LECTURE 1

Preferences

Preferences

Although we are on our way to constructing a model of rational choice, we begin the course with an "exercise": formulating the notion of "preferences" independently of the concept of choice. We view preferences as the mental attitude of an individual (economic agent) toward alternatives.We seek to develop a "proper" formalization of this concept, which plays such a central role in economics.

Imagine that you want to fully describe the preferences of an agent toward the elements in a given set X. For example, imagine that you want to describe your own attitude toward the universities you apply to before finding out to which of them you have been admitted. What must the description include? What conditions must the description fulfill?

We take the approach that a description of preferences should fully specify the attitude of the agent toward each pair of elements in X. For each pair of alternatives, it should provide an answer to the question of how the agent compares the two alternatives. We present two versions of this question. For each version we formulate the consistency requirements necessary to make the responses "preferences" and examine the connection between the two formalizations.

The Questionnaire Q

Let us think about the preferences on a set X as *answers* to a "long" questionnaire Q which consists of all quiz questions of the type:

Q(x,y) (for all distinct x and y in X)

How do you compare x and y? Tick one and only one of the following three options:

☐ I prefer x to y (this answer is denoted as $x \succ y$).
☐ I prefer y to x (this answer is denoted by $y \succ x$).
☐ I am indifferent (this answer is denoted by I).

A "legal" answer to the questionnaire is a response in which the respondent ticks exactly one of the boxes in each question. We do not allow the decision maker to refrain from answering a question or to tick more than one answer. Furthermore, we do not allow him to respond with answers that demonstrate a lack of ability to compare, such as:

☐ They are incomparable.
☐ I don't know what x is.
☐ I have no opinion.

Or a dependence on other factors such as:

☐ It depends on what my parents think.
☐ It depends on the circumstances (sometimes I prefer x but usually I prefer y).

Or the intensity of preferences such as:

☐ I somewhat prefer x.
☐ I love x and I hate y.

Or confusion such as:

☐ I both prefer x over y and y over x.
☐ I can't concentrate right now.

The constraints that we place on the legal responses of the agents constitute our implicit assumptions. Particularly important are the assumptions that the elements in the set X are all comparable, that the individual has an opinion about all elements in the set X and that we do not allow him to specify the intensity of preferences.

A legal answer to the questionnaire can be formulated as a function f which assigns to any pair (x, y) of distinct elements in X exactly one of the three "values": $x \succ y$ or $y \succ x$ or I, with the interpretation that $f(x, y)$ is the answer to the question $Q(x, y)$. (Alternatively, we can use the terminology of the soccer "betting" industry and say that $f(x, y)$ must be 1, 2, or $\times$ with the interpretation that

$f(x, y) = 1$ means that x is better than y, $f(x, y) = 2$ means that y is better than x and $f(x, y) = \times$ means indifference.)

Not all legal answers to the questionnaire Q qualify as *preferences over the set* X. We will adopt two "consistency" restrictions:

First, the answer to $Q(x, y)$ must be identical to the answer to $Q(y, x)$. In other words, we want to exclude the common "framing effect" by which people who are asked to compare two alternatives tend to prefer the "first" one.

Second, we require that the answers exhibit "transitivity." In other words, the answers to $Q(x, y)$ and $Q(y, z)$ must be consistent with the answer to $Q(x, z)$ in the following sense: If "x is preferred to y" and "y is preferred to z" then "x is preferred to z," and if the answers to the two questions $Q(x, y)$ and $Q(y, z)$ are "indifference" then so is the answer to $Q(x, z)$.

To summarize, here is my favorite formalization of the notion of preferences:

Definition 1

Preferences on a set X are a function f that assigns to any pair (x, y) of distinct elements in X exactly one of the three "values" $x \succ y$, $y \succ x$ or I so that for any three different elements x, y and z in X, the following two properties hold:

- *No order effect*: $f(x, y) = f(y, x)$.
- *Transitivity*:
  if $f(x, y) = x \succ y$ and $f(y, z) = y \succ z$ then $f(x, z) = x \succ z$ and
  if $f(x, y) = I$ and $f(y, z) = I$ then $f(x, z) = I$.

Note again that I, $x \succ y$, and $y \succ x$ are merely symbols representing verbal answers. Needless to say, the choice of symbols is not an arbitrary one. (Why do I use the notation I and not $x \sim y$?)

