same holds if the subscripts 1 and 2 are interchanged throughout. (Note that the relations $>$ and $=$ refer to the numerical components of the options; whereas $>_r$ and $=_r$ refer to the preference between options in reference state r .) Loss aversion implies that the slope of the indifference curve through y is steeper

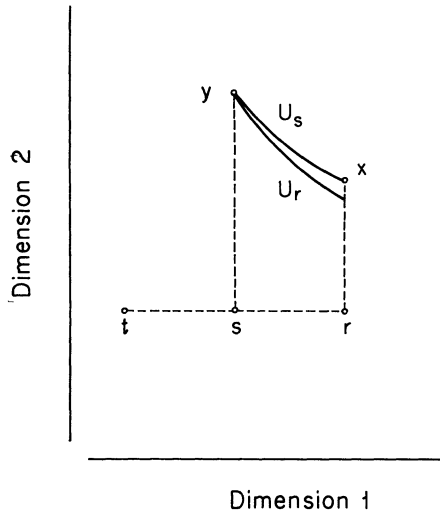


FIGURE III
A Graphic Illustration of Loss Aversion

when y is evaluated from r than when it is evaluated from s . In other words, $U_r^*(y) > U_s^*(y)$, where $U_r^*(y)$ is the marginal rate of substitution of U_r at y .

To motivate the definition of loss aversion, it is instructive to restate it in terms of advantages and disadvantages, relative to a reference point r . An ordered pair $[x_i, r_i]$, $i = 1, 2$, is called an advantage or a disadvantage, respectively, if $x_i > r_i$, or $x_i < r_i$. We use brackets to distinguish between the pair $[x_i, r_i]$ and the two-dimensional option (x_1, x_2) . Suppose that there exist real-valued functions v_1, v_2 such that $U_r(x)$ can be expressed as $U(v_1[x_1, r_1], v_2[x_2, r_2])$. To simplify matters, suppose that $x_1 = r_1$ and $x_2 > r_2$, as in Figure III. Hence, $x =_s y$ implies that the combination of the two advantages, $[x_1, s_1]$ and $[x_2, s_2]$, relative to the reference state s , has the same impact as the combination of the advantage $[y_2, s_2]$ and the null interval $[y_1, y_1]$. Similarly, $x >_r y$ implies that the combination of the advantage $[x_2, r_2]$ and the null interval $[x_1, x_1]$ has greater impact than the combination of the advantage $[y_2, r_2]$ and the disadvantage $[y_1, r_1]$. As the reference state shifts from s to r , therefore, the disadvantage $[y_1, r_1] = [s_1, r_1]$, enters into the evaluation of y , and the advantage $[x_1, s_1] = [r_1, s_1]$ is deleted from the evaluation of x . But since $[s_1, r_1]$ and $[r_1, s_1]$ differ by sign only, loss aversion implies that the introduction of a disadvantage has a bigger effect than the deletion of the corresponding advantage. A similar argument applies to the case where $x_1 > r_1 > s_1$.

The present notion of loss aversion accounts for the endowment effect and the status quo bias described in the preceding section. Consider the effect of different reference points on the preference between x and y , as illustrated in Figure II. Loss aversion entails that a decision maker who is indifferent between x and y from t will prefer x over y from x , and y over x from y . That is, $x =_t y$ implies that $x >_x y$ and $y >_y x$. This explains the different valuations of a good by sellers and choosers and other manifestations of the status quo bias.

Diminishing Sensitivity

Recall that, according to the value function of Figure I, marginal value decreases with the distance from the reference point. For example, the difference between a yearly salary of \$60,000 and a yearly salary of \$70,000 has a bigger impact when current salary is \$50,000 than when it is \$40,000. A reference structure satisfies *diminishing sensitivity* (DS) if the following condition holds for all x, y, s, t in X . Suppose that $x_1 > y_1, y_2 > x_2$,

$s_2 = t_2$, and either $y_1 \geq s_1 \geq t_1$ or $t_1 \geq s_1 \geq x_1$; see Figure III. Then $y =_s x$ implies that $y \geq_t x$; the same holds if the subscripts 1 and 2 are interchanged throughout. *Constant sensitivity* is satisfied if the same hypotheses imply that $y =_t x$. DS states that the sensitivity to a given difference on a dimension is smaller when the reference point is distant than when it is near. It follows from DS that the slope of the indifference curve through x is steeper when evaluated from s than from t , or $U_s^*(x) > U_t^*(x)$. It is important to distinguish between the present notion of diminishing sensitivity, which pertains to the effect of the reference state, and the standard assumption of diminishing marginal utility. Although the two hypotheses are conceptually similar, they are logically independent. In particular, diminishing sensitivity does not imply that the indifference curves are concave below the reference point.

