Lecture Notes in

Microeconomic Theory

second edition

The Economic Agent

Ariel Rubinstein

Lecture Notes in Microeconomic Theory

The Economic Agent

Second Edition

Ariel Rubinstein

Lecture Notes in
Microeconomic Theory

Preface

This is the second edition of my lecture notes for the first quarter of a microeconomics course for PhD (or MA) economics students. The lecture notes were developed over a period of 20 years during which I taught the course at Tel Aviv, Princeton, and New York universities.

I published the book for the first time in 2007 with some hesitation since several superb books were already on the shelves. Foremost among them is Kreps (1990), which pioneered the shift of the game theoretic revolution from research papers into textbooks. His book covers the material in depth and includes many ideas for future research. Mas-Colell, Whinston, and Green (1995) continued this trend with a very comprehensive and detailed textbook. There are three other books on my short list: Bowles (2003), which brings economics back to its authentic political economics roots; Jehle and Reny (1997), with its very precise style; and the classic Varian (1984). These five books constitute an impressive collection of textbooks for the standard advanced microeconomics course. My book covers only the first quarter of the standard course. It does not aim to compete with these books, but rather to supplement them. I published it only because I think that some of the didactic ideas presented might be beneficial to both students and teachers, and it is to this end that I insisted on retaining its lecture notes style.

Downloading Updated Versions
The book is posted on the Internet, and access is entirely free. I am grateful to Princeton University Press for allowing it to be downloaded for free right after publication. Since 2007, I have updated the book annually, adding material and correcting mistakes. My plan is to continue revising the book annually. To access the latest electronic version go to: http://arielrubinstein.tau.ac.il.

Solution Manual
Teachers of the course can also get an updated solution manual. I do my best to send the manual only to teachers of a graduate course in microeconomics. Requests for the manual should be made at: http://gametheory.tau.ac.il/microtheory.

Gender

Throughout the book I use only male pronouns. This is my deliberate choice and does not reflect the policy of the editors or the publishers. I believe that continuous reminders of the he/she issue simply divert readers' attention. Language is of course very important in shaping our thinking, and I don't dispute the importance of the type of language we use. But I feel it is more effective to raise the issue of discrimination against women in the discussion of gender-related issues rather than raising flags on every page of a book on economic theory.

Acknowledgments

I would like to thank all my teaching assistants, who made helpful comments during the many years I taught the course: Rani Spiegler, Kfir Eliaz, Yoram Hamo, Gabi Gayer, and Tamir Tshuva at Tel Aviv University; Bilge Yilmaz, Ronny Razin, Wojciech Olszewski, Attila Ambrus, Andrea Wilson, Haluk Ergin, and Daisuke Nakajima at Princeton; and Sophie Bade and Anna Ingster at NYU. Sharon Simmer and Rafi Aviav helped me with the English editing. Avner Shlain prepared the index. Special thanks to Rafi Aviav and Benjamin Bachi for their devoted work in producing the revised versions of the book.

Introduction

As a new graduate student, you are at the beginning of a new stage of your life. In a few months you will be overloaded with definitions, concepts, and models. Your teachers will be guiding you into the wonders of economics and will rarely have the time to stop to raise fundamental questions about what these models are supposed to mean. It is not unlikely that you will be brainwashed by the professional-sounding language and hidden assumptions. I am afraid I am about to initiate you into this inevitable process. Still, I want to use this opportunity to pause for a moment and alert you to the fact that many economists have strong and conflicting views about what economic theory is. Some see it as a *set of theories* that can (or should) be tested. Others see it as a *bag of tools* to be used by economic agents. Many see it as a *framework* through which professional and academic economists view the world.

