Invited Presentation

Rational Roots of "Irrational" Behavior: New Theories of Economic Decision-Making

James J. Opaluch and Kathleen Segerson

The neoclassical paradigm has proven to be a rich approach for evaluating a variety of issues for individual and social decision-making. However, an increasing body of literature suggests that actual behavior systematically violates the neoclassical utility model. This paper reviews a number of alternative models for decision-making. Results from the literature show several examples of apparently "irrational" behavior that can be explained in terms of these alternative motivations. The paper also extends the received literature by examining in some detail the implications of one such model which is based on the psychological feeling of ambivalence. The paper demonstrates that ambivalence has the potential for explaining the appearance of intransitive choices, the use of rules of thumb in decision-making and the large discrepancies between stated willingness-to-pay and willingness-to-accept, all of which have been observed in various settings. There are potentially great rewards from innovative research that expands the neoclassical paradigm to incorporate additional motivational factors in decision-making.

Introduction

The neoclassical models of utility maximization and expected utility maximization have proven to be rich models that are widely applicable to a variety of individual and social choice problems. The models are based on simple and intuitively appealing axioms of behavior and have proven surprisingly powerful for a variety of conceptual and empirical applications. They imply restrictions on behavior that can be tested with empirical data and provide insight into both individual choice and social policy formation.

Despite the great success of the use of the standard neoclassical models, economists have no cause to become complacent. Recent work by both economists and others have found that behavior is systematically inconsistent with the implications of these models. Given the normative appeal of the neoclassical behavioral axioms, such violations might be viewed as evidence of "irrational" behavior. However, the apparent persistence of these violations has led some researchers to seek a "rational" explanation for them, i.e. to develop intuitively appealing alternatives to, or generalizations of, the standard models that actually predict and possibly explain these violations.

It is our belief that these alternative behavioral models could play an important role in applied economic analysis, not as substitutes for the neoclassical models but rather as complements to them. We have learned much from applications of the neoclassical models, and they will undoubtedly continue to provide insight into unsolved problems. However, there is clearly room for expanding the paradigm to include additional factors, in much the same way that expected utility theory expanded the standard utility model to include risk. The resulting models have the potential to provide insights that are not forthcoming from the utility or expected utility models, such as explanations of behavior that might otherwise appear "irrational." Such insights should enrich our understanding of private decision-making and ultimately public policy.

The authors are, respectively, Associate Professor of Resource Economics and Assistant Professor of Economics. The authors wish to acknowledge the support of the Rhode Island Agricultural Experiment Station (AES Contribution No. 2501), Rhode Island Water Resources Center and the Northeast Agricultural and Resource Economics Association.

This paper examines some of these issues in two parts. First, we briefly outline three general approaches to modelling decision-making, and in that context we discuss a few of the factors that might influence individual choices, but are not included in the standard neoclassical models. We do not attempt to provide a comprehensive review of all alternative models since the literature on this topic is now quite large and useful surveys exist elsewhere (e.g. Machina; Schoemaker). Our hope is simply to introduce some of these kinds of modelling efforts to those who are not familiar with them and to encourage others in the profession to consider expanding their research to include some of these concerns, where appropriate.

In the second part of the paper we speculate on the implications of one alternative model that has not received much attention thus far. In particular, we explore the implications of the psychological feeling of ambivalence through an informal examination of its implications for observable behavior and for compensating welfare measures. Again, our aim is to be suggestive, rather than comprehensive or rigorous.

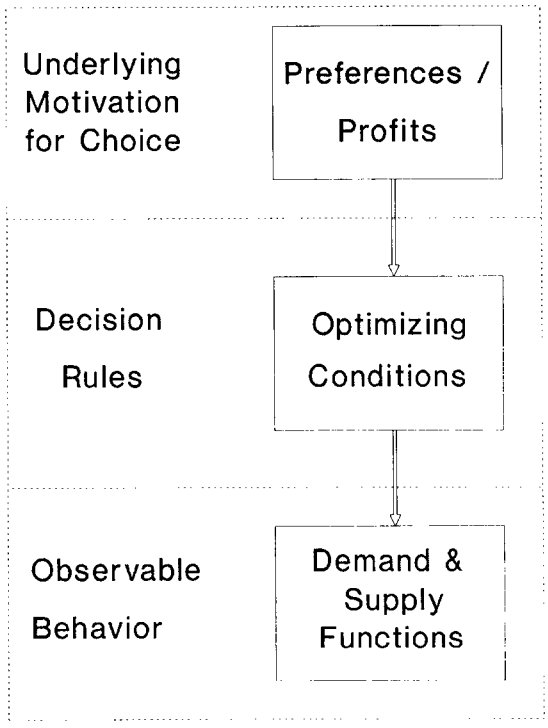

Figure 1. Depiction of General Framework for Neoclassical Behavioral Modelling

Alternative Approaches for Modelling Behavior

Complete behavioral models provide linkages among three components of choice:

(1) underlying motivations for behavior
(2) decision rules for making choices and
(3) potentially observable behavior.

For example, Figure 1 depicts these three components for the behavioral paradigm of neoclassical theory. In terms of motivations, neoclassical consumer theory assumes that the consumers are motivated by desires described by preference orderings, where these preferences are defined over bundles of goods and services consumed. Neoclassical producer theory assumes that firms are motivated by maximization of profits, as determined by production technology and input/output prices. The decision rules of neoclassical optimization are described by a set of first order necessary conditions which equate marginal benefits and marginal costs, subject to the second order conditions. These decision rules describe optimizing behavior and provide insights into trade-offs faced in decision-making. Finally, observable behavior is a series of demand and supply functions on which neoclassical theory imposes certain restrictions.

In examining alternatives to the neoclassical behavioral models, one can start by specifying an alternative underlying motivation; by specifying alternative decision rules for action; or by directly describing alternative rules for observed behavior. Each of these approaches has advantages and disadvantages. Specifying underlying motivations is the most structural approach, in the sense that these motivations provide the origin of decision rules and, ultimately, of observable behavior. This general method of behavioral modelling has been termed the consistent modelling method (Weaver and Stefanou). It specifies a complete and consistent framework for modelling, starting with the underlying motivation and tracing the implications through to actual behavior.