A Discussion of Transitivity

The transitivity property is an appealing property of preferences. How would you react if somebody told you he prefers x to y, y to z and z to x? You would probably feel that his answers are "confused." Furthermore, it seems that, when confronted with an intransitivity in their responses, people are embarrassed and want to change their answers.

Before the lecture, students in Tel Aviv had to fill out a questionnaire similar to Q regarding a set X that contains nine alternatives, each specifying the following four characteristics of a travel package: location (Paris or Rome), price, quality of the food, and quality of the lodgings. The questionnaire included only thirty six questions since for each pair of alternatives x and y, only one of the questions, $Q(x, y)$ or $Q(y, x)$, was randomly selected to appear in the questionnaire (thus the dependence on order of an individual's response could not be checked within the experimental framework). In the 2004 group, out of eighteen MA students, only two had no intransitivities in their answers, and the average number of triples in which intransitivity existed was almost nine. Many of the violations of transitivity involved two alternatives that were actually the same, but differed in the order in which the characteristics appeared in the description. "A weekend in Paris at a four-star hotel with food quality Zagat 17 for \$574," and "A weekend in Paris for \$574 with food quality Zagat 17 at a four-star hotel." All students expressed indifference between the two alternatives, but in a comparison of these two alternatives to a third alternative—"A weekend in Rome at a five-star hotel with food quality Zagat 18 for \$612"—half of the students gave responses that violated transitivity.

In spite of the appeal of the transitivity requirement, note that when we assume that the attitude of an individual toward pairs of alternatives is transitive, we are excluding individuals who base their judgments on "procedures" that cause systematic violations of transitivity. The following are two such examples.

1. *Aggregation of considerations as a source of intransitivity.* In some cases, an individual's attitude is derived from the aggregation of more basic considerations. Consider, for example, a case where $X = \{a, b, c\}$ and the individual has three primitive considerations in mind. The individual finds one alternative better than the other if a majority of considerations support the first alternative. This aggregation process can yield intransitivities. For example, if the three considerations rank the alternatives as follows: $a \succ_1 b \succ_1 c$, $b \succ_2 c \succ_2 a$ and $c \succ_3 a \succ_3 b$, then the individual determines a to be preferred over b, b over c, and c over a, thus violating transitivity.

2. *The use of similarities as an obstacle to transitivity*. In some cases, the decision maker expresses indifference in a comparison between two elements that are too "close" to be distinguishable. For example, let $X = \Re$ (the set of real numbers). Consider an individual whose attitude is "the more the better"; however, he finds it impossible to determine whether a is greater than b unless the difference is at least 1. He will assign $f(x, y) = x \succ y$ if $x \geq y - 1$ and $f(x, y) = I$ if $|x - y| < 1$. This is not a preference relation since $1.5 \sim 0.8$ and $0.8 \sim 0.3$, but it is not true that $1.5 \sim 0.3$.

Did we require too little? Another potential criticism of our definition is that our assumptions might have been too weak and that we did not impose some reasonable further restrictions on the concept of preferences. That is, there are other similar consistency requirements we may impose on a legal response to qualify it as a description of preferences. For example, if $f(x, y) = x \succ y$ and $f(y, z) = I$, we would naturally expect that $f(x, z) = x \succ z$. However, this additional consistency condition was not included in the above definition since it follows from the other conditions: If $f(x, z) = I$, then by the assumption that $f(y, z) = I$ and by the no order effect, $f(z, y) = I$, and thus by transitivity $f(x, y) = I$ (a contradiction). Alternatively, if $f(x, z) = z \succ x$, then by no order effect $f(z, x) = z \succ x$, and by $f(x, y) = x \succ y$ and transitivity $f(z, y) = z \succ y$ (a contradiction).

Similarly, note that for any preferences f, we have if $f(x, y) = I$ and $f(y, z) = y \succ z$, then $f(x, z) = x \succ z$.

The Questionnaire *R*

A second way to think about preferences is through an imaginary questionnaire R consisting of all questions of the type:

R(x,y) (for all $x, y \in X$, not necessarily distinct).

"Is x at least as preferred as y?" Tick one and only one of the following two options:

☐ Yes
☐ No

By a "legal" response we mean that the respondent ticks exactly one of the boxes in each question. To qualify as preferences a legal response must also satisfy two conditions:

1. The answer to at least one of the questions $R(x, y)$ and $R(y, x)$ must be Yes. (In particular, the "silly" question $R(x, x)$ which appears in the questionnaire must get a Yes response.)
2. For every $x, y, z \in X$, if the answers to the questions $R(x, y)$ and $R(y, z)$ are Yes, then so is the answer to the question $R(x, z)$.

