Lecture 1: Non-Rational Choice Behavior and its Rationalizations

Ran Spiegler

August 14, 2007

Abstract

In our first meeting, we recall the concept of rational choice, discuss salient experimental departures from rationality, and address the ways economists often try to rationalize non-rational choice behavior.

Introduction to the Course

This course examines recent attempts to construct models of individual and interactive decision making which depart from the standard rationality assumption, and to embed them in larger economic models. The departures are informed, as usual in economics, by introspection and everyday observation, but also by listening to what experimental psychologists have to say about decision making, and by paying attention to procedural aspects of decision making.

There are several motivations for these attempts to enrich the psychology of the decision maker in economic models. First, there is a sentiment that some economic phenomena cannot be explained without a theory of decision making that departs from rationality. For instance, rationality implies that people maximize according to stable preferences. In many situations, however, it appears that our behavior is a result of a constant battle between long-range planning and visceral urges. Can we understand phenomena such as the diet industry, addiction and rehabilitation, the pricing strategies of credit card companies, or saving for retirement, without models that account for this arguably non-rational aspect of decision making?

Another motivation for these attempts is a sentiment that independently of a particular economic application, certain aspects of the rationality assumption are unrealistic, and that it would be interesting to examine the robustness of standard economic analyses to modifications of the standard model of decision making. For instance, people have limited memory, whereas standard models assume perfect memory. Another example is that people's sense of well-being is often directly affected by their beliefs, whereas rationality implies a purely instrumental attitude to beliefs. How can we model decision making processes that relax these assumptions? Do such models raise interesting new issues? What are the implications of such models in economic environments?

The first motivation takes economic phenomena as a starting point, and seeks to explain them on the basis of decision models that incorporate aspects of the psychology of decision making absent from, or violated by the standard model. Indeed, if we want to understand aspects of the credit card industry, it is good economics to guess that consumers' internal battle between long-term planning and visceral urges is an important part of the story. The second motivation begins with a set of assumptions on individual behavior, and tries to explore its economic implications without any pre-intended desire to explain a particular phenomenon. For instance, people often deliberately choose to hold over-optimistic beliefs. Such "choice of beliefs" is forbidden in the standard model. How can we incorporate deliberate optimism into models of decision making? What are the implications of deliberate optimism for the concept of economic welfare? How does deliberate optimism affect the demand for information? These questions are interesting whether or not deliberate optimism is absolutely necessary for the explanation of some economic phenomenon, because they are motivated by realistic assumptions on individual behavior. This course will generally be based on the second kind of motivations.

Although our discussion in this course relies on a distinction between rational and non-rational behavior, the boundaries between the two are not as clear-cut as one might expect. First, the rationality assumption is agnostic about what motivates decision makers. For instance, a model of rational choice that assumes self-destructive behavior, or a model that assumes that people care only about justifying their actions to others, or a model that assumes that people do not care about absolute consumption levels but about changes relative to a reference point - all these models do not necessarily violate rationality in the broadest sense of the word. However, they do violate rationality as ordinarily practiced by economists. Second, we will see that certain kinds of behavior (e.g., the self-control issues that are arguably so important for the credit card industry) may be viewed as rational or non-rational, depending on the economist's modelling strategy. Perhaps the best title for the course is "economic theory in which agents have a richer psychology than traditionally assumed by economists".

1 The Maximizing Procedure

Rationality of choice is defined in two ways: procedurally and behaviorally. The procedural definition is as follows. The givens are a set X, and a preference relation $\succsim$ on X. For every choice problem $A \subseteq X$, choose a $\succsim$-maximal element in A. For expositional purposes, assume that X is finite, so that existence of a $\succsim$-maximal element in any choice problem is guaranteed. (In economic models, $\succsim$ is typically replaced by a utility function u which represents it. When X is uncountably infinite, some additional conditions are needed for a preference relation to be represented by a utility function.)

In many economic applications, X is endowed with additional structure because we want to incorporate elements of time or uncertainty into the decision model. In these more structured decision models, $\succsim$ is typically assumed to satisfy additional properties (exponential discounting, expected utility). These properties are sometimes so convincing from a normative perspective that their violation is often viewed as violation of rational behavior.

The maximizing procedure means that the agent perfectly identifies the set of feasible alternatives, that he has stable preferences (in particular, the ranking of alternatives is independent of the choice situation he faces), and that he is able to solve the maximization problem. Since these are unrealistic assumptions, economists have developed an "as if" justification for the model: if individual choice behavior satisfies certain consistency properties, then it can be described as if it is induced by the maximization procedure.

The precise consistency conditions depend on whether $\succsim$ is a strict or a weak order. Let us focus on the former case, where the maximizing procedure induces a choice function that selects a unique element from every finite choice set. An arbitrary choice function c assigns an element in A to every $A \subseteq X$. The following consistency condition is known as "*Independence of Irrelevant Alternatives*" (IIA): whenever $x \in A \subset B$, $x = c(B)$ implies $x = c(A)$.