Neoclassical theory is an example of a consistent modelling approach. Indeed modern duality theory is really a comprehensive statement of the relationships between underlying

motivations of utility/profit maximization (or cost minimization) and observable demand/supply functions. For example, Hotelling's Lemma provides a tie between the unobserved indirect profit function and potentially estimable input demand functions. Similarly, the approach suggested by Hausman provides a tie between potentially observable demand functions and the expenditure function. Thus, a link is provided between potentially observable market behavior and the underlying motivations of the firm and consumer. This duality framework also places constraints on behavior, such as upward-sloping output supply functions, downward-sloping input demand functions, and symmetric, negative semi-definite Hicksian price effects. These duality relationships have been widely exploited in both theoretical and empirical work.

One difficulty of the consistent modelling approach is that the resultant structure is incapable of describing the full richness of actual behavior. Instead, the consistent modelling approach identifies a critical motivation for decision-making and traces out the logic of the resultant model. In many cases where little information is available, a less structured and restrictive approach based on general observations of behavior may be useful, as emphasized by Weaver and Stefanou. Furthermore, there is generally no conclusive way of verifying the appropriateness of an hypothesized motivation from observable behavior because there is not necessarily a unique correspondence between underlying motivations and decision rules or observable behavior. However, partial validation can be carried out when certain types of behavior are identified that are consistent with some underlying motivation. Thus, although a particular hypothesized motivation cannot generally be positively validated, it may be invalidated as being incompativle with a particular observed behavior.

The second approach to behavioral modelling is to start by specifying some particular set of decision rules, such as optimizing rules, satisficing rules, rules of thumb or lexicographic rules. One can then determine the set of observable behavior that is implied by such a decision rule. Decision rules are neither motivations for behavior, nor actual behavior itself, but rather they are a means of linking the two—that is they are a means of translating motivations into behavior. Decision rules could be precise, formal means of implementing these motivations or could merely be a convenient 'philosophy' for making choices. In addition, the decision rule may be an actual rule used by an individual making choices or can be an 'as if' rule, where potentially observable behavior that is consistent with the underlying motivations and the decision environment can be viewed 'as if' the decision-maker were following some particular rule. Further, decision rules may be equivalent to or formal restatements of the underlying motivation or can be a convenient means of implementing a particular motivation in making otherwise complex decisions in an uncertain environment.

Basing behavioral modelling on decision rules can be very useful for predicting observed behavior, but can lead to difficulties for welfare measurement and social policy analysis. That is, because this approach does not necessarily specify an underlying value or objective function, it may not allow one to define a welfare measure to use in judging alternative outcomes. Thus, this approach is necessarily restrictive in term of its ultimate use in policy analysis.

The third approach is to start by specifying rules for observed behavior. This approach can be very useful for predicting behavior as it is capable of integrating broad observations on actual choices. Thus, such a specification does not necessarily restrict behavioral modelling to the extent that the consistent modelling approach may. This approach can also be expected to result in estimates that are likely to be consistent with and perhaps be a good predictor of observed behavior. Furthermore, one need not understand why individuals behave as they do in order to use this approach to describe or predict behavior. Nevertheless, taking stock of observed behavior can be a useful method for the initial specification of hypothesized motivations. By carefully examining observed behavior we can consider the dimensions of the behavioral modelling problem that underlying motivations need to explain. Indeed, behavior in various experimental environments has led to many of the observed violations of utility theory that have been the driving force behind the development of many of the alternative behavioral models discussed below. As with the second approach, however, this approach can also result in potential difficulties for social welfare measurement, as it does not generally imply a

unique underlying value function.

In the literature on alternative paradigms, examples of each of these three modelling approaches can be found. Since we are interested in the "roots of 'irrational' behavior," out primary focus in the remainder of the paper will be on the first approach, namely starting from underlying motivations for decision-making. It should be clear from the following discussion, however, that the distinction between motivations, decision rules and behavioral implications may sometimes be blurred. Nonetheless, we feel that those factors that we have termed "motivations" below satisfy the spirit, if not the letter, of the categorization of modelling approaches that we have just outlined.

Motivating Factors

All economic theories of behavioral motivation either implicitly or explicitly embody assumptions regarding the factors that influence choice. These factors represent the psychological considerations involved in making a decision and provide the motivation for choosing one feasible alternative over another.

As noted above, in neoclassical consumer theory under certainty, the factors influencing choice are the quantities of goods and services consumed and the individual's preferences over alternative combinations of those quantities. Utility depends only on these factors. Given two potential bundles of goods and services to be consumed, bundles A and B, the individual is assumed to prefer A over B, to prefer B over A, or to be indifferent between the two. These preferences are assumed to be complete, reflexive, continuous and transitive.

The von Neumann-Morgenstern expected utility theory is similar but incorporates risk. Under this model, the factors influencing choice are the quantities of goods/services under different states of the world, the (objective) probabilities of those states occurring, and the individual's preferences over the lotteries represented by the alternative combinations of quantities and probabilities.

Subjective Probabilities

One of the earliest modifications of the expected utility model was the inclusion of subjective, rather than objective, probabilities for different states of the world (e.g. Ramsey; Savage). It was recognized that in many cases decision-makers are unlikely to know the true (objective) probabilities and will thus base their decisions on their subjective beliefs about those probabilities.

Although decisions being motivated by subjective rather than objective probabilities is a minor conceptual change, it can result in fairly complex and important changes in behavior when it is recognized that information and processing power of individuals are imperfect. People are often argued to have particular difficulty estimating the probability of rare events. One reason is that within the individual's experience there is not sufficient information to provide exact estimates of these events, precisely because they are so rare. Alternatively, the individual may not be capable of perceiving or differentiating between probabilities of, say, 10^{-6} and 10^{-9}. Given an inability to perceive of such small numbers, a decision-maker may assume that the probability is zero, with the idea that they are unrepresentative or unlikely occurrences. Alternatively, when the individual is unable to comprehend such small numbers, the individual may perceive of these probabilities as an arbitrary small number, like 10^{-3}.

Because of the lack of experience and the difficulties inherent in preceiving low probabilities, the information available to the person is often argued to bias the perception of the likelihood of the event. Current observations may be given too much weight (Tversky and Kahneman, 1973). For example, an individual who works in a dangerous job but never personally observes an accident occurring may underestimate the probability of an accident. Alternatively, media coverage of events such as plane crashes may cause an over-estimation of the probability of a crash since the public is aware of all crashes but not of all safe trips. Likewise, subjective estimates of the probability of groundwater contamination from agricultural pesticides may be strongly influenced by the number of reported cases of contamination of which the individual is aware. As a final example, the popularity of purchasing lottery tickets may be explainable in terms of people's inability to comprehend the true probability of winning. Instead the perception of the subjective probability may be based on news accounts of 'real' people who win multi-million dollar prizes. These ex-

amples suggest that subjective probabilities are closely tied to an individual's experience with or knowledge of such events occurring.