We identify a response to this questionnaire with the binary relation $\succsim$ on the set X defined by $x \succsim y$ if the answer to the question $R(x, y)$ is Yes.

(*Reminder*: An *n-ary relation* on X is a subset of X^n. Examples: "Being a parent of" is a binary relation on the set of human beings; "being a hat" is an unary relation on the set of objects; "$x + y = z$" is a 3-nary relation on the set of numbers; "x is better than y more than x' is better than y'" is 4-nary relation on a set of alternatives, etc. An *n-ary relation* on X can be thought of as a response to a questionnaire regarding all n-tuples of elements of X where each question can get only a Yes/No answer.)

This brings us to the "traditional" definition:

Definition 2

A *preference on* a set X is a binary relation $\succsim$ on X satisfying:

- *Completeness*: For any $x, y \in X$, $x \succsim y$ or $y \succsim x$.
- *Transitivity*: For any $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.

The Equivalence of the Two Definitions

We have presented two definitions of preferences on the set X. We now proceed to show their equivalence. There are many ways to construct "a one-to-one correspondence" between the objects satisfying the two definitions. But, when we think about the equivalence of two definitions in economics we are thinking about much more than the existence of a one-to-one correspondence: the correspondence has to *preserve the interpretation*. Note the similarity to the notion of an isomorphism in mathematics. For example, an isomorphism between two topological spaces X and Y is a one-to-one

Table 1.1

A response to $Q(x, y)$ and $Q(y, x)$	A response to $R(x, y)$ and $R(y, x)$
$x \succ y$	Yes, No
I	Yes, Yes
$y \succ x$	No, Yes

function from X onto Y that is required to preserve the open sets. In economics, the one-to-one correspondence is required to preserve the more informal concept of interpretation.

We will now construct a one-to-one and onto correspondence, *Translation*, between answers to Q that qualify as preferences by the first definition and answers to R that qualify as preferences by the second definition, such that the correspondence preserves the meaning of the responses to the two questionnaires. In other words, *Translation* is a "bridge" between the responses to Q that qualify as preferences and the responses to R that qualify as preferences.

To illustrate the correspondence imagine that you have two books. Each page in the first book is a response to the questionnaire Q which qualifies as preferences by the first definition. Each page in the second book is a response to the questionnaire R which qualifies as preferences by the second definition. The correspondence matches each page in the first book with a unique page in the second book, so that a reasonable person will recognize that the different responses to the two questionnaires reflect the same mental attitudes towards the alternatives.

Since we assume that the answers to all questions of the type $R(x, x)$ are "Yes," the classification of a response to R as a preference only requires the specification of the answers to questions $R(x, y)$, where $x \neq y$. Table 1.1 presents the translation of responses.

This translation preserves the interpretation we have given to the responses, that is, "I prefer x to y" has the same meaning as the statement "I find x to be at least as good as y, but I don't find y to be at least as good as x."

The following observations complete the proof that *Translation* is indeed a one-to-one correspondence from the set of preferences, as given by definition 1, onto the set of preferences as given by definition 2.

By the assumption on Q of a no order effect, for any two alternatives x and y, one and only one of the following three answers was received for both $Q(x, y)$ and $Q(y, x)$: $x \succ y$, I and $y \succ x$. Thus, the responses to $R(x, y)$ and $R(y, x)$ are well defined.

Next we verify that the response to R that we have constructed with the table is indeed a preference relation (by the second definition).

Completeness: In each of the three rows, the answers to at least one of the questions $R(x, y)$ and $R(y, x)$ is affirmative.

Transitivity: Assume that the answers to $R(x, y)$ and $R(y, z)$ are affirmative. This implies that the answer to $Q(x, y)$ is either $x \succ y$ or I, and the answer to $Q(y, z)$ is either $y \succ z$ or I. Transitivity of Q implies that the answer to $Q(x, z)$ must be $x \succ z$ or I, and therefore the answer to $R(x, z)$ must be affirmative.

To see that *Translation* is indeed a one-to-one correspondence, note that for any two different responses to the questionnaire Q there must be a question $Q(x, y)$ for which the responses differ; therefore, the corresponding responses to either $R(x, y)$ or $R(y, x)$ must differ.