Proposition 1 *c satisfies IIA if and only if there exists a linear ordering $\succ$ on X, such that for every $A \subseteq X$, $c(A)$ is the $\succ$-maximal element in A.*

Proof. To see that IIA is a necessary condition for maximization is easy. Let $A \subset B$. If $c(B)$ is the $\succ$-maximal element in B and $c(B) \in A$, then obviously $c(B)$ is the $\succ$-maximal element in A. Let us now show that IIA is a sufficient condition for maximization. Given a choice function c that satisfies IIA, define $\succ$ as follows: $a \succ b$ if $c\{a, b\} = a$. Since we are restricting attention to choice *functions*, this means that

$b \not\succ a$. Suppose that $a \succ b$ and $b \succ c$. Then, $c\{a, b, c\} = a$ - otherwise, there would be a violation of IIA w.r.t either $c\{a, b\}$ or $c\{b, c\}$. We have established that $\succ$ is complete, asymmetric and transitive, hence it is a linear ordering. It remains to show that $c(A)$ is the $\succ$-maximal element for every $A \subseteq X$. By IIA, $c(A) = c\{a, c(A)\}$ for every $a \in A$, hence $c(A) \succ a$ for every $a \in A$, $a \neq c(A)$. ■

Thus, if a choice function violates IIA, then it is inconsistent with the maximizing procedure. Conversely, a choice function may be a result of a different procedure than the maximizing procedure, and yet be consistent with it. Note that the mere fact that choice behavior can be described by a well-defined choice function already implies a consistency property which is necessary for rational behavior: any two choice situations which admit the same set of feasible alternatives are also supposed to admit the same choices.

In general, the maximizing procedure is based on a weak order $\succsim$ rather than a linear ordering, and as such allows for indifference. The behavioral characterization of rational choice is subtler in this case, because choice behavior is given by a choice *correspondence*, which selects a non-empty *subset* from any choice set. It is not clear how to interpret situations in which $c(A)$ contains more than one element. First, can we distinguish between indifference and indecision? Second, suppose that the decision maker will choose a different element from $c(A)$ in different choice situations in which the set of feasible elements is A but their presentation is different (e.g., the order by which the elements are introduced). Then, choice correspondences are an insufficiently rich description of choice behavior. Finally, the axioms which characterize rational choice correspondences are somewhat less elegant than in the case of choice functions.

2 Systematic Violations of Rationality

Psychologists who study decision making have designed numerous clever choice experiments that elicit violations of rational choice. These experiments are of interest not only because they demonstrate the shortcomings of the maximizing procedure as a descriptive theory of decision making, but also because they uncover some decision procedures other than maximization, which decision makers employ.

Amos Tversky and Daniel Kahneman, who were the chief contributors to this area of research, proposed a nice analogy to justify their research strategy. Scientists who study human vision pay a lot of attention to optical illusions. An optical illusion means that human vision doesn't function well in certain environments. However, human vision functions remarkably well in other environments, which are more relevant for everyday human functioning. However, optical illusions are more informative about the inner workings of human vision, because one learns more about a system when it fails than when it works. The same holds for the study of violations of the axioms of rational choice. By focusing on the violations, we may learn more about the procedures underlying choice behavior.

This section is devoted to a number of salient experimental deviations from rational choice. I bundle these deviations into two categories: choice inconsistencies in choice problems that involve pairs of alternatives, and violations of IIA (which therefore involve choice sets containing more than two alternatives).

A comment on experimental methodology. The experiments we will survey do not present multiple choice problems to the same experimental subject. Instead, two groups are randomly drawn from the same population and offered different choice problems. A choice experiment that identifies a choice inconsistency confronts the two groups of subjects with two choice situations involving a choice between two alternatives. The two problems are meant to be identical as far as the model of rational choice is concerned, and if the two groups display significantly different choice frequencies, we say that the experiment has elicited a choice inconsistency.

A choice experiment that identifies violation of IIA confronts the two groups of subjects with two choice situations, which are meant to differ only in that choice situation A involves a larger set of feasible alternatives than choice situation B. Consider an element a that belongs to the feasible set in both choice situations. If the choice frequency of x is significantly *higher* in choice situation A than in choice situation B, we say that the experiment has elicited a violation of IIA. The reason is that under the null hypothesis, all experimental subjects have well-defined preferences which are drawn from the same distribution. Therefore, if x is a $\succsim$-maximal alternative in the large choice set, then it must be a $\succsim$-maximal alternative in the small choice set.

2.1 Choice Inconsistencies

Let us begin with choice experiments that document inconsistencies in choices from choice sets that consist of two alternatives.

2.1.1 Dynamically Inconsistent Preferences

One real-life violation of this consistency requirement results from the fact that people's choices are sometimes dictated by long-term planning and sometimes by visceral urges, and two often conflict each other. For instance, consider a decision maker who faces a choice between two kinds of dessert for dinner: an unattractive, low-fat dessert and a tasty, fatty dessert. When the choice is made on the spot, the decision maker chooses the fatty dessert. In contrast, when the choice is made twelve hours in advance, he goes for the low-fat dessert.