On the other hand, the literature on "anchoring" (Edwards, 1968) suggests that in some cases too little weight may be given to current observations, i.e. that individuals are reluctant to modify their subjective probabilities in light of new information (observations). For example, individuals living in flood plains have been known to rebuild houses destroyed by a flood in exactly the same location. Having experienced a flood, they do not revise upward their subjective probability of a flood, as suggested by Bayes rule. Rather they view the flood as an "aberration" that won't be repeated.

While the underlying process for estimating subjective probabilities is complex and not well understood, it is clear that calculations based on subjective probabilities can easily be inaccurate or imprecise due to biases in the available information and the inherent difficulty of carrying out complex calculations based on limited information. Thus, a divergence between objective and subjective probabilities may explain some of the observed violations of expected utility theory. However, replacing objective probabilities with subjective ones does not change the fundamental nature of the individual's choice problem. Choice would still be determined solely by probabilities, final outcomes and preferences over lotteries.

Regret/Disappointment

Some researchers have suggested that other factors, such as psychological feelings of regret or disappointment, may also be important motivations for choice. Bell (1982) and Loomes and Sugden have proposed models that incorporate anticipated feelings of regret or rejoicing that might be experienced when an individual learns that his decision turned out to be "wrong" or "right" in the sense of yielding an unfavorable or a favorable outcome *ex post*.

The fundamental assumption in regret theory is that final utility depends not only on what you get, but also on what you could have gotten had you chosen an alternative action. Thus, a catastrophe is compounded if the individual could have chosen an action that would have avoided the catastrophe. Similarly, the individual may rejoice if the action results in an outcome that would have been worse had some other action been chosen. If such feelings are anticipated, individuals would be expected to compare the outcomes resulting from different decisions under a given state of the world and use this information in their decision-making process.

Suppose, for example, that there is uncertainty about the health effects associated with pesticide contamination of groundwater and that a policy maker is contemplating a ban on use of the pesticide. If the policy maker chooses to ban it and then it turns out to be highly toxic, *ex post* he would have made the right decision. If the potential danger is revealed *ex post*, the policy maker might experience an additional payoff (rejoicing) due to his having made the correct choice, in addition to the direct utility resulting from avoiding the contamination. Alternatively, if he had chosen not to ban the pesticide, then *ex post* his decision would have been wrong. He could have avoided the negative effects by choosing an alternative action, the ban, but did not do so. The responsibility he bears (i.e. the regret he feels for not having chosen the ban) might make the outcome seem even worse than it would have been if there were no choice by the decision-maker that would have avoided the problem. In this case, the contamination incident occurs and, in addition, the policy maker must accept the blame and guilt for having 'caused' it by not choosing an action that could have avoided the incident. If the policy maker anticipates such feelings, they can be expected to influence his decision regarding the ban. In fact, it seems likely that such feelings played a role in the recent ban on importation of Chilean grapes that followed the discovery of cyanide in two grapes. Similarly, the recent Pan Am tragedy may have been compounded by the fact that there had been prior warning which was not fully heeded.

The concept of regret/rejoicing reflects an assumption that individuals judge an outcome relative to the outcome that would have occurred in that state of the world if a different decision had been made. Alternatively, the concept of disappointment (Bell, 1985) suggests that actual outcomes are judged relative to the expected outcome. For example, a groundwater contamination level of x ppm might be viewed differently depending upon whether the level was expected to be higher or

lower than x. Thus, expectations may influence the valuation of any given outcome.

The psychological feelings of regret and disappointment both imply that utility depends not only on the realized outcome, such as the level of damages from contamination, but also on how that outcome compares to some alternative outcome, either one that would have resulted from an alternative choice or one that was expected. Thus, the realized outcome is judged relatively rather than absolutely. It is this feature that distinguishes these models from the standard neoclassical models and results in behavior that violates neoclassical assumptions.

For example, under neoclassical utility theory, choices are assumed to be transitive. However, intransitive choices have been observed in a variety of experiment contexts (e.g. Edwards, 1954; May; Tversky). Transivity need not hold when the motivational assumptions embodied in the neoclassical model are relaxed. For example, Loomes and Sugden have shown that anticipated feelings of regret or rejoicing can lead to intransitive pairwise preferences. Regret implies that preferences are not independent of the choice set, i.e. the set of options from which the decision-maker can choose, since each option is evaluated relative to the outcomes that would have been obtained had any of the other options been chosen. Thus, the evaluation of any given option depends on the other options that were available when the choice is made.

Under regret theory apparently intransitive choices can be rational. Consider, for example, the hypothetical situation where the individual chooses option A from the choice set {A, B, C} and chooses option B from the choice set {A, B}. Under neoclassical theory, these choices are irrational since they violate the weak axiom of revealed preference. The first decision is interpreted as the individual revealing that A is preferred to B, while the second decision is interpreted as revealing that B is preferred to A. However, under regret theory no such inconsistency is indicated because *ex post* feelings of regret imply that *ex post* utility is determined by the entire choice set, rather than just the bundle chosen. Thus, the choice must be defined over the entire choice set, and cannot be defined over a part of the set. In the first choice described above the individual merely reveals that, given a choice between A, B and C, they would choose bundle A, not necessarily that A is preferred to B. Suppose, for example, if the individual chooses option A there is a chance that great regret will be felt for not having chosen option C under some possible state. On the other hand, if the individual chooses option B there may be much less regret for not having chosen option C. Thus, when C is an alternative, the individual may choose B rather than A. In this case the individual's anticipated feelings of regret for not having chosen option C may be the factor which results in the choice of B rather than A. On the other hand, when C is not in the choice set no regret would be felt for not having chosen C and the individual may choose A rather than B.

Anticipated regret or rejoicing might also be an explanation for the observed disparity between willingness-to-pay (WTP) and willingness-to-accept (WTA) measures of value. Knetsch and Sinden have suggested that decision-makers might view the reference alternative (for example, "no policy change" in the case of WTP and the "policy change" in the case of WTA) as choiceless in the sense that it is the position the individual starts from rather than a position to which he explicitly chose to move. If this is true, the individual might not experience any regret or rejoicing from staying at that position. This implies, however, that anticipated regret/rejoice feelings would differ in WTP and WTA scenarios, since the choiceless reference alternative would differ. Opaluch and Segerson have shown that with this asymmetry in regret/rejoicing, WTP and WTA measures of value will generally differ in ways other than simply through the income effect.