Each reference state r partitions X into four quadrants defined by treating r as the origin. A pair of options, x and y , belong to the same quadrant with respect to r whenever $x_i \geq r_i$ iff $y_i \geq r_i$, $i = 1, 2$. A reference structure satisfies *sign dependence* if for all x, y, r, s in X $x \geq_r y$ iff $x \geq_s y$ whenever (i) x and y belong to the same quadrant with respect to r and with respect to s , and (ii) r and s belong to the same quadrant with respect to x and with respect to y . This condition implies that reference independence can be violated only when a change in reference turns a gain into a loss or vice versa. It is easy to verify that sign dependence is equivalent to constant sensitivity. Although sign dependence may not hold in general, it serves as a useful approximation whenever the curvature induced by the reference state is not very pronounced.

The assumption of diminishing (or constant) sensitivity allows us to extend the implications of loss aversion to reference states that do not coincide with x or y on either dimension. Consider the choice between x and y in Figure IV. Note that r is dominated by x but not by y , whereas s is dominated by y but not by x . Let t be the meet of r and s ; that is, $t_i = \min(r_i, s_i)$, $i = 1, 2$. It follows from loss aversion and diminishing sensitivity that if $x =_t y$, then $x >_r y$ and $y >_s x$. Thus, x is more likely to be chosen over y when evaluated from r than when evaluated from s . This proposition is illustrated by our earlier observation that a gift was more attractive when evaluated as a moderate improvement on one attribute than when evaluated as a combination of a large improvement and a loss (see example c above).

Consider two exchangeable individuals (i.e., hedonic twins), each of whom holds position t , with low status and low pay; see

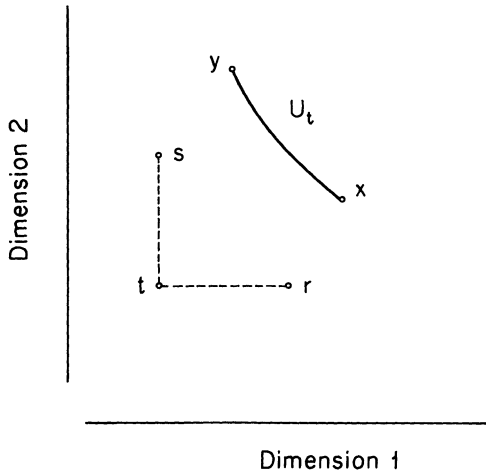


FIGURE IV
An Illustration of Reference-Dependent Preferences

Figure IV. Suppose that both are indifferent between position x (very high status, moderate pay) and position y (very high pay, moderate status). Imagine now that both individuals move to new positions, which become their respective reference points; one individual moves to r (high status, low pay), and the other moves to s (high pay, low status). LA and DS imply that the person who moved to r now prefers x , whereas the person who moved to s now prefers y , because they are reluctant to give up either salary or status.

Constant Loss Aversion

The present section introduces additional assumptions that constrain the relation among preference orders evaluated from different reference points. A reference structure (X, \geq_r) , $r \in X$, is *decomposable* if there exists a real-valued function U , increasing in each argument, such that for each $r \in X$, there exist increasing functions $R_i: X_i \rightarrow \text{Reals}$, $i = 1, 2$ satisfying

$$U_r(x_1, x_2) = U(R_1(x_1), R_2(x_2)).$$

The functions R_i are called the reference functions associated with reference state r . In this model the effect of the reference point is captured by separate monotonic transformations of the two axes. Decomposability has testable implications. For example, suppose

that U_r is additive; that is, $U_r(x_1, x_2) = R_1(x_1) + R_2(x_2)$. It follows then that, for any $s \in X$, U_s is also additive although the respective scales may not be linearly related.

In this section we focus on a special case of decomposability in which the reference functions assume an especially simple form. A reference structure (X, \geq_r) satisfies *constant loss aversion* if there exist functions $u_i: X_i \rightarrow \text{Reals}$, constants $\lambda_i > 0$, $i = 1, 2$, and a function U such that $U_r(x_1, x_2) = U(R_1(x_1), R_2(x_2))$, where

$$R_i(x_i) = \begin{cases} u_i(x_i) - u_i(r_i) & \text{if } x_i \geq r_i \\ (u_i(x_i) - u_i(r_i))/\lambda_i & \text{if } x_i < r_i. \end{cases}$$

Thus, the change in the preference order induced by a shift of reference is described in terms of two constants, λ_1 and λ_2 , which can be interpreted as the coefficients of loss aversion for dimensions 1 and 2, respectively. Figure V illustrates constant loss aversion, with $\lambda_1 = 2$ and $\lambda_2 = 3$. For simplicity, we selected a linear utility function, but this is not essential.