The following experiment is very well-known. Each group of subjects is faced with a choice between getting \$11 on day t and \$10 on day $t+1$. Group 1 is requested to make the choice on day t, whereas Group 2 is requested to commit to a choice on day $t-1$. Many experimental subjects prefer the earlier payoff in the first treatment and the later payoff in the second treatment.

These are examples of what is called *dynamically inconsistent preferences*: the choice inconsistency is due to the passage of time, or due to a change in circumstances due to the passage of time. We will devote the next lecture to this case.

2.1.2 Reference Point Effects

In a very famous experiment, one group of doctors was presented with the following choice problem. "As a result of an epidemic disease, 600 people are going to die. There are two mutually exclusive treatments: (A) exactly 200 people will be saved ; (B) with a probability of $\frac{1}{3}$, all 600 people will be saved and with a probability of $\frac{2}{3}$, no one will be saved. Which treatment would you choose?" Another group of doctors was presented with the following choice problem. "As a result of an epidemic disease, 600 people are going to die. There are two mutually exclusive treatments: (C) exactly 400 people will die ; (D) with a probability of $\frac{2}{3}$, everyone will die and with a probability of $\frac{1}{3}$, no one will die. Which treatment would you choose?"

The subjects displayed a strong tendency to choose A over B and D over C. Since A is logically equivalent to C and B is logically equivalent to D, a rational decision maker should prefer A to B if and only if he prefers C to D. The difference between A and C, like the difference between B and D, amounts to phrasing alone. The preference reversal is thus an example of what is called a *framing effect*: people display different choice behavior in two choice problems which are logically identical in their content, yet different in presentation: verbal description of alternatives, order in which alternatives are presented, the number of times that an alternative is presented, etc. Of course, we could say that decision makers care about these things as such, and therefore the equivalence between the two choice problems does not hold. However, this is very far from the way the model of rational choice is practiced by economists, and therefore we construe the experimental behavior as evidence for rationality being violated.

One explanation for the preference reversal in this experiment is that the phrasing in each experiment suggests a reference point for assessing outcomes. Describing the outcomes of medical treatments in terms of lives saved implies a reference point in which everybody is dead, so that the prospects are all gains relative to the reference point. Conversely, describing outcomes in terms of lives lost implies a reference point in which everybody is alive, so that the prospects are all losses relative to the reference point. If people tend to be risk averse with respect to prospects that lie above the reference point and risk seeking with respect to prospects that lie below the reference point, then this explains their behavior.

This experiment uncovers an aspect of the underlying decision process, namely the tendency to think about outcomes in terms of changes relative to a reference point. This tendency is the main feature of Tversky and Kahneman's *Prospect Theory*. This feature is especially nice, because it resonates with general psychological principles whose scope extends beyond decision making. Humans can perceive changes in temperature, noise or brightness much more easily than absolute levels.

There is an important connection between choice inconsistencies due to reference points and dynamically inconsistent preferences. When a decision maker faces a temporal sequence of decision problems, his decision today may affect tomorrow's reference point. When preferences are sensitive to reference points, this results in dynamically inconsistent preferences.

2.1.3 Reason-Based Choice

Consider the following experiment. Two groups of subjects are confronted with the same choice between two holiday packages. Each package is described according to its features (weather, beaches, nightlife, hotel). Package 1 is characterized by features which are either very good or very bad, while Package 2 is characterized by rather neutral features. In the first treatment, subjects are told that no reservation has been made yet, and that they need to choose between the two packages. In the second treatment, subjects are told that reservations for both packages have been made, yet the travel agent cannot continue to hold both of them and therefore one of them must be cancelled.

Subjects in the first treatment are more likely to choose Package 1, while subjects in the second treatment are more likely to *cancel* the same package. This behavioral pattern violates (statistically) choice consistency. The choice procedure that seems to underlie this behavioral pattern is a search for reasons to support a decision. When the decision problem is framed in terms of accepting, the decision maker is motivated to search for reasons to accept any alternative, and since Package 1 has an abundance of such reasons, he ends up choosing it. On the other hand, when the decision problem is framed in terms of rejecting, the decision maker is impelled to search for reasons to reject any alternative, and since Package 1 has an abundance of those as well, he ends up rejecting it. The picture of decision making suggested by this experiment is that the decision maker searches for reasons to enable him to construct a preference ranking, and this search process can be manipulated through phrasing .

2.1.4 Dependence on State-Space Description

The following experiment demonstrates another framing effect. In the first treatment, experimental subjects face the following choice between two roulette turns:

A	Color	White	Red	Green	Yellow
	Prob. (%)	90	6	1	3
	Prize ($)	0	45	30	-15

B	Color	White	Red	Green	Yellow
	Prob. (%)	90	7	1	2
	Prize ($)	0	45	-10	-15

In the second treatment, they face a choice between two other roulette turns:

C	Color	White	Red	Green	Blue	Yellow
	Prob. (%)	90	6	1	1	2
	Prize ($)	0	45	30	-15	-15

D	Color	White	Red	Green	Blue	Yellow
	Prob. (%)	90	6	1	1	2
	Prize ($)	0	45	45	-10	-15

Experimental subjects correctly perceive that C is dominated by D, in the sense that for each of the five outcomes of the roulette turn, D prescribes a higher monetary prize than C. In contrast, experimental subjects tend to choose A over B in the first treatment. This is anomalous because D and B induce the same probability distribution over prizes, and B and C induce the same probability distribution over prizes. Since the null hypothesis is that the decision makers do not care about the roulette colors as such, but only about the probability distributions over monetary prizes, their choices are inconsistent.