Reference Points

Further support for the hypothesis that individuals base decisions on relative rather than absolute outcomes is provided by the evidence suggesting the use of reference points (Kahneman and Tversky; Markowitz; Simon; Tversky and Kahneman, 1981). In making decisions, individuals often identify a reference point and judge possible outcomes relative to that reference point. For example, if the reference or starting point in a discussion of groundwater quality is that all wells in a given area are uncontaminated, then contamination of 50% of the wells might be viewed as a "loss." Alternatively, if the reference point is that all of the wells are contaminated, then

finding that only 50% of the wells are contaminated would be viewed as a "gain." In either case, the outcome is that 50% of the wells are contaminated, but it might be judged differently with the different reference points.

Reference points would be a factor motivating behavior if the 'utility' of the individual is based on changes from a reference point. In this case, preference functions underlying decision-making need to be defined in terms of gains or losses (i.e. movements from the reference point) rather than realized outcomes (e.g. the number of contaminated wells). Alternatively, reference points could simply be a convenient way of implementing 'true' absolute preferences in a complex world characterized by uncertainty, as opposed to risk.

Reference points are also a possible explanation for "framing" effects. Framing effects arise when changes in the wording or description of the options alter the individual's choice (e.g. Slovic; Tversky and Kahneman, 1981), possibly due changes in the individual's reference point. For example, in a decision regarding banning the use of a toxic substance, the decision-maker may choose differently if the policy choice is described as losing y lives in order to save x dollars as opposed to describing it as spending x dollars to save y lives. By stating the policy in the first way the decision-maker might be induced to choose implicitly a reference point of "zero exposure." He starts from a position of zero lives lost and then considers whether losing y lives is worth a savings of x dollars. Alternatively, the reference point is implicitly "high exposure" if the policy is stated as spending x dollars to save y lives. Here the decision-maker starts from a position of zero lives saved and then considers whether saving y lives is worth x dollars. This change in the reference point or framing of the option can lead to changes in the choice regarding adoption of the policy.

Reference points are also another possible explanation for the difference between WTP and WTA measures noted above (Gregory). WTP scenarios may elicit feelings about payment for a *gain*, while WTA measures might capture required compensation for a *loss*. If identical final wealth positions are viewed differently depending upon whether they were attained through a gain or a loss, as suggested by the use of reference points, then the observed disparity between WTP and WTA measures of value could be considered "rational."

Complexity

Another factor that could influence the decision-making process is the complexity of the decision environment. Simon's seminal work on the bounds to rationality suggests that limits on cognitive ability may prevent a decision-maker from being able to (or wanting to take the time to) perform the calculations necessary to maximize the expected value of some preference function. If a choice situation involves many possible states of the world and there is uncertainty about both the probabilities of the different states occurring and the possible outcomes under these states, then the complexity of the choice problem may prevent the decision-maker from considering all "relevant" information. Instead, he may structure his preferences lexicographically, focusing on a single characteristic of the problem as the main determinant of choice, with some priority sequence on the various characteristics.

Alternatively, if the probabilities of outcomes or the outcomes themselves are highly complex, uncertain and/or subject to large errors, then rather than trying to use very imperfect information to balance possible costs and benefits, a decision-maker may instead choose to rely on the use of a rule of thumb. For example, rather than trying to weigh costs and benefits from the reduction of each toxic substance individually, policy makers may instead set a single "acceptable" risk level to be applied to several or all substances even though the costs and benefits associated with that risk would vary across substances. The implications of using rules of thumb are discussed in more detail below.

Ambivalence

A key feature of the neoclassical paradigm is the assumption that the decision-maker has the ability to rank options unambiguously. Given a choice between bundles A and B, the neoclassical model assumes that either A is preferred to B, B is preferred to A or the individual is indifferent between A and B. This, however, ignores a potentially important psychological feeling, that of ambivalence (Opaluch). Ambivalence arises when the individual faces strongly opposing feelings when making a decision. In cases where the decision is a difficult one to resolve, the decision-maker faces a dilemma where the outcome cannot be

correctly described as indifference. Indifference suggests that the individual does not care which alternative occurs. This implies a complacency in decision-making that is not an appropriate description of the psychological demeanor for choice where the individual has strongly opposing feelings for two options. Rather, under extreme ambivalence the choice embodies a deep psychological conflict which must be resolved. Thus, rather than making decisions according to a balancing of costs and benefits, conflict resolution determines the choices to be made and this may result in behavior that violates neoclassical assumptions.

Akerlof and Dickens did the first work by economists that incorporated ambivalence in decision-making. They examined the case of individuals choosing dangerous jobs and argued that psychological conflict arises because individuals must somehow reconcile the fact that they choose to risk their lives at work every day. Akerlof and Dickens argue that individuals will resolve this psychological conflict by modifying their beliefs to conform with their choice of occupation. Thus, in order to resolve the internal conflict, people in dangerous occupations may modify their subjective probabilities of an accident so that they choose to believe that their job is not dangerous. As Akerlof and Dickens demonstrate, this implies people in dangerous jobs may not take sufficient safety precautions since they subsequently believe that their job is not as dangerous as it really is. Further, if workers resolve their internal conflict by choosing to believe that their jobs are really safe, they will not require compensating wage payments so that hazards will not be internalized to the firm nor will they tend to act through unions to demand safer working conditions. Thus, the optimal level of precautions for accident avoidance in the workplace will require regulations on worker safety. In contrast, the neoclassical decision paradigm suggests that workers will have the incentive to see that costs associated with hazards at the workplace are internalized to the firm.

The concept of ambivalence may be more general and pervasive than the type of problem discussed by Akerlof and Dickens. Ambivalence can arise whenever an individual's choice involves tradeoffs among characteristics that cannot be easily compared. Suppose, for example, that an individual has both tastes and values, where tastes rank outcomes in terms of personal payoffs, while values rank outcomes according to how the individual believes that the world 'ought' to be. In this example, tastes relate to the usual concept of individual preferences over goods and services, while values relate to feelings about such things as ethics and moral principles. The usual concept of indifference may not be appropriate for decisions concerning tradeoffs between personal payoffs and values, such as the amount of money one must be paid to violate one's moral principles.