Although we do not have an axiomatic characterization of constant loss aversion in general, we characterize below the special case where U is additive, called additive constant loss aversion. This case is important because additivity serves as a good approxi-

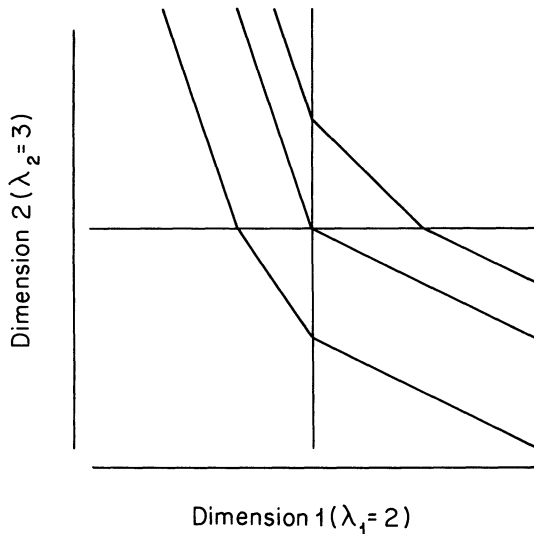


FIGURE V
A Set of Indifference Curves Illustrating Constant Loss Aversion

mation in many contexts. Indeed, some of the commonly used utility functions (e.g., Cobb-Douglas, or CES) are additive. Recall that a family of indifference curves is additive if the axes can be monotonically transformed so that the indifference curves become parallel straight lines. The following cancellation condition, also called the Thomsen condition, is both necessary and sufficient for additivity in the present context [Debreu, 1960; Krantz, Luce, Suppes, and Tversky, 1971].

For all $x_1, y_1, z_1 \in X_1, x_2, y_2, z_2 \in X_2$, and $r \in X$,

if $(x_1, z_2) \geq_r (z_1, y_2)$ and $(z_1, x_2) \geq_r (y_1, z_2)$, then $(x_1, x_2) \geq_r (y_1, y_2)$.

Assuming cancellation for each \geq_r , we obtain an additive representation for each reference state. In order to relate the separate additive representations to each other, we introduce the following axiom. Consider $w, w', x, x', y, y', z, z'$ in X that (i) belong to the same quadrant with respect to r as well as with respect to s , and (ii) satisfy $w_1 = w'_1, x_1 = x'_1, y_1 = y'_1, z_1 = z'_1$ and $x_2 = z_2, w_2 = y_2, x'_2 = z'_2, w'_2 = y'_2$; see Figure VI. A reference structure (X, \geq_r) , $r \in X$, satisfies *reference interlocking* if, assuming (i) and (ii) above, $w =_r x, y =_t z$ and $w' =_s x'$ imply that $y' =_s z'$. Essentially the same condition was invoked by Tversky, Sattah, and Slovic [1988] in the treatment of preference reversals, and by Wakker [1988] and

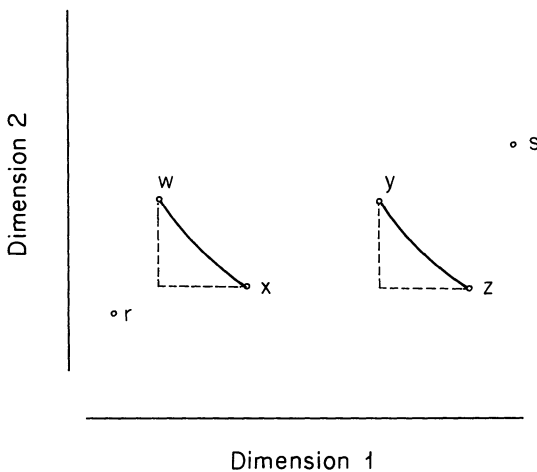


FIGURE VI
A Graphic Illustration of Reference Interlocking

Tversky and Kahneman [1991] in the analysis of decision under uncertainty.