It remains to be shown that the range of the *Translation* function includes all possible preferences as defined by the second definition. Let $\succsim$ be preferences in the traditional sense (a response to R). We have to specify a function f, a response to Q, which is converted by *Translation* to $\succsim$. Read from right to left, the table provides us with such a function f.

By the completeness of $\succsim$, for any two elements x and y, one of the entries in the right-hand column is applicable (the fourth option, that the two answers to $R(x, y)$ and $R(y, x)$ are "No," is excluded), and thus the response to Q is well defined and by definition satisfies no order effect.

We still have to check that f satisfies the transitivity condition. If $F(x, y) = x \succ y$ and $F(y, z) = y \succ z$, then $x \succsim y$ and not $y \succsim x$ and $y \succsim z$ and not $z \succsim y$. By transitivity of $\succsim$, $x \succsim z$. In addition, not $z \succsim x$ since if $z \succsim x$, then the transitivity of $\succsim$ would imply $z \succsim y$. If $F(x, y) = I$ and $F(y, z) = I$, then $x \succsim y$, $y \succsim x$, $y \succsim z$ and $z \succsim y$. By transitivity of $\succsim$, both $x \succsim z$ and $z \succsim x$, and thus $F(x, z) = I$.

Summary

From now on we will use the second definition, that is, a *preference on X* is a binary relation $\succsim$ on a set X satisfying Completeness and

Transitivity. For a preference relation $\succsim$, we will use the notation $x \sim y$ when both $x \succsim y$ and $y \succsim x$; the notation $x \succ y$ will stand for if $x \succsim y$ and not $y \succsim x$.

Bibliographic Notes

Recommended readings: Kreps 1990, 17–24; Mas-Colell et al. 1995, chapter 1, A–B.

Fishburn (1970) contains a comprehensive treatment of preference relations.

Problem Set 1

Problem 1. (*Easy*)
Let $\succsim$ be a preference relation on a set X. Define $I(x)$ to be the set of all $y \in X$ for which $y \sim x$.

Show that the set (of sets!) $\{I(x) | x \in X\}$ is a partition of X, i.e.,

- For all x and y, either $I(x) = I(y)$ or $I(x) \cap I(y) = \emptyset$.
- For every $x \in X$, there is $y \in X$ such that $x \in I(y)$.

Problem 2. (*Standard*)
Kreps (1990) introduces another formal definition for preferences. His primitive is a binary relation P interpreted as "strictly preferred." He requires P to satisfy:

- *Asymmetry*: For no x and y do we have both xPy and yPx.
- *Negative-Transitivity*: For all x, y, and $z \in X$, if xPy, then for any z either xPz or zPy (or both).

Explain the sense in which Kreps' formalization is equivalent to the traditional definition.

Problem 3. (*Standard*)
In economic theory we are often interested in other types of binary relations, for example, the relation xSy: "x and y are almost the same." Suggest properties that would correspond to your intuition about such a concept.

Problem 4. (*Difficult.* Based on Kannai and Peleg 1984.)
Let Z be a finite set and let X be the set of all nonempty subsets of Z. Let $\succsim$ be a preference relation on X (not Z).

Consider the following two properties of preference relations on X:

a. If $A \succsim B$ and C is a set disjoint to both A and B, then $A \cup C \succsim B \cup C$, and
   if $A \succ B$ and C is a set disjoint to both A and B, then $A \cup C \succ B \cup C$.
b. If $x \in Z$ and $\{x\} \succ \{y\}$ for all $y \in A$, then $A \cup \{x\} \succ A$, and
   if $x \in Z$ and $\{y\} \succ \{x\}$ for all $y \in A$, then $A \succ A \cup \{x\}$.

Discuss the plausibility of the properties in the context of interpreting $\succsim$ as the attitude of the individual toward sets from which he will have to make a choice at a "second stage."

Provide an example of a preference relation that

- Satisfies the two properties.
- Satisfies the first but not the second property.
- Satisfies the second but not the first property.

Show that if there are x, y, and $z \in Z$ such that $\{x\} \succ \{y\} \succ \{z\}$, then there is no preferene relation satisfying both properties.

Problem 5. (*Fun*)
Listen to the illusion called the Shepard Scale. (Currently, it is available at http://www.sandlotscience.com/Ambiguous/ShpTones1.htm and http://asa.aip.org/demo27.html.)

Can you think of any economic analogies?