My own view may disappoint those of you who have come to this course with practical motivations. In my view, economic theory is no more than an arena for the *investigation of concepts* we use in thinking about economics in real life. What makes a theoretical model "economics" is that the concepts we are analyzing are taken from real-life reasoning about economic issues. Through the investigation of these concepts, we indeed try to understand reality better, and the models provide a language that enables us to think about economic interactions in a systematic way. But I do not view economic models as an attempt to describe the world or to provide tools for predicting the future. I object to looking for an ultimate truth in economic theory, and I do not expect it to be the foundation for any policy recommendation. Nothing is "holy" in economic theory and everything is the creation of people like yourself.

Basically, this course is about a certain class of economic *concepts* and *models*. Although we will be studying formal concepts and models, they will always be given an interpretation. An economic model differs substantially from a purely mathematical model in that it is a *combination* of a mathematical model and its *interpretation*. The names of the mathematical objects are an integral part of an economic model. When mathematicians use terms such as "field" or "ring" that are in everyday use, it is only for the sake of convenience. When they name a

collection of sets a "filter", they are doing so in an associative manner; in principle, they could call it "ice cream cone". When they use the term "good ordering", they are not making an ethical judgment. In contrast to mathematics, interpretation is an essential ingredient of any economic model.

The word "model" sounds more scientific than "fable" or "fairy tale", but I don't see much difference between them. The author of a fable draws a parallel to a situation in real life and has some moral he wishes to impart to the reader. The fable is an imaginary situation that is somewhere between fantasy and reality. Any fable can be dismissed as being unrealistic or simplistic, but this is also the fable's advantage. Being something between fantasy and reality, a fable is free of extraneous details and annoying diversions. In this unencumbered state, we can clearly discern what cannot always be seen from the real world. On our return to reality, we are in possession of some sound advice or a relevant argument that can be used in the real world. We do exactly the same thing in economic theory. Thus, a good model in economic theory, like a good fable, identifies a number of themes and elucidates them. We perform thought exercises that are only loosely connected to reality and have been stripped of most of their real-life characteristics. However, in a good model, as in a good fable, something significant remains. One can think about this book as an attempt to introduce the characters that inhabit economic fables. Here, we observe the characters in isolation. In models of markets and games, we further investigate the interactions between the characters.

It is my hope that some of you will react and attempt to change what is currently called economic theory and that you will acquire alternative ways of thinking about economic and social interactions. At the very least, this course should teach you to ask hard questions about economic models and the sense in which they are relevant to real-life economics. I hope that you walk away from this course with the recognition that the answers are not as obvious as they might appear.

Microeconomics

In this course we deal only with microeconomics, a collection of models in which the primitives are details about the behavior of units called *economic agents*. Microeconomic models investigate assumptions about economic agents' activities and about interactions between these agents. An economic agent is the basic unit operating in the model. When we

construct a model with a particular economic scenario in mind, we might have some degree of freedom regarding whom we take to be the economic agents. Most often, we do have in mind that the economic agent is an individual, a person with one head, one heart, two eyes, and two ears. However, in some economic models, an economic agent is taken to be a nation, a family, or a parliament. At other times, the "individual" is broken down into a collection of economic agents, each operating in distinct circumstances, and each regarded as an economic agent.

We should not be too cheerful about the statement that an economic agent in microeconomics is not constrained to being an individual. The facade of generality in economic theory might be misleading. We have to be careful and aware that when we take an economic agent to be a group of individuals, the reasonable assumptions we might impose on it are distinct from those we might want to impose on a single individual. For example, although it is quite natural to talk about the will of a person, it is not clear what is meant by the will of a group when the members of the group differ in their preferences.

An economic agent is described in our models as a unit that responds to a scenario called a *choice problem*, where the agent must make a choice from a set of available alternatives. The economic agent appears in the microeconomic model with a specified deliberation process he uses to make a decision. In most of current economic theory, the deliberation process is what is called *rational* choice. The agent decides what action to take through a three-step process:

1. He asks himself, what is desirable?
2. He asks himself, what is feasible?
3. He chooses the most desirable from among the feasible alternatives.