To appreciate the content of reference interlocking, note that, in the presence of additivity, indifference can be interpreted as a matching of an interval on one dimension to an interval on the second dimension. For example, the observation $w =_r x$ indicates that the interval $[x_1, w_1]$ on the first dimension matches the interval $[w_2, x_2]$ on the second dimension. Similarly, $y =_r z$ indicates that $[z_1, y_1]$ matches $[y_2, z_2]$. But since $[w_2, x_2]$ and $[y_2, z_2]$ are identical by construction (see Figure VI), we conclude that $[x_1, w_1]$ matches $[z_1, y_1]$. In this manner we can match two intervals on the *same* dimension by matching each of them to an interval on the *other* dimension. Reference interlocking states that if two intradimensional intervals are matched as gains, they are also matched as losses. It is easy to verify that reference interlocking follows from additive constant loss aversion. Furthermore, the following theorem shows that in the presence of cancellation and sign-dependence, reference interlocking is not only necessary but it is also sufficient for additive constant loss aversion.

THEOREM. A reference structure (X, \geq_r) , $r \in X$, satisfies additive constant loss aversion iff it satisfies cancellation, sign-dependence, and reference interlocking.

The proof of the theorem is presented in the Appendix. An estimate of the coefficients of loss aversion can be derived from an experiment described earlier, in which two groups of subjects assigned a monetary value to the same consumption good: sellers who were given the good and the option of selling it, and choosers who were given the option of receiving the good or a sum of money [Kahneman, Knetsch, and Thaler, 1990]. The median value of the mug for sellers was \$7.12 and \$7.00 in two separate replications of the experiments; choosers valued the same object at \$3.12 and \$3.50. According to the present analysis, the sellers and the choosers differ only in that the former evaluate the mug as a loss, the latter as a gain. If the value of money is linear in that range, the coefficient of loss aversion for the mug in these experiments was slightly greater than two.

There is an intriguing convergence between this estimate of the coefficient of loss aversion and estimates derived from decisions under risk. Such estimates can be obtained by observing the ratio G/L that makes an even chance to gain G or lose L just acceptable. We have observed a ratio of just over 2:1 in several experiments. In

one gambling experiment with real payoffs, for example, a 50-50 bet to win \$25 or lose \$10 was barely acceptable, yielding a ratio of 2.5:1. Similar values were obtained from hypothetical choices regarding the acceptability of larger gambles, over a range of several hundred dollars [Tversky and Kahneman, 1990]. Although the convergence of estimates should be interpreted with caution, these findings suggest that a loss aversion coefficient of about two may explain both risky and riskless choices involving monetary outcomes and consumption goods.

Recall that the coefficient of loss aversion could vary across dimensions, as illustrated in Figure V. We surmise that the coefficient of loss aversion associated with different dimensions reflects the importance or prominence of these dimensions [Tversky, Sattath, and Slovic, 1988]. For example, loss aversion appears to be more pronounced for safety than for money [Viscusi, Magat, and Huber, 1987], and more pronounced for income than for leisure.

III. IMPLICATIONS OF LOSS AVERSION

Loss aversion is an important component of a phenomenon that has been much discussed in recent years: the large disparity often observed between the minimal amount that people are willing to accept (WTA) to give up a good they own and the maximal amount they would be willing to pay (WTP) to acquire it. Other potential sources of this discrepancy include income effect, strategic behavior, and the legitimacy of transactions. The buying-selling discrepancy was initially observed in hypothetical questions involving public goods (see Cummings, Brookshire, and Schulze [1986], for a review), but it has also been confirmed in real exchanges [Heberlein and Bishop, 1985; Kahneman, Knetsch, and Thaler, 1990; Loewenstein, 1988]. It also survived, albeit reduced, in experiments that attempted to eliminate it by the discipline of market experience [Brookshire and Coursey, 1987; Coursey, Hovis, and Schulze, 1987]; see also Knetsch and Sinden [1984, 1987]. Kahneman, Knetsch, and Thaler [1990] showed that the disparate valuations of consumption goods by owners and by potential buyers inhibits trade. They endowed half the participants with a consumption good (e.g., a mug) and set up a market for that good. Because the mugs were allocated at random, standard theory predicts that half the sellers should trade their mugs to buyers who value them more. The actual volume of trade was consistently

observed to be about half the predicted amount. Control experiments in which subjects traded tokens redeemable for cash produced nearly perfect efficiency and no disparity between the values assessed by buyers and sellers.