How can we explain this framing effect? The choice between C and D is easy, because the domination of D is immediately obvious. In contrast, the representation of the lotteries in the first treatment makes it harder to see that B dominates A. The decision makers seem to be trying to resolve the difficulty by simplifying the decision problem. They notice that the two lotteries are exactly identical in their "White" outcome and *virtually* identical in their "Red" and "Yellow" outcomes. If the decision maker cancels out the similarities, he is left with the "Green" outcome, and here A beats B. We see that two ways to frame the same decision problem lead to different decisions. One way encourages the decision maker to apply domination while the other encourages him to simplify the decision problem by canceling out similarities.

2.2 Violations of IIA

Let us now turn to choice experiments that document violations of the axiom of Independence of Irrelevant Alternatives.

2.2.1 The Attraction and Compromise Effects

Here is a compilation of two well-known experiments, once again involving travel agents and holiday packages. The compiled experiment consists of three treatments:

1. A hypothetical travel agent offers the experimental subject two holiday packages: 4 days in London and 7 days in Paris, vs. 7 days in London and 4 days in Paris.

2. Same as the first treatment, except that the travel agent offers an additional, third package: 3 days in London and 6 days in Paris.

3. Same as the second treatment, except that the third package consists of 2 days in London and 9 days in Paris.

In experiments, the package $(4,7)$ has a higher choice frequency in the second and third treatments than in the first scenario, violating regularity. In particular, this means that there are experimental subjects for whom $c\{(4,7),(7,4)\} = (7,4)$ and $c\{(3,6),(4,7),(7,4)\} = (4,7)$, and that there are experimental subjects for whom $c\{(4,7),(7,4)\} = (7,4)$ and $c\{(2,9),(4,7),(7,4)\} = (4,7)$.

One explanation for the behavior in the second treatment is that when $(4, 7)$ dominates another alternative in the choice set, this provides the decision maker with an argument that could justify his choice. We sometimes justify our choices by saying that they are better than other, counter-factual choices. The asymmetric domination in the second treatment provides such an argument. Similarly, when $(4, 7)$ is naturally between the two other options in the third treatment, a decision maker can appeal to a standard argument of moderation to justify his choice of $(4, 7)$. This argument, however, is only valid relative to the particular choice problem. As in the accepting-versus-rejecting experiment discussed earlier, these violations of IIA suggest a choice procedure that involves a search for reasons. But while in the accepting-versus-rejecting experiment the reasons were given and external, here the reasons are internal, and relative to other alternatives in the choice set.

Note that both violations of IIA imply that a clever marketer can manipulate a consumer who chooses in this way into buying a particular product, by adding suitable "irrelevant" alternatives to the choice set, which provide the needed internal reasons that the consumer seeks to justify choosing that product.

2.2.2 Reason-Based Choice

The following is another experiment that provides evidence for a phenomenon discussed above - namely, decision makers' preference for choices that they can justify with good reasons. A group of surgeons need to choose which patient from a set of candidates will be operated on more urgently. In one treatment, there are two patients, a and b, and the characteristics of their cases are quite different. In another treatment, a third patient c is added, and the characteristics of his case are similar to those of patient a. The surgeons display a greater tendency to choose patient b in the second treatment.

The explanation for this violation of IIA seems to be that the surgeons prefer to make a decision they can justify (to potential outside critics or to their own "professional conscience") their choice of one patient over the others. Since the medical cases of the patients contain considerable detail, it is easier to cling to a reason that favors b over both a and c, whereas it is hard to justify why a was chosen over the similar candidate c.

2.2.3 Choice Overload

Several choice experiments have documented a violation of IIA with a particular structure: as the choice set increases in size, decision makers have a tendency to reduce choice complexity by increasing their propensity to choose default alternatives. For instance, in one experiment shoppers had a significantly lower tendency to purchase jam in a large-variety treatment (24 kinds of jam) than in a small-variety treatment (6 kinds of jam). In a field study, workers were found to make lower contributions to pension plans when the set of feasible pension funds was larger.

2.3 Recurrent Themes

Our discussion of choice experiments suggests a picture of decision making, which is quite different from the standard model. The decision maker does not come to the choice experiment with existing, fixed preferences. Instead, he constructs his preferences as he deliberates about the decision problem, and this means that his reasoning is sensitive to many contextual features, including "irrelevant alternatives", defaults, and the description of individual alternatives. The decision maker shuns difficult choices and tries to simplify his decision problem, either by evading choices that are based on intricate trade-offs or by simplifying his description of individual alternatives. He looks for reasons he could use to justify his decisions, and these reasons often involve "irrelevant" comparisons to other alternatives in the choice set.