Essentially, under ambivalence the decision-maker has non-scalar preferences which are based on two different objective functions, one reflecting his social values and the other reflecting his personal tastes, and the individual is unable to make precise trade-offs between these two. Hence, when values and tastes clash, decisions may not be based on the types of calculations implicit in the neoclassical utility model and its generalizations. Indeed, even the concept of making a decision through trading moral values for personal rewards may be viewed as reprehensible and unacceptable social behavior. Thus, the balancing involved in neoclassical models cannot be viewed as the basis of choice. It is possible that in these circumstances that individual switches into a different mode of decision-making.

For example, to resolve the conflict one might simply determine that values will not be compromised for any amount of money. This can lead to the use of decision rules based on rules of thumb or a lexicographic ordering. To avoid making comparisons between factors that the decision-maker views as non-comparable, such as dollars and deaths, he may limit his choice set to those options that do not require such a comparison to be made. For example, setting the allowable level of a cancer-causing substance at zero or at some arbitrary level effectively rules out options that require an explicit trade between lives and dollars. By banning the use of a substance, the government eliminates the need to decide whether the benefits obtained from that use outweigh the costs or whether the incremental benefits of an increase in use outweigh the incremental cost. Likewise, banning the importation of Chilean grapes when two were found to contain cyanide eliminated the need to decide whether the reduction in the risk of poisonings from other tainted grapes was worth the cost of taking all grapes off the

market.

When choices are based on rules of thumb, due either to ambivalence or complexity, the observed behavior may not directly reveal the decision-maker's underlying preferences. For example, the behavior resulting from a "safety first" rule would suggest that the decision-maker is concerned only with safety and not with monetary costs. If a lexicographical ordering is used in the final choice, the role of other factors may not be observed. Since the decision-making criterion is designed specifically to avoid the need to balance costs and benefits, we cannot use the resulting observed behavior to infer the decision-makers preferences in making such a balance.

Thus, we must be careful in making inferences about underlying preferences from these limited observations on behavior. In some cases where behavior is governed by rules of thumb, the only way in which preferences are potentially revealed in observable behavior is when the decision-maker changes the rule of thumb to account for special circumstances of a particular decision. Suppose, for example, that EPA sets standards for groundwater contamination according to the rule of thumb of one cancer case in a million. Thus, the agency would set allowable concentrations for all pollutants, independent of the associated costs, such that the expected incidence of cancer would be one for each million individuals who drink the water. In such a case the agency does not reveal a unique monetary value for avoiding cancer cases. However, there are invariably decisions that arise where it is 'unacceptably' expensive to achieve this standard. In such a case the agency would need to relax the constraints to allow a higher concentration of the pollutant and an associated higher cancer risk. For this example the only observable trade-off between money and lives is in determining what is 'unacceptably' expensive, hence requiring a change in standards. Thus, any attempt at inferring relative preferences of money versus health risk for such a case would have to be targeted at evaluating the cases where exceptions are made in using the rule of thumb.

Decision-Making Under Ambivalence

We now attempt to extend concepts of neoclassical theory to decision-making under ambivalence and speculate on possible differences in behavior that may result from ambivalence. Consider the case discussed above, where an individual faces two motivation functions, one for social values and one for personal payoffs or tastes. Implicitly we assume here that the individual is able to specify unique preference orderings on social values and on personal payoffs, but that the individual cannot make precise tradeoffs between values and tastes. Thus, there is no precise scalar measure of ultimate preferences that includes both values and tastes.

While there is no single, precise measure of preferences, the individual may have a limited ability to make such trades. Thus, for example, the individual might be willing to accept a small compromise in values for a large personal payoff. The individual would clearly be willing to tell a white lie for a million dollars. On the other hand, the individual would likely not accept a large compromise in values for a small improvement in personal payoffs. The individual would clearly not kill someone for $5, even if the probability of being caught is zero. Thus, the individual may be able to make some comparisons between values and tastes. However, there may be a large middle ground where the individual may find that decision-making is very difficult, possibly uncertain and unreliable, and where indifference is not an appropriate description of the individual's psychological state. Rather, when choosing among these strongly ambivalent alternatives, the individual faces a state of psychological conflict which must be resolved in some manner.

The two alternative means for resolving the psychological conflict resulting from ambivalence that were discussed above are lexicographical rules, such as not allowing values to be compromised for tastes, or modifying beliefs such that decisions can be rationalized as not being in conflict with values. Using these two decision rules the decision-maker can make trades between personal payoffs and values, but these trades may be inconsistent with neoclassical theory, as indicated above.

Consider the case depicted in Figure 2. Here the individual starts at point I, with personal payoffs at level T and values at level V. All alternatives to the Northeast of I, denoted by set A, are clear improvements, as the situation leads to an improvement in terms of the individual's personal payoffs without compromising values. In contrast, for points to the

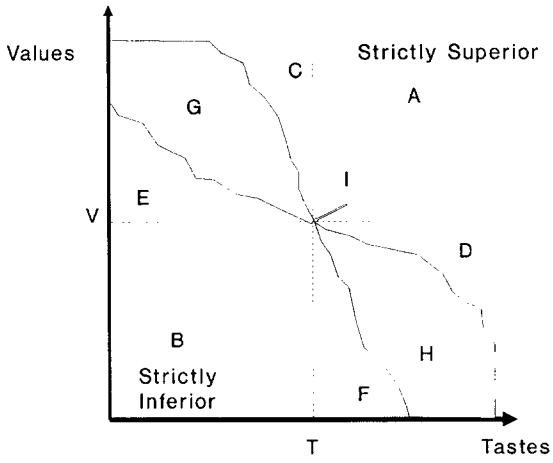

Figure 2. Depiction of Preference Ordering with Fuzzy Ambivalence Region

Southwest, denoted by B, values are compromised without a countervailing improvement in terms of personal payoffs. For points to the southeast and the northwest the individual faces alternatives that force a comparison between personal payoffs and values. For some set such as C, little is lost in terms of personal payoffs, while substantial improvements are forthcoming in terms of the value ordering. Similarly for set D, little is lost in terms of the value ordering but substantial improvements are made in terms of personal payoffs. As depicted in the figure, the individual prefers points within regions C and D to the initial point, I, and hence views these alternatives as desirable trades. In contrast, E and F represent the sets where much is lost for little gain, so that the individual prefers point I to points within these regions, and they are viewed as clearly unacceptable trades. Finally, the shaded areas, H and G, represent the 'fuzzy' ambivalence sets where changes in personal payoffs and values are in conflict and where the individual finds decision-making difficult. Similar to cases of complexity, the individual may have limits to 'rational,' consistent decision-making within the ambivalence range simply because the individual is incapable of making precise tradeoffs within this region. Nevertheless, individuals are sometimes forced to choose among such points and it is in making these kinds of choices that we have speculated individuals may make decisions according to lexicographic rules or rationalization.