A trade involves two dimensions, and loss aversion may operate on one or both. Thus, the present analysis suggests two ways in which loss aversion could contribute to the disparity between WTA and WTP. The individual who states WTA for a good considers giving it up; the individual who states WTP for that good considers acquiring it. If there is loss aversion for the good, the owner will be reluctant to sell. If the buyer views the money spent on the purchase as a loss, there will be reluctance to buy. The relative magnitude of the two effects can be estimated by comparing sellers and buyers to choosers, who are given a choice between the good and cash, and are therefore not susceptible to loss aversion. Results of several comparisons indicated that the reluctance to sell is much greater than the reluctance to buy [Kahneman, Knetsch, and Thaler, 1990]. The buyers in these markets do not appear to value the money they give up in a transaction as a loss. These observations are consistent with the standard theory of consumer choice, in which the decision of whether or not to purchase a good is treated as a choice between it and other goods that could be purchased instead.

Loss aversion is certainly not involved in the exchange of a \$5 bill for \$5, because the transaction is evaluated by its net outcome. Similarly, reluctance to sell is surely absent in routine commercial transactions, in which goods held for sale have the status of tokens for money. However, the present analysis implies that asymmetric evaluations of gains and losses will affect the responses of both buyers and sellers to changes of price or profit, relative to the reference levels established in prior transactions [Kahneman, Knetsch, and Thaler, 1986; Winer, 1986]. The response to changes is expected to be more intense when the changes are unfavorable (losses) than when they are for the better. Putler [1988] developed an analysis of demand that incorporates an asymmetric effect of price increases and decreases. He tested the model by estimating separate demand elasticities for increases and for decreases in the retail price of shell eggs, relative to a reference price estimated from the series of earlier prices. The estimated elasticities were -1.10 for price increases and -0.45 for price decreases, indicating that price increases have a significantly greater impact on consumer decisions. (This analysis assumes that the availability of

substitutes eliminates loss aversion in the response to the reduced consumption of eggs.) A similar result was observed in scanner-panel data in the coffee market [Kalwani, Yim, Rinne, and Sugita, 1990]. The reluctance to accept losses may also affect sellers: a study of the stock market indicated that the volume of trade tends to be higher when prices are rising than when prices are falling [Shefrin and Statman, 1985].

Loss aversion can complicate negotiations. Experimental evidence indicates that negotiators are less likely to achieve agreement when the attributes over which they bargain are framed as losses than when they are framed as gains [Bazerman and Carroll, 1987]. This result is expected if people are more sensitive to marginal changes in the negative domain. Furthermore, there is a natural asymmetry between the evaluations of the concessions that one makes and the concessions offered by the other party; the latter are normally evaluated as gains, whereas the former are evaluated as losses. The discrepant evaluations of concessions significantly reduces the region of agreement in multi-issue bargaining.

A marked asymmetry in the responses to favorable or unfavorable changes of prices or profits was noted in a study of the rules that govern judgments of the fairness of actions that set prices or wages [Kahneman, Knetsch, and Thaler, 1986]. In particular, most people reject as highly unfair price increases that are not justified by increased costs and cuts in wages that are not justified by a threat of bankruptcy. On the other hand, the customary norms of economic fairness do not absolutely require the firm to share the benefits of reduced costs or increased profits with its customers or its employees. In contrast to economic analysis, which does not distinguish losses from forgone gains, the standards of fairness draw a sharp distinction between actions that impose losses on others and actions (or failures to act) that do not share benefits. A study of court decisions documented a similar distinction in the treatment of losses and forgone gains; in cases of negligence, for example, compensation is more likely to be awarded for out-of-pocket costs than for unrealized profits [Cohen and Knetsch, 1990].

Because actions that are perceived as unfair are often resisted and punished, considerations of fairness have been invoked as one of the explanations of wage stickiness and of other cases in which markets clear only sluggishly [Kahneman, Knetsch, and Thaler, 1986; Okun, 1981; Olmstead and Rhode, 1985]. For example, the

difference in the evaluation of losses and of forgone gains implies a corresponding difference in the reactions to a wage cut and to a failure to increase wages when such an increase would be feasible. The terms of previous contracts define the reference levels for collective as well as for individual bargaining; in the bargaining context the aversion to losses takes the form of an aversion to concessions. The rigidity induced by loss aversion may result in inefficient labor contracts that fail to respond adequately to changing economic circumstances and technological developments. As a consequence, new firms that bargain with their workers without the burden of previous agreements may gain a competitive advantage.