This description of decision processes is extremely vague by the standards of economic theory. In order to assess the relevance of these ideas, we need to find a way to formalize them. So far, economic theorists haven't been successful at that.

3 Rationalizations

When economists discuss departures from the rational-choice model (at the procedural level or at the level of observed behavior), a question that is invariably asked is whether there is something wrong with the description of the decision maker and in reality he does choose rationally. Throughout the course, as we examine decision models which depart from the standard model, we will have to face this question. Since this turns out to be a very tricky question, I devote the rest of this lecture to it.

3.1 A Non-Rational Procedure which Satisfies IIA: Satisficing

Consider the case in which the decision maker is non-rational at the procedural level yet rational in the behavioral sense. Herbert Simon's "satisficing" procedure is an example. The givens of this procedure are a set X, an enumeration of the elements in X, and a subset $S \subset X$. The procedure is as follows. For every choice problem $A \subseteq X$, the decision maker goes over the elements of A according to the basic enumeration, and picks the first element that belongs to S. If no element in A belongs to S, the decision maker selects the last element in A, according to the basic enumeration.

(We shall restrict attention to finite X. Note that when the choice set is infinite, the procedure may fail to yield a decision, in case $A \cap S$ is empty. This predicament is also possible under rationality, because an infinite set may fail to have a maximal element.)

This procedure describes a search for an alternative that achieves some aspiration level. At the procedural level, it is non-rational because nothing is being maximized. However, the choice function c^* induced by this procedure satisfies IIA, hence it is rational in the behavioral sense. To see why, let $A \subset B$ and suppose that $c^*(B) \in A$. There are two cases in which this is possible: *(i)* the first element in $B \cap S$ according to the basic ordering belongs to A; since $A \subset B$, this element is also the first in $A \cap S$ according to the basic ordering, hence it is chosen in A; *(ii)* the set $B \cap S$ is empty, hence the last element in B according to the basic ordering is chosen; but this means that $A \cap S$ is also empty and the last element in B according to the basic ordering is also the last element in A.

Thus, from a purely behavioral point of view, it is unimportant whether the decision maker employs the maximizing or the satisficing procedures. However, when we try to extrapolate beyond the narrow environment for which the equivalence has been established, the two procedures may diverge. First, when X is a compact, convex set, and choice sets are uncountably infinite, the satisficing procedure makes no sense because it is impossible to enumerate the elements in the choice set. Second, it is often natural in applications to assume that outside parties (e.g., marketers) are able to influence the enumeration through their presentation of the choice set. Thus, it makes sense to ask which of the two procedures seems like a better explanation of choice behavior, even in the domain in which they are behaviorally equivalent. When we construct the utility function that rationalizes the choice function induced by the satisficing procedure, we should ask whether the utility function looks intuitive. I leave this as an exercise to the student.

3.2 Procedures which Over-determine Preferences in Restricted Domains

It often happens that a psychological theory about decision making is incomplete, in the sense that its scope does not extend to the entire set of choice problems. For instance, Tversky and Kahneman's Prospect Theory restricts attention to choices between pairs of monetary lotteries with a particular structure. However, even though the psychological theory does not induce a completely well-defined choice function over the entire domain of choice problems, it may contain enough information to allow us to examine the extent to which the induced behavior can be rationalized by a preference relation over the grand set of alternatives. This sub-section provides two illustrations for this observation.

3.2.1 Similarity-Based Reasoning and Hyperbolic Discounting

First, consider a decision maker who chooses between dated prizes - i.e., pairs $(x,t) \in \mathbb{R}^2_{++}$, where x is the dollar amount received and t is the date on which the amount is received. The decision maker prefers to have more money and dislikes delay, but he finds it hard to calculate trade-offs across the two dimensions, and therefore reasons in the following way when choosing between two pairs, (x,t) and (y,s). If one pair dominates the other - e.g., if $x \geq y$ and $t \leq s$, with at least one strict inequality, the decision is simple and the decision maker chooses the dominant pair. If there is no domination, the decision maker looks for similarities in one dimension and cancels them out, and then chooses the bundle which is better in the remaining dimension. Specifically, let $\lambda \in (0,1)$. If $x/y \in [\lambda, 1/\lambda]$ whereas $t/s < \lambda$, the decision maker chooses (x,t). Similarly, if $t/s \in [\lambda, 1/\lambda]$ whereas $x/y > 1/\lambda$, the decision maker chooses (x,t). If these first two stages do not lead to any resolution, the decision maker chooses according to some unspecified rule. The number λ is a "psychological constant", which determines how close the projection of two vectors need to be in order to count as similar.

This decision model is incomplete in two major respects. First, it restricts attention to choices between pairs. Second, it does not provide a complete algorithm for choice, because it is silent over what happens when the first two stages of the procedure are indeterminate. However, it turns out that even this partial description places strong restrictions on potential rationalizations of any choice behavior which might be consistent with it.