Discrete choice modelling may be useful for predicting behavior within the ambivalence region or for characterizing the region itself. Ignore for the moment all sources of randomness in decision-making except for that arising from ambivalence. Because decision-making is uncertain within the ambivalence region, the region that is ambivalent to point I in Figure 2 may be defined as the region where the probability of accepting a trade is greater than zero but less than one. Although behavior may not be completely predictable within the ambivalence region, one might expect the probability of accepting an ambivalent trade to be a function of where the alternative lies within the ambivalent region. That is, as one moves within the ambivalence zone from the boundary of region F, which contains points which are clearly inferior to I, to the boundary of region D, which contains points which are clearly superior to I, the probability of accepting the trade will likely increase from zero to one.

At one extreme, as ambivalence becomes more difficult to resolve, the ambivalence set expands. In the limit as tradeoffs between preferences and values become impossible the ambivalence set fills the entire Northwest and Southeast quadrants as shown in Figure 3. As ambivalence becomes easier to resolve the ambivalence set becomes sharper and narrower until it collapses to a curve similar to the indifference curve of ordinary utility theory, as depicted by U in Figure 4. Implicitly economists assume that individuals are always able to resolve ambivalence in this precise manner. Here we argue that ambivalence

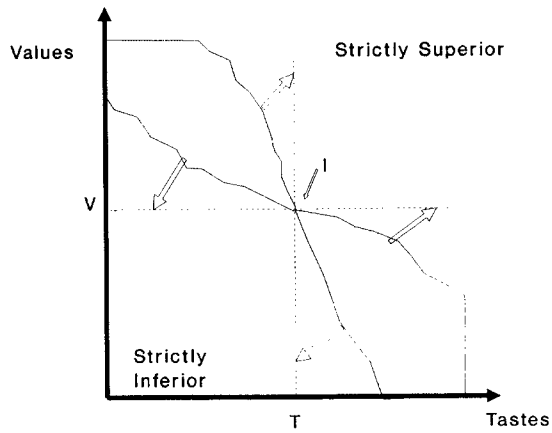

Figure 3. Ambivalence Region Expands to Fill NW and SE Quadrants

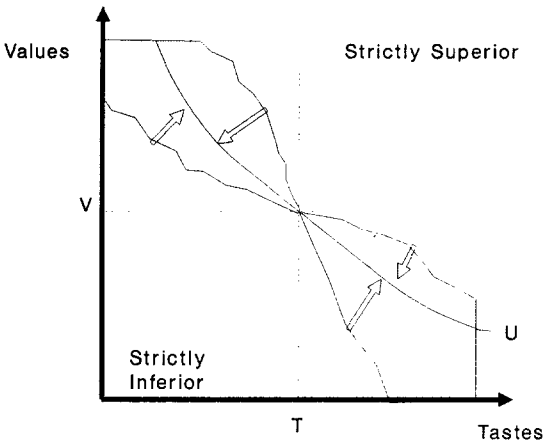

Figure 4. Ambivalence Set Collapses to Ambivalence Curve

spans a continuous range, from complete congruity, where the individual is able to establish precise trades of values and tastes, to complete incongruity where the individual is unable to make any comparison between values and tastes. Complete congruity can be handled using ordinary economic methods, such as state dependent utility functions, but under incomplete congruity neoclassical assumptions can be violated. This kind of decision-making is likely unreliable, inconsistent over time and possibly irreversible.

Consider the Akerlof and Dickens example, where individuals change their beliefs in order to rationalize the decision to accept a trade between dollars and risk of death. Under neoclassical theory, the individual would accept the job if the pay is high enough to compensate for the risk of death. The individual would then act to reduce or avoid the risk. However, if individuals instead choose to change their beliefs to rationalize the choice, a number of 'anomalies' in decision-making may occur. Once their beliefs are changed, individuals will act according to these new beliefs, even if it is potentially detrimental. For example, even if risk reduction is 'inexpensive,' individuals may not act to reduce the risk since they no longer believe that the job is risky. Further, the decision will not be perfectly reversible, so that if the comparative wage differential for the risky occupation is eroded over time, the individual will likely not reverse the occupational decision, since the individual's perception is now that the job is not dangerous.

Individuals who experience ambivalence may also exhibit apparently intransitive choices. As an example, consider the case shown in Figure 5. Here, the individual clearly prefers bundle J to bundle I. However, bundle K is ambivalent to both I and J. Thus, in choosing between bundles J and K, the individual is ambivalent and may choose either bundle so that the individual may be observed choosing bundle K over bundle J. At some other point in time, making the ambivalent decision between bundles K and I, the individual may choose bundle I. Thus, in making two ambivalent decisions, the individual may end up choosing the inferior bundle, I when an alternative set of decisions could have led to the choice of bundle J. Given a direct choice between bundles I and J, however, the individual would clearly choose J. Under the neoclassical paradigm, these choices would be viewed as irrational since they are intransitive.

Ambivalence is also a possible explanation for the observed disparity between WTP and WTA measures of value. To see this we must first derive money measures of compensation under ambivalence. It is relatively straightforward using ordinary economic methods to provide a money measure of compensation for changes in tastes when values are held constant. The non-trivial problem is determining compensation for changes in values. However, given the ability to convert changes in tastes into dollar equivalents, we can frame