Is loss aversion irrational? This question raises a number of difficult normative issues. Questioning the values that decision makers assign to outcomes requires a criterion for the evaluation of preferences. The actual experience of consequences provides such a criterion: the value assigned to a consequence in a decision context can be justified as a prediction of the quality of the experience of that consequence [Kahneman and Snell, 1990]. Adopting this predictive stance, the value function of Figure I, which was initially drawn to account for the pattern of risky choices, can be interpreted as a prediction of the psychophysics of hedonic experience. The value function appropriately reflects three basic facts: organisms habituate to steady states, the marginal response to changes is diminishing, and pain is more urgent than pleasure. The asymmetry of pain and pleasure is the ultimate justification of loss aversion in choice. Because of this asymmetry a decision maker who seeks to maximize the experienced utility of outcomes is well advised to assign greater weight to negative than to positive consequences.

The demonstrations discussed in the first part of this paper compared choices between the same two objective states, evaluated from different reference points. The effects of reference levels on decisions can only be justified by corresponding effects of these reference levels on the experience of consequences. For example, a bias in favor of the status quo can be justified if the disadvantages of any change will be experienced more keenly than its advantages. However, some reference levels that are naturally adopted in the context of decision are irrelevant to the subsequent experience of outcomes, and the impact of such reference levels on decisions is normatively dubious. In evaluating a decision that has long-term consequences, for example, the initial response to these conse-

quences may be relatively unimportant, if adaptation eventually induces a shift of reference. Another case involves principal-agent relations: the principal may not wish the agent's decisions to reflect the agent's aversion to losses, because the agent's reference level has no bearing on the principal's experience of outcomes. We conclude that there is no general answer to the question about the normative status of loss aversion or of other reference effects, but there is a principled way of examining the normative status of these effects in particular cases.

APPENDIX

THEOREM. A reference structure (X, \geq_r) , $r \in X$, satisfies additive constant loss aversion iff it satisfies cancellation, sign dependence, and reference interlocking.

Proof. Necessity is straightforward. To establish sufficiency, note that, under the present assumptions, cancellation implies additivity [Debreu, 1960; Krantz et al., 1971]. Hence, for any $r \in X$ there exist continuous functions $R_i: X_i \rightarrow \text{Reals}$, unique up to a positive linear transformation, such that $R(x) = R_1(x_1) + R_2(x_2)$ represents \geq_r . That is, for any $x, y \in X$, $x \geq_r y$ iff $R(x) \geq R(y)$. We next establish the following two lemmas.

LEMMA 1. Let A be a set of options that belong to the same quadrant with respect to r and with respect to s . Then there exist $\lambda_i > 0$ such that for all x, y in A ,

$$R_i(y_i) - R_i(x_i) = (S_i(y_i) - S_i(x_i))/\lambda_i, \quad i = 1, 2.$$

Proof. We wish to show that for all $r, s, w, x, y, z \in X$,

$$\begin{aligned} R_i(z_i) - R_i(y_i) = R_i(x_i) - R_i(w_i) \text{ implies that} \\ S_i(z_i) - S_i(y_i) = S_i(x_i) - S_i(w_i), \quad i = 1, 2. \end{aligned}$$

This proposition follows from continuity, additivity, and reference interlocking whenever the i -intervals in question can be matched by intervals on the other dimension. If such matching is not possible, we use continuity to divide these i -intervals into sufficiently small subintervals that could be matched by intervals on the other dimension. Because equality of R_i differences implies equality of S_i differences, Lemma 1 follows from continuity and additivity.

LEMMA 2. Suppose that $r, s \in X$, with $s_1 < r_1$ and $s_2 = r_2$. Let S be a representation of \geq_s satisfying $S_1(s_1) = 0$. If sign-dependence and reference interlocking hold, then there exist $\lambda_1 > 0, \lambda_2 = 1$, such that $R^*(x) = R_1^*(x_1) + R_2^*(x_2)$ represents \geq_r , where

$$R_1^*(x_1) = \begin{cases} S_1(x_1) - S_1(r_1) & \text{if } x_1 \geq r_1 \\ (S_1(x_1) - S_1(r_1))/\lambda_1 & \text{if } s_1 \leq x_1 \leq r_1 \\ S_1(x_1) - S_1(r_1)/\lambda_1 & \text{if } x_1 \leq s_1 \end{cases}$$

and $R_2^*(x_2) = S_2(x_2) - S_2(r_2)/\lambda_2$. The same holds if the indices 1 and 2 are interchanged throughout.