Note the order of the stages. In particular, the stage in which desires are shaped precedes the stage in which feasible alternatives are recognized, and therefore the rational economic agent's desires are independent of the set of alternatives. Note that rationality in economics does not contain judgments about desires. A rational agent can have preferences that the entire world views as being against the agent's interest.

Furthermore, economists are fully aware that almost all people, almost all the time, do not practice this kind of deliberation. Nevertheless, until recently the practice of most economists was to make further assumptions that emphasize the materialist desires of the economic agent and minimize the role of the psychological motives. This practice has been

somewhat changed in the past few years with the development of the "Economics and Psychology" approach. Still, we find the investigation of economic agents who follow the rational process to be important, because we often refer to rational decision making in life as an ideal process. It is meaningful to talk about the concept of "being good" even in a society where all people are evil; similarly, it is meaningful to talk about the concept of a "rational man" and about the interactions between rational economic agents even if all people systematically behave in a nonrational manner.

Bibliographic Notes

For an extended discussion of my views about economic theory, see Rubinstein (2006).

Lecture Notes in
Microeconomic Theory

Contents

Preferences

Preferences

Our economic agent will soon be advancing to the stage of economic models. Which of his characteristics will we be specifying in order to get him ready? We might have thought name, age and gender, personal history, cognitive abilities and knowledge, and his mental state. However, in most of economic theory, we specify an economic agent only by his attitude toward the elements in some relevant set, and usually we assume that his attitude is expressed in the form of *preferences*.

We begin the course with a modeling "exercise": we seek to develop a "proper" formalization of the concept of preferences. Although we are on our way to constructing a model of rational choice, we will think about the concept of preferences here independently of choice. This is quite natural. We often use the concept of preferences not in the context of choice. For example, we talk about an individual's tastes over the paintings of the masters even if he never makes a decision based on those preferences. We refer to the preferences of an agent were he to arrive tomorrow on Mars or travel back in time and become King David even if he does not believe in the supernatural.

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 that 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 exactly one of the boxes is ticked in each question. We do not allow refraining from answering a question or ticking more than one answer. Furthermore, by allowing only the above three options we exclude responses that demonstrate:

a lack of ability to compare, such as

☐ They are incomparable.
☐ I don't know what x is.
☐ I have no opinion.
☐ I prefer both x over y and y over x.

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).

and, most importantly, intensity of preferences, such as

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

The constraints that we place on the legal responses of the agents constitute our implicit assumptions. Particularly important are the assumption that the elements in the set X are all comparable and the fact that we ignore 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 notation 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 preferred to y, $f(x, y) = 2$ means that y is preferred to 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 to $Q(x, y)$ and $Q(y, z)$ are consistent with the answer to $Q(x, z)$ in the following sense. If the answers to the two questions $Q(x, y)$ and $Q(y, z)$ are "x is preferred to y" and "y is preferred to z", then the answer to $Q(x, z)$ must be "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

Transitivity 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.

On some occasions before giving this lecture, I asked students 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 was not checked within the experimental framework). Out of 458 students who responded to the questionnaire, only 57 (12%) had no intransitivities in their answers, and the median number of triples in which intransitivity existed was 7. 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 4-star hotel with food quality Zagat 17 for $574", and "A weekend in Paris for $574 with food quality Zagat 17 at a 4-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 5-star hotel with food quality Zagat 18 for $612"— a quarter 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 an alternative x better than an alternative y if a majority of considerations supports x. This aggregation process can yield intransitivities. For example, if the three considerations rank the alternatives as $a \succ_1 b \succ_1 c$, $b \succ_2 c \succ_2 a$, 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, an individual may express indifference in a comparison between two elements that are too "close" to be distinguishable. For example, let $X = \mathbb{R}$ (the set of real numbers). Consider an individual whose

attitude toward the alternatives is "the larger 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 because $1.5 \sim 0.8$ 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 want to 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 because 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 the 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 that 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 $R(x, x)$ that appears in the questionnaire must get a Yes response.)