Consider the two points (x,t) and $(y,s) = (x/\lambda, t/\lambda)$. If we wish to rationalize the decision maker's choice behavior with a continuous utility function u over $\mathbb{R}^2_{++}$, then $u(x,t) = u(y,s)$. The reason is simple. According to the procedure, (x,t) is preferred to the pair $(y',s') = (y-\varepsilon, s+\varepsilon)$, where $\varepsilon > 0$ is arbitrarily small. This is because (x,t) and (y',s') are considered similar along the money dimension and dissimilar along the time dimension. The second stage of the choice procedure then implies that

(x,t) is chosen over (y',s'). Therefore, the rationalizing utility function u must satisfy $u(x,t) > u(y',s')$. In contrast, (x,t) is inferior to the bundle $(y'',s'') = (y+\varepsilon, s-\varepsilon)$. This is because (x,t) and (y'',s'') are considered similar along the time dimension and dissimilar along the money dimension. Therefore, the rationalizing utility function u must satisfy $u(x,t) < u(y'',s'')$. Since u is assumed to be continuous, it follows that $u(x,t) = u(y,s)$.

Since this equality is true for every $(x,t) \in \mathbb{R}^2_{++}$, we conclude that any rationalizing utility function must be *approximated* by $u(x,t) = x/t$. This functional form is known as *hyperbolic discounting* in the literature on intertemporal choice. The approximation gets better as λ gets closer to 1. Thus, even without knowing how the decision maker chooses when the first two stages in the procedure do not lead to a clear resolution, these two stages place strong restrictions on the utility function that could possibly rationalize any behavior that is consistent with the procedure.

3.2.2 Small-Scale Risk Aversion and Expected Utility Theory

Imagine a scenario in which an experimental subject is requested to decide whether to accept the following lottery: win \$11 or lose \$10, with probability $\frac{1}{2}$ each. Suppose that the subject rejects the lottery, regardless of his initial wealth level. This behavior is consistent with the following psychological theory (which is, broadly speaking, a specification of Prospect Theory). The decision maker registers outcomes as gains or losses relative to his initial wealth, which serves as a reference point. He applies expected utility with a vNM utility function over outcomes which increases in gains and decreases in losses. However, the vNM function is steeper in the loss range than in the gain range, and there is a "kink" at the reference point. This means that the vNM function is concave around the reference point, which explains why the decision maker rejects the lottery, regardless of his initial wealth. Note that when we repeat the experiment at different levels of initial wealth, we are running the same experiment as far as the above theory is concerned, since outcomes always enter his considerations as gains or losses relative to initial wealth, and so the exact level of his initial wealth is immaterial.

The question is, can we rationalize his behavior by maximization of expected utility over *absolute wealth*? Note that here we are trying to rationalize the decision maker's behavior not with an arbitrary preference relation, but with a preference relation having a particular structure - namely, expected utility over absolute wealth, which is such a workhorse in economic applications.

It turns out that in order to provide such a rationalization for our decision maker's behavior, one must accept preferences which are intuitive very implausible.

Proposition 2 *Suppose that a decision maker rejects the above lottery for every initial wealth $w \in (-\infty, +\infty)$. If this behavior is induced by maximization of expected utility over absolute wealth (according to some vNM utility function u), then he will reject the following lottery: win \$$\infty$ and lose \$100, with probability $\frac{1}{2}$ each.*

Proof. Without loss of generality, let $u(w) - u(w-10) = 1$. By concavity, $u(w-10) - u(w-20) \geq 1$. Again by concavity, $u(w-9) - u(w-10) \geq \frac{1}{10}$. Therefore, $u(w-9) - u(w-20) \geq \frac{11}{10}$. By assumption, $u(w-20) \geq \frac{1}{2}u(w-9) + \frac{1}{2}u(w-30)$. Rearranging, we get: $u(w-20) - u(w-30) \geq \frac{11}{10}$. This argument can be reiterated, such that in general, for every $k \geq 1$:

$$u[w-(2k+1)\cdot 10] - u[w-2k\cdot 10] \geq \frac{11}{10}\cdot\{u[w-(2k-1)\cdot 10] - u[w-(2k-2)\cdot 10]\}$$

This estimate places a lower bound on how steep u gets as we subtract multiplies of 10 from w. Let us now obtain an upper bound on how marginal utility as we add multiples of 11 to w. As before, $u(w) - u(w-10) = 1$. By assumption, $u(w) \geq \frac{1}{2}u(w+11) + \frac{1}{2}u(w-10)$. Therefore, $u(w+11) - u(w) \leq 1$. By concavity, $u(w+11) - u(w+1) \leq \frac{10}{11}$. By assumption, $u(w+11) \geq \frac{1}{2}u(w+22) + \frac{1}{2}u(w+1)$. Rearranging, we get: $u(w+22) - u(w+11) \leq \frac{10}{11}$. This argument can be reiterated, such that in general, for every $k \geq 1$:

$$u[w+(k+1)\cdot 11] - u[w+k\cdot 11] \leq \frac{10}{11}\cdot\{u[w+k\cdot 11] - u[w+(k-1)\cdot 11]\}$$