Proof. By sign-dependence \geq_r and \geq_s coincide for all pairs of elements of $\{x \in X : x_1 \geq r_1, x_2 \geq r_2\}$ and of $\{x \in X : x_1 \geq r_1, x_2 \leq r_2\}$. To prove that \geq_r and \geq_s also coincide on their union, suppose that y belongs to the former set and z belongs to the latter. It suffices to show that $y =_r z$ implies that $y =_s z$. By monotonicity and continuity there exists w such that $y =_r w =_r z$ and $w_2 = r_2 = s_2$. Since w belongs to the intersection of the two sets, $y =_r w$ implies that $y =_s w$ and $z =_r w$ implies that $z =_s w$ hence $y =_s z$.

Therefore, we can select the scales so that $R_i = S_i, i = 1, 2$, on $\{x \in X : x_1 \geq r_1\}$. Next we show that $R^*(x) + S(r) = R(x)$. We consider each dimension separately. For $i = 2, R_2^*(x_2) + S_2(r_2) = S_2(x_2)$. We show that $S_2(x_2) = R_2(x_2)$. Select an $x_1 \geq r_1$. By construction, $S(x) = R(x)$ —hence $S_2(x_2) = R_2(x_2)$.

For $i = 1$, if $x_1 \geq r_1$, we get $R_1^*(x_1) + S_1(r_1) = S_1(x_1)$ and $R_1(x_1) = S_1(x_1)$, by construction. Hence

$$R_1^*(x_1) + S_1(r_1) = S_1(x_1) = R_1(x_1).$$

For $s_1 < x_1 < r_1$, we want to show that there exists λ_1 such that

$$\begin{aligned} R_1(x_1) &= S_1(r_1) + (S_1(x_1) - S_1(r_1))/\lambda_1, \text{ or} \\ R_1(x_1) - R_1(r_1) &= (S_1(x_1) - S_1(r_1))/\lambda_1, \end{aligned}$$

which follows from Lemma 1.

For $x_1 \leq s_1, \geq_r$ and \geq_s coincide, by sign-dependence—hence $R_1 = \alpha S_1 + \beta, \alpha > 0$. Because $R_2 = S_2, \alpha = 1$, and because $S_1(s_1) = 0, \beta = R_1(s_1)$ —hence $R_1(x_1) = S_1(x_1) + R_1(s_1)$. Consequently,

$$\begin{aligned} R_1(x_1) - R_1^*(x_1) &= S_1(x_1) + R_1(s_1) - (S_1(x_1) - S_1(r_1)/\lambda_1) \\ &= R_1(s_1) + S_1(r_1)/\lambda_1. \end{aligned}$$

It suffices to show that this expression equals $S_1(r_1)$. Consider

the case $s_1 < x_1 < r_1$, by continuity at s_1 ,

$$\begin{aligned} R_1(s_1) - R_1(r_1) &= (S_1(s_1) - S_1(r_1))/\lambda_1, \text{ hence} \\ R_1(s_1) + S_1(r_1)/\lambda_1 &= R_1(r_1) \\ &= S_1(r_1), \text{ by construction,} \end{aligned}$$

which completes the proof of Lemma 2.

Next we show that $\lambda_i, i = 1, 2$, is independent of r . Select $r, s, t \in X$ such that $r_2 = s_2 = t_2$ and $s_1 < r_1 < t_1$. By the previous lemma there exist R^* and T^* , defined in terms of S , with constants $\lambda^{(r)}$ and $\lambda^{(t)}$, respectively. Because \geq_r and \geq_t coincide on $\{x \in X : x_1 \leq r_1\}$, by sign-dependence, $\lambda^{(r)} = \lambda^{(t)}$. The same argument applies when indices 1 and 2 are interchanged, and when $r_1 < s_1$.

To establish sufficiency for the general case, consider $r, s \in X$, with $r_1 > s_1, r_2 \leq s_2$ and $t = (r_1, s_2)$. By applying the previous (one-dimensional) construction twice, once for (s, t) and then for (t, r) , we obtain the desired result.

