Econometrica, Vol. 74, No. 4 (July, 2006), 865–883

DILEMMAS OF AN ECONOMIC THEORIST

BY ARIEL RUBINSTEIN [1]

What on earth are economic theorists like me trying to accomplish? This paper discusses four dilemmas encountered by an economic theorist:

The dilemma of absurd conclusions: Should we abandon a model if it produces absurd conclusions or should we regard a model as a very limited set of assumptions that will inevitably fail in some contexts?

The dilemma of responding to evidence: Should our models be judged according to experimental results?

The dilemma of modelless regularities: Should models provide the hypothesis for testing or are they simply exercises in logic that have no use in identifying regularities?

The dilemma of relevance: Do we have the right to offer advice or to make statements that are intended to influence the real world?

KEYWORDS: Economic theory, game theory, hyperbolic discounting, time response, jungle, economic education, fable.

1. AN ECONOMIC THEORIST'S MOTIVATION

I COULD SAY that this talk will be about some of the research I have been involved in over the past few years. I could also say that it will express my dilemmas over the efficacy of economic theory with the realization that my views constitute an inseparable part of who I am. My comments might even be interpreted as "an outpouring from a therapist's couch," as the referee described them. However, underlying this paper is one major question that I ask myself obsessively: What on earth am I doing? What are we trying to accomplish as economic theorists? We essentially play with toys called models. We have the luxury of remaining children over the course of our entire professional lives and we are even well paid for it. We get to call ourselves economists and the public naively thinks that we are improving the economy's performance, increasing the rate of growth, or preventing economic catastrophes. Of course, we can justify this image by repeating some of the same fancy sounding slogans we use in our grant proposals, but do we ourselves believe in those slogans?

I recall a conference I attended in Lumini, France, in the summer of 1981 that was attended by the giants of the game theory profession. They were standing around in a beautiful garden waiting for dinner after a long day of sessions. Some of us, the more junior game theoreticians, were standing off to the side eavesdropping on their conversation. They loudly discussed the relevance of game theory and one of them suggested that we are just "making a living."

[1]This paper was presented as the Presidential Address to the Econometric Society in Madrid in 2004. I am grateful to all those who provided comments, especially Rani Spiegler and a coeditor of the journal.

I think he merely intended to be provocative, but nonetheless his response traumatized me. Are we no more than "economic agents" maximizing our utility? Are we members of an unproductive occupation that only appears to others to be useful?

Personally, I did not fulfill any childhood fantasy by becoming a professor. It was never my dream to become an economist. Frankly, I respect philosophers, teachers, writers, and nurses more than I do economists. I don't care about stock market prices and I'm not sure I know what "equities" are. I am reluctant to give policy advice to the government, and I am not happy with the idea that I may be acting in the interest of fanatic profit maximizers. Fortunately, people seldom ask me what I do. Perhaps I am a proud skeptic. Nevertheless, after many years in the profession, I still get excited when formal abstract models are successfully constructed and meaning emerges from the manipulation of symbols. It is moving when I observe that same excitement in students' faces. Thus, my greatest dilemma is between my attraction to economic theory, on the one hand, and my doubts about its relevance, on the other.

In this lecture I will try to decompose this basic dilemma into four parts:

The dilemma of absurd conclusions: Should we abandon a model if it produces absurd conclusions or should we regard it as a very limited set of assumptions that will inevitably fail in some contexts?

The dilemma of responding to reality: Should our models be judged according to experimental results?

The dilemma of modelless regularities: Should models provide the hypothesis for testing or are they simply exercises in logic that have no use in identifying regularities?

The dilemma of relevance: Do we have the right to offer advice or to make statements that are intended to influence the real world?

Many economists are aware of these dilemmas in one form or another. Nevertheless, I hope that bringing them together and linking them to recent research will have some impact.

2. THE DILEMMA OF ABSURD CONCLUSIONS

Formal models have a number of functions. Sometimes they are simply used as a tool to paint a clear picture of what we wish to express. As economic theorists, we use formal models to produce conclusions. Should we be as concerned with an absurd conclusion reached from sound assumptions as we would be by a contradiction in a mathematical model? Does an absurd conclusion require us to abandon an economic model?

Adam in the Garden of Eden

Consider Adam in the Garden of Eden who is taking a crash course[2] in life. He is endowed with a stream of apples that he can pick from the trees in the garden. Each period he chooses whether or not to pick the apples available that day; however, once he picks an apple, he has to eat it right away. In other words, he cannot store apples from one day to the next.

Adam was created rational and he is aware of the fact that a rational decision maker first has to identify what a final consequence is.[3] Adam adopts the standard economic view that a final consequence is a list of quantities of apples to be consumed each day. For example, the sequence that describes eating one apple on April 13th, 2071 is a final consequence (not only for the apple) that is independent of the day on which the decision is made to consume this sequence.

Assume that when Adam enters Eden, the following assumptions are satisfied:

(i) Adam possesses preferences $\succsim$ over the set of streams of apple consumption (sequences of nonnegative integers).