To get an upper bound on $u(w+\infty)$, all we need to do is calculate the sum of the geometric series $1, \frac{10}{11}, \frac{100}{121}, ...$, and the upper bound is $11 + u(w)$. Now, it is easy to use the divergent series $1, \frac{11}{10}, \frac{121}{100}, ...$ to calculate an upper bound on $u(w-100)$. It is easy to see that $u(w) > \frac{1}{2}u(w+\infty) + \frac{1}{2}u(w-100)$. ∎

Thus, under the assumption of expected utility maximization over absolute wealth, if the decision maker is always risk averse with respect to a small-scale lottery, he must be extremely risk averse with respect to large-scale lotteries. But whereas his risk aversion over small stakes is perfectly reasonable and indeed largely observed in experiments, the induced risk aversion with respect to the large-scale lottery seems absurd. Of course, our intuitions are formed by exposure to initial wealth levels of intermediate size, whereas the proposition deals with the entire range of initial wealth levels. Moreover, the antecedent to the proposition is of course not borne out by actual experiment. In this sense, we are dealing with a thought experiment and not with an actual experiment. Note, however, that calibration results in the same vein can be obtained for restricted ranges of initial wealth levels.

The intuition behind this result is very simple. Expected utility over absolute wealth means that agents are nearly risk neutral with respect to small stakes. The degree of concavity that is needed to rationalize rejection of the "win $11, lose $10 with equal probabilities" is so large, that it generates enormous risk aversion with respect to large stakes. Note that the expected utility assumption itself is not refuted. By re-defining alternatives in terms of gains and losses relative to initial wealth (which constitutes a reference point), rather than in terms of absolute wealth, we can rationalize both small-scale and large-scale risk attitudes.

Like the similarity-based decision model, this exercise demonstrates that when we try to rationalize the choice behavior induced by a non-standard procedure in a restricted domain of choice sets, we may obtain conclusions for the entire domain which are extremely restrictive, and sometimes paradoxically so.

3.3 Rationalization by Multiple Rationales: Choice Sets which Convey Information

In some of the examples examined above, choice behavior violates rationality for one specification of X yet becomes consistent with rationality for another specification of X. For instance, choices may appear inconsistent if we omit reference points from the description of consequences, but consistency is restored if reference points are accounted for. The question of whether choice behavior is rational thus boils down to the question of how to specify the consequence space X. Our judgment of which is the "right" specification depends on common sense and additional details about the context of the choice situation.

The following is a simple illustration of this idea. Imagine that you go into a restaurant and the menu says that you can choose between Chicken and Steak. You opt for Chicken. Then, the waiter comes in and announces that there is also a "special": frogs' legs. You change your mind and choose Steak instead of Chicken. If the set of alternatives X is specified to be the set of dishes, your choice behavior violates IIA. However, the context suggests that this specification is inappropriate - namely, you do not care only about dishes as such, but also about their quality. You have a stable preference: high-quality Steak is better than high-quality Chicken but low-quality Steak is worse than low-quality Chicken. The presence of frogs' legs in the menu conveys information about the quality of the other dishes: it signals that the restaurant offers high-quality dishes. Your choice behavior is rational with respect to this specification of the alternatives.

Here is another example. Whenever you are invited to a tea party where cake is served and slices of different size are available, you consistently choose the second-largest available piece. If the specification of alternatives consists of the "labels" of the cake slices, your choice behavior violates IIA - e.g., $c\{2,3\} = 3$ and $c\{1,2,3\} = 2$. However, if you also care about whether you appear to be polite in the context of a tea party, this might explain why you never want to take the largest slice, because this is less polite than taking the second-largest slice.

The fact that the soundness of the rationalization of choice behavior depends on the context means that we lack a rigid criterion that tells us which specification of X is the "right" one. In addition, the more we tend to believe a priori that the decision maker follows the maximizing procedure, the greater our tendency to look for specifications of X which rationalize choice behavior. Note, however, that without clear constraints on the specification of X, rationalization of a decision maker's choice behavior runs the risk of being no more than a post-hoc explanation.

Let us formalize the idea that non-rational choice behavior can be rationalized if we assume that the choice set convey information to the decision maker and thus systematically influences his preferences. Using this formalization, we can shed some light on the danger that this mode of rationalization may be a trivial, post-hoc explanation. Let $X = \{1, ...n\}$. A choice function c is rationalized by K rationales if there is a collection of K linear orderings $(\succ_k)_{k=1,...,K}$ over X with the property that for every choice set $A \subseteq X$, there exists $k \in \{1, ..., K\}$ such that $c(A)$ is the $\succ_k$-maximal element in A. The interpretation is that the collection of K rationales partitions the set of all choice problems into K cells, so that each cell corresponds to a state of the world. When the decision maker confronts a choice set that belongs to cell k, he learns the state of the world and knows that he should apply the preference relation $\succ_k$.