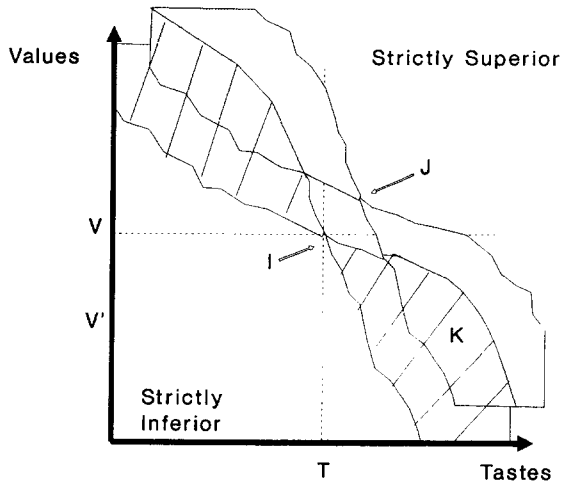

Figure 5. Depiction of 'Intransitive' Preferences Resulting from Ambivalence

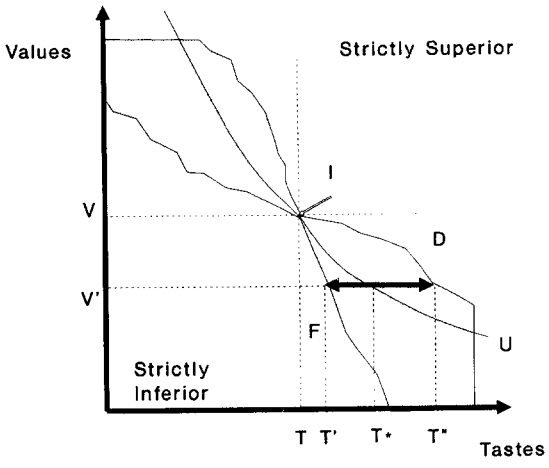

Figure 6. Depiction of Range for Compensation Under Ambivalence

compensation for changes in values in terms of ordering for tastes. That is, if we can find the increase in the ordering for tastes that is required to compensate for a change in values, this is asumed to be translated into a money measure using the usual techniques.

Assume that the individual starts with bundle I, as depicted in Figure 6, which has taste ranking T and value ranking V. Suppose that values are reduced to V′ and we wish to compensate the individual with personal payoffs. The level of personal payoff required to compensate is at least (T′ − T) since otherwise the individual will be in region F which is unambiguously inferior to bundle I. Compensation by more than (T″ − T), on the other hand, will bring the individual to region D, which is superior to the initial bundle. Compensation within the range [(T′ − T), (T″ − T)] will leave the individual within the ambivalence region, where the individual is neither unambiguously better off nor worse off than at the initial point, I. As the ambivalence region collapses to the two dimensional ambivalence curve depicted in the diagram, the range of compensation collapses to the relatively straightforward scalar measure (T* − T). In contrast, as the ambivalence region expands to fill the NW and SE quadrants, the range of compensation expands to [0, ∞] and compensation becomes meaningless.

Thus, compensation is generally not a unique amount which achieves indifference, as defined under neoclassical utility theory. Rather it is a range over which the individual is unsure and full of doubt. If individuals feel ambivalence towards trades between two factors, then they will naturally feel uncomfortable with the idea of increasing the level of one factor in compensation for losses in the other, as the two factors are not precisely comparable. Clearly, an individual facing such an ambivalent decision would find it very difficult to respond precisely and consistently to willingness-to-pay questions in a contingent valuation survey and would feel uncomfortable with the outcome in terms of providing appropriate compensation. This is the same sort of uncomfortable feeling that non-economists seem to exhibit for the concept of estimating the monetary compensation for losses in many types of environmental services.

Note that ambivalence does not imply that no finite amount of money is enough to compensate for losses in values, but rather that the individual cannot make precise trade-offs between tastes and values, so that no unique monetary compensation can be determined. Further, within the imprecise range for compensation the individual would feel unsure and ambivalent, rather than being 'made whole' or indifferent. In the case of complete incongruity, monetary compensation for losses in values is not a logical question, in the same way that it does not make sense to ask the question "How many pounds does the color red weigh?" The two are non-comparable so that they conform to different standards and are subject to different measures.

In the case of partial congruity, the range for compensation under ambivalence could be the cause for some cases of observed divergence between stated WTP and WTA in contingent valuation studies. Faced with a contingent valuation question that attempts to elicit WTP or WTA for losses in values, the individual would be unable to provide a precise response in the case where the ambivalence region is a non-trivial but strict subset of the SE and NW quadrants. Under such circumstances the individuals may behave conservatively when asked questions regarding WTP and WTA. If so, when asked how much they would be willing to pay to avoid a reduction in values, the individual may be unable to determine a 'true,' precise maximum WTP, but they may know that they would certainly be willing to pay at least (T′ − T). When asked how much compensation is required, they again could not provide a precise answer. However, they know that they would certain-

ly be adequately compensated by the amount $(T'' - T)$. Thus, the inability to compare two fundamentally incongruous options combined with a conservative response to the questions can lead to a large divergence in responses to WTA versus WTP questions. This occurs because there does not exist a unique money measure of indifference. Instead there is a range for both WTP and WTA, and a 'conservative' response to questions might reflect the opposite ends of this range for WTP versus WTA.

Discrete choice models could also be applied to WTP for changes in values. The individual's response to a once-and-for-all WTA type question would be uncertain for bids in the ambivalence region, i.e. bids between $(T' - T)$ and $(T'' - T)$. Again by ignoring any source of randomness other than ambivalence, the individual should respond positively with probability 0 if compensation is less than $(T' - T)$ and positively with probability 1 if compensation is greater than $(T'' - T)$. However, within the range $[(T' - T), (T'' - T)]$ one might speculate that the probability of a positive response varies in some sense systematically from 0 to 1. Again, there would be a small probability of a positive response to bids near $(T' - T)$. In contrast, the probability of a positive response would be near one for bids near $(T'' - T)$.

In searching for a unique value for compensation one might be tempted to consider measures that have been used in the context of other discrete choice problems (Hanemann). These include the 'median' bid where the probability of a positive response is 50% and the 'mean' bid which could be estimated by multiplying the bid times the probability of a positive response and integrating over the range from $(T' - T)$ to $(T'' - T)$. Note, however, these measures merely represent two points that leave the individual within the ambivalence region. Thus, it is questionable whether there is any normative significance to these measures in the context of ambivalence.

Returning to the general problem of decision-making under ambivalence, we can represent the preference relations using the following mathematical notation. Bundle I represents the value-taste combination $[V^I, T^I]$, where V^I represents the value ordering of bundle I and T^I represents the ordering of bundle I in terms of tastes or personal payoffs. Bundle J being preferred to I is denoted as $\{J\} > \{I\}$, for which $T^J > T^I$ and $V^J > V^I$ are sufficient, but possibly not necessary. For example, a situation may be unambiguously preferred by the individual if T^J is greatly preferred to T^I, while V^J is only slightly inferior to V^I or if T^J is slightly inferior to T^I and V^J is greatly preferred to V^I, as depicted above by areas C and D in Figure 2. Indifference between I and J is denoted as $\{I\} = \{J\}$, where $T^I = T^J$ and $V^I = V^J$ are sufficient for indifference, but not necessary if the individual is able to make at least some precise tradeoffs between tastes and values. Finally, ambivalence between I and J is denoted as $\{I\} < > \{J\}$. Using this notation, the individual could face bundles such that:

$$\{J\} > \{I\}, \text{ but } \{I\} < > \{K\} \quad \text{and } \{K\} < > \{J\}$$

so that the ambivalence operator is not strictly transitive, as was depicted above in Figure 5.