STANFORD UNIVERSITY
UNIVERSITY OF CALIFORNIA, BERKELEY

REFERENCES

- Bazerman, Max, and John S. Carroll, "Negotiator Cognition," in B. Staw and L. L. Cummings, eds., *Research in Organizational Behavior*, Vol. VIII (Greenwich, CT: JAI Press, 1987), pp. 247-88.
- Brookshire, David S., and Don L. Coursey, "Measuring the Value of a Public Good: An Empirical of Elicitation Procedures," *American Economic Review*, LXXVII (1987), 554-66.
- Cohen, David, and Jack L. Knetsch, "Judicial Choice and Disparities Between Measures of Economic Values," Simon Fraser University Working Paper, 1990.
- Coursey, Don L., John L. Hovis, and William D. Schulze, "The Disparity between Willingness to Accept and Willingness to Pay Measures of Value," *Quarterly Journal of Economics*, CII (1987), 679-90.
- Cummings, Ronald G., David S. Brookshire, and William D. Schulze, eds., *Valuing Environmental Goods* (Totowa, NJ: Rowman and Allanheld, 1986).
- Debreu, Gerald, "Topological Methods in Cardinal Utility Theory," in Kenneth J. Arrow, Sam Karlin, and Patrick Suppes, eds., *Mathematical Methods in the Social Sciences* (Stanford, CA: Stanford University Press, 1960), pp. 16-26.
- Easterlin, Richard A., "Does Economic Growth Improve the Human Lot? Some Empirical Evidence," in P. A. David and M. W. Reder, eds., *Nations and Households in Economic Growth* (New York, NY: Academic Press, 1974), pp. 89-125.
- Heberlein, Thomas A., and Richard C. Bishop, "Assessing the Validity of Contingent Valuation: Three Field Experiments," Paper presented at the International Conference on Man's Role in Changing the Global Environment, Italy, 1985.
- Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, LXXVI (1986), 728-41.
- , —, and —, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, XCVIII (1990), 1325-48.

- Kahneman, Daniel, and Jackie Snell, "Predicting Utility," in Robin Hogarth, ed., *Insights in Decision Making* (Chicago, IL: University of Chicago Press, 1990).
- Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, XLVII (1979), 263-91.
- , and —, "Choices, Values and Frames," *American Psychologist*, XXXIX (1984), 341-50.
- Kalwani, Manohar U., Chi Kin Yim, Heikki J. Rinne, and Yoshi Sugita, "A Price Expectations Model of Customer Brand Choice," *Journal of Marketing Research*, XXVII (1990), 251-262.
- Knetsch, Jack L., "The Endowment Effect and Evidence of Nonreversible Indifference Curves," *American Economic Review*, LXXIX (1989), 1277-84.
- , and J. A. Sinden, "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, XCIX (1984), 507-21.
- , and —, "The Persistence of Evaluation Disparities," *Quarterly Journal of Economics*, CII (1987), 691-95.
- Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky, *Foundations of Measurement*, Vol. I (New York, NY: Academic Press, 1971).
- Loewenstein, George, "Frames of Mind in Intertemporal Choice," *Management Science*, XXXIV (1988), 200-14.
- Okun, Arthur, *Prices and Quantities: A Macroeconomic Analysis* (Washington, DC: The Brookings Institution, 1981).
- Olmstead, Alan L., and Paul Rhode, "Rationing Without Government: The West Coast Gas Famine of 1920," *American Economic Review*, LXXV (1985), 1044-55.
- Putler, Daniel S., "Reference Price Effects and Consumer Behavior," unpublished, Economic Research Service, U. S. Department of Agriculture, Washington, DC, 1988.
- Samuelson, William, and Richard Zeckhauser, "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, I (1988), 7-59.
- Shefrin, Hershey, and Meir Statman, "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence," *Journal of Finance*, XL (1985), 777-90.
- Thaler, Richard, "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, I (1980), 39-60.
- Tversky, Amos, Shmuel Sattath, and Paul Slovic, "Contingent Weighting in Judgment and Choice," *Psychological Review*, XCV (1988), 371-84.
- Tversky, Amos, and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," unpublished, Stanford University, 1991.
- van de Stadt, Huib, Arie Kapteyn, and Sara van der Geer, "The Relativity of Utility: Evidence from Panel Data," *Review of Economics and Statistics*, LXVII (1985), 179-87.
- van Praag, Bernard M. S., "The Individual Welfare Function of Income in Belgium: An Empirical Investigation," *European Economic Review*, XX (1971), 337-69.
- Varian, Hal R., *Microeconomic Analysis* (New York, NY: Norton, 1984).
- Viscusi, W. Kip, Wesley A. Magat, and Joel Huber, "An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks," *Rand Journal of Economics*, XVIII (1987), 465-79.
- Wakker, Peter P., *Additive Representations of Preferences: A New Foundation of Decision Analysis* (Dordrecht, The Netherlands: Kluwer Academic Publishers, 1989).
- Winer, Russell S., "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, XIII (1986), 250-56.