When $K = 1$ - i.e., when the state space is trivial - choice behavior can be rationalized by a single linear ordering, and thus it is rational in the usual sense. At the other extreme, any choice behavior can be rationalized by n rationales. The construction is trivial: we group all choice sets A for which $c(A) = x$ into a single cell, such that x is the maximal element according to the ordering that corresponds to this cell. Clearly, such rationalization is trivial - essentially, it amounts to saying: "whatever I chose had to be optimal".

In fact, every choice function can be rationalized by $n-1$ rationales, by slightly modifying the above construction. Without loss of generality, denote $c(X) = 1$. For every $x = 1, ..., n-1$, construct $\succ_x$ such that x is the maximal element and n is the second-maximal element. To see how this profile of linear orderings rationalizes c, note that the choice from the grand set X is clearly rationalized by $\succ_1$. For any other choice set A, either $c(A) \in \{1, ..., n-1\}$, in which case the ordering $\succ_{c(A)}$ clearly rationalizes the choice, or $c(A) = n$, in which the ordering for which an element $y \notin A$ is the maximal element (because $A \neq X$, such an element must exist) rationalizes the choice.

Thus, as long as we are allowed to enrich the description of consequences with $n-1$ states, any choice function can be rationalized. Moreover, it can be shown that as n tends to infinity, the proportion of choice functions which cannot be rationalized by less than $n-1$ rationales tends to one. That is, almost all choice functions cannot be rationalized except in the trivial way that fits all choice functions. These results put some content into the impression that without imposing restrictions on rationalization via modification of the consequence space, rationalization may be a trivial post-hoc explanation of choice behavior. The problem is that we lack rigid criteria for judging such restrictions. It follows that the judgment that a particular non-rational choice behavior can be rationalized by modifying X remains a matter of common sense and intuition.

4 Bibliographic Notes

This lecture is partly based on Chapter 1 in Rubinstein (1998). The presentation of the IIA axiom and its relation to the maximizing procedure is standard (see, e.g., Kreps (1988). The experimental evidence on violations of rational choice is taken from Huber, Payne, and Puto (1982), Tversky and Kahneman (1979, 1981,1986), Simonson (1989), Shafir, Simonson, and Tversky (1993), and Redelmeier and Shafir (1995). The discussion of choice overload is based on Iyengar and Lepper (2000) and Iyengar and Kamenica (2006). The satisficing procedure is due to Simon (1982). The calibration result is due to Rabin (2000). The analysis of rationalizing a similarity-based choice procedure is due to Rubinstein (1988). The discussion of rationalization by multiple rationales is based on Sen (1993) and Kalai, Rubinstein and Spiegler (2002).

5 Exercises

1. Find a utility function, whose maximization induces the same choice behavior as the satisficing procedure.

2. Consider the following procedure. The primitives are two numerical functions $u : X \to \mathbb{R}$ and $v : X \to \mathbb{R}$, as well as a number v^*. The decision maker chooses an element that maximizes u as long as its v-value is at least v^*, and chooses an element that maximizes v otherwise. Propose an interpretation for this procedure and check whether it satisfies IIA.

References

[1] Huber, J., J. Payne, and C. Puto. (1982). "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." Journal of Consumer Research 9, 90–98.

[2] Iyengar, S. and E. Kamenica (2006), "Choice Overload and Simplicity Seeking." Mimeo.

[3] Iyengar, S. and M. Lepper (2000), "When Choice is Demotivating: Can One Have Too Much of a Good Thing?" Journal of Personality and Social Psychology, 79, 995–1006.

[4] Kalai G., A. Rubinstein and R. Spiegler (2002). "Rationalizing Choice Functions by Multiple Rationales." Econometrica 70, 2481-2488.

[5] Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica, 47, 263-292.

[6] Kreps, D. M. (1988). Notes on the Theory of Choice. Boulder, Colorado: Westview Press.

[7] Rabin, M. (2000). "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." Econometrica 68, 1281-1292.

[8] Redelmeier, D., and E. Shafir (1995). "Medical decision making in situations that offer multiple alternatives." Journal of the American Medical Association 273, 302-305.

[9] Rubinstein, A. (1988). "Similarity and Decision-Making Under Risk". Journal of Economic Theory 46, 145-153.

[10] Shafir, E., I. Simonson, and A. Tversky. (1993). "Reason-Based Choice." Cognition 49, 11–36.

[11] Sen A. (1993). "Internal Consistency of Choice." Econometrica 61, 495-521.

[12] Simon, H. A. (1982). Models of Bounded Rationality. Cambridge, Mass.: MIT Press.

[13] Simonson, I. (1989). "Choice Based on Reasons: The Case of Attraction and Compromise Effects." Journal of Consumer Research 16, 158–174.

[14] Tversky, A., and D. Kahneman. (1981). "The Framing of Decisions and the Psychology of Choice." Science 211, 453–458.

[15] Tversky, A., and D. Kahneman. (1986). "Rational Choice and the Framing of Decisions." Journal of Business 59, 251–278.