Conclusions

Neoclassical theory has proven to be a rich paradigm for understanding decision-making within various types of environments. However, despite the great many successes of the standard neoclassical models, an increasing body of evidence has detected systematic violations of the predictions of neoclassical theory. These violations should not be viewed as anomalies of behavior to be ignored or to be explained away as irrational. Rather, this type of behavior poses a challenge to extend the neoclassical paradigm to incorporate other aspects of decision-making in an attempt to explain these results. In order to provide a more comprehensive understanding of decision-making, standard models need to be modified to capture the essential components of certain types of decision environments.

We do not mean to suggest that efforts to expand the neoclassical paradigm would be straightforward; on the contrary, this is quite a difficult task. If we are to use the consistent modelling approach in order to explain fully the decision process, we must begin with the underlying motivations for choice and use them to derive a theory of behavior. If the decision-making process is more complicated than implied by neoclassical theory, then the resulting models of that behavior will be more complicated as well. Even if individuals take steps to simplify decision-making, for example by using rules of thumb, predicting how those simplifications are made or when an individual

will switch into that mode of decision-making may be a difficult empirical task. Richer behavioral theories will also tend to complicate the interpretation of observable behavior for the purpose of inferring preferences from behavior. Finally, verification of the theory is complicated by the absence of a unique correspondence between motivations and behavior.

Despite these difficulties further work in this area has the potential for tremendous payoffs in explaining behavior that is inconsistent with neoclassical theory. We have discussed possible "rational" explanations for some of this apparently irrational behavior. For example, intransitive choices can arise when decision-makers experience regret or rejoicing in making choices (Loomes and Sugden; Bell, 1982). We suggested here that observed intransitivities might also arise from ambivalence. Likewise, the use of rules of thumb or lexicographical choice might be due to complexity (Simon) or to ambivalence. Finally, the observed disparity between willingness-to-pay and willingness-to-accept measures of value can be explained by feelings of regret/rejoicing and the use of reference points (Opaluch and Segerson; Gregory). As suggested here, a third possible explanation is ambivalence.

Given the number of possible "rational" explanations for observed phenomena, generalizations of neoclassical theory may prove to be a highly productive path for research. Thus, rather than discounting "irrational" behavior, our time might be more productively spent exploring when and how to expand upon the neoclassical paradigm in order to explain this behavior. The challenge is to strike a balance between having a paradigm that is richly descriptive of observed behavior and having a tractable approach for modelling that behavior. Although the task before us is difficult, the rewards to success are correspondingly great. Through innovative research efforts in this area, we have the potential to enrich our thought process, our understanding of decision-making and, ultimately, our paradigm in order to encompass a broader spectrum of human behavior.

References

Akerlof, George A. and William T. Dickens. "The Economic Consequences of Cognitive Dissonance." *American Economic Review* 72(1982):307–319.

Bell, David E. "Regret in Decision Making Under Uncertainty." *Operations Research* 30(1982):961–981.

Bell, David E. "Disappointment in Decision Making Under Uncertainty." *Operations Research* 33(1985): 1–27.

Edwards, Ward. "The Theory of Decision Making." *Psychological Bulletin* 51(1954):380–417.

Edwards, Ward. "Conservatism in Human Information Processing." *Formal Representation of Human Judgment*, ed. Benjamin Kleinmuntz. New York: Wiley, 1968.

Gregory, Robin. "Interpreting Measure of Economic Loss: Evidence from Contingent Valuation and Experimental Studies." *Journal of Environmental Economics and Management* 13(1986):325–337.

Hanemann, W. Michael, 1984. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses" *American Journal of Agricultural Economics* 66(1984):332–341.

Hausman, Jerry. "Exact Consumer's Surplus and Deadweight Loss" *American Economic Review* 71(1981):662–676.

Kahneman, Daniel and Amos Tversky. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47(1979):263–291.

Knetsch, Jack L. and J. A. Sinden. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics* 99(1984):507–521.

Lichtenstein, Sarah, et al. "Judged Frequency of Lethal Events." *Journal of Experimental Psychology* 4(1978):551–578.

Loomes, Graham and Robert Sugden. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." *Economic Journal* 92(1982):805–24.

Machina, Mark J. "Choice Under Uncertainty: Problems Solved and Unsolved." *Journal of Economic Perspectives* 1(1987):121–154.

Markowitz, Harry. "The Utility of Wealth." *Journal of Political Economy* 60(1952):151–158.

May, Kenneth O. "Intransitivity, Utility, and the Aggregation of Preference Patterns." *Econometrica* 22(1954):1–13.

Opaluch, James J. "Valuing Natural Resource and Environmental Amenities: Can Economic Valuation Techniques Be Made Defensible: Discussion." *Northeast Journal of Agricultural and Resource Economics* (1984):138–141.

Opaluch, James J. and Kathleen Segerson. "Hicksian Welfare Measures within a Regret Theory Framework." *American Journal of Agricultural Economics* 70(1988):1100–1106.

Ramsey, Frank P. *The Foundations of Mathematics*. New York: Harcourt Brace, 1931.

Savage, Leonard J. *The Foundations of Statistics*. New York: Wiley, 1954.

Schoemaker, Paul J. H. "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations." *Journal of Economic Literature* 20(1982):529–563.

Simon, Herbert A. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*

69(1955):174–83.

Slovic, Paul. "Manipulating the Attractiveness of a Gamble Without Changing its Expected Value." *Journal of Experimental Psychology* 79(1969):139–145.

Tversky, Amos. "Intransitivity of Preferences." *Psychological Review* 76(1969):31–48.

Tversky, Amos and Daniel Kahneman. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5(1973):207–32.

Tversky, Amos and Daniel Kahneman. "The Framing of Decisions and the Psychology of Choice." *Science* 211(1981):453–48.

Weaver, R. D. and S. E. Stefanou. "Toward a Behavioral Approach to Modelling Dynamic Production Choice Structures." *Northeast Journal of Agricultural and Resource Economics* (1984):163–176.