(ii) Given a consumption stream $c = (c_s)$ and a day t, his preferences $\succsim_{t,c}$ over the changes in his consumption from time t onward are derived from $\succsim$ (that is, for any two vector of integers Δ and Δ', interpreted as changes in apple consumption from period t onward, $\Delta \succsim_{t,c} \Delta'$ if and only if $(c_1, \dots, c_t + \Delta_1, c_t + \Delta_2, \dots) \succsim (c_1, \dots, c_t + \Delta'_1, c_t + \Delta'_2, \dots)$).

(iii) Adam likes to eat up to two apples a day and cannot bear to eat more than two apples a day.

(iv) Adam is impatient. In each period he would be delighted to increase his consumption right away from zero to one apples in exchange for two apples the next day and from one to two apples in exchange for one apple the next day. (This strong impatience assumption is not implausible even for individuals outside the Garden of Eden. In fact, one of the primary motivations of the hyperbolic discounting literature is the fact that there are people who prefer one apple today over two apples tomorrow and, at the same time, prefer two apples in 21 days to one in 20 days.)

(v) Adam does not expect to live for more than 120 years.

The First Traumatic Experience

Adam is endowed with a stream of one apple per day starting on day 18 for the rest of his life. We will now put Adam through his first traumatic experience

[2]The "course" follows Rubinstein (1998, 2001).

[3]This is an opportunity to say that I am more than a little confused about the meaning of this concept. (See Savage (1972, Sections 2.5 and 5.2).) Can there be a "final consequence" when it appears that most of us do in fact care about events after our death? Shouldn't the term "consequence" be interpreted as subjective, corresponding to what the decision maker considers "final" in a particular context?

in Eden. Adam proves a simple "calibration theorem" for his case: he should be willing to exchange his endowment for a single apple right away!

The proof can be understood from the following observation: Denote by $\langle a_1, \ldots, a_K \rangle$ the stream $(a_1, \ldots, a_K, 0, 0, \ldots)$. The stream of one apple per day for 2^1 days after a delay of 1 day, namely $\langle 0, 1, 1 \rangle$, is inferior to $\langle 0, 2, 0 \rangle$ and also to $\langle 1, 0, 0 \rangle$. Similarly, the stream of one apple per day for 2^2 days with a delay of 2 days, namely, the stream $\langle 0, 0, 1, 1, 1, 1 \rangle$ is inferior to $\langle 0, 1, 0, 1, 0, 0 \rangle$ and thus to $\langle 0, 1, 1, 0, 0, 0 \rangle$ and $\langle 1, 0, 0, 0, 0, 0 \rangle$. By induction we conclude that he must find the stream of 2^{17} days of one apple per day with a delay of 17 days inferior to receiving one apple right away. It is only left to calculate that in 120 years there are less than $2^{17} + 17$ days and we are done.

Thus, we have here a case in which a set of reasonable assumptions yields an absurdity. This is an alarming situation. If a basic model of decision making yields conclusions that are absurd, what is the validity of reasonable conclusions from models that use the decision making model as a building block?

The reader might notice a similarity between the above observation and an argument made in Rabin (2000) in the context of decision making under uncertainty.[4] When I initially added Rabin's argument to the material for my graduate microeconomics course, I added a sarcastic remark: "Do we economists take our own findings seriously?" Apparently, some economists like Rabin and Thaler (2001) have called for the replacement of expected utility with an alternative theory and are so sure of themselves that they feel "much like the customer in the pet shop, beating at a dead parrot." Let us follow this path and try to change the model so as to get rid of the absurd conclusion reached by Adam.

Recovering from the First Traumatic Experience

Let us return to Adam. Following his first traumatic experience (and following Strotz (1956)), Adam realizes that he should split his personality. He withdraws from the assumption that the consequences are independent of time. He now thinks of himself as a collection of egos, each with a different perspective. The consequences of an agent's choice at time t are streams of apples from time t onward. Thus, the meaning of eating one apple on day 27 will not necessarily be the same at $t = 0$ as at $t = 26$. It might be that at any time t he is ready to replace two apples at time $t + 1$ for one at time t, but not two apples on day $t + 27$ for one apple on day $t + 26$. Thus, Adam will be modeled as a

[4]Following is one of its versions: Consider a decision maker who behaves according to expected utility theory, is risk averse, and takes the final consequence to be the amount of money he will hold after all uncertainties have been resolved. Such a decision maker, who rejects the lottery $0.5[-10] \oplus 0.5[+11]$ at all levels of wealth in the interval [0, \$4,000], will reject an equal chance of losing a moderate amount like \$100 and making a large gain like \$64,000 when he holds the initial wealth of \$3,000.

sequence of preference relations ($\succsim_t$), one for each date, where each is defined on the streams of future consumption streams.

Note that this alteration of the model has an analogy in the context of decision making under uncertainty. Rabin's absurd conclusion was an outcome not only of expected utility theory assumptions, but also of the assumption that there is a single preference relation $\succsim$ over the set of lotteries, with prizes being the "final wealth levels" such that a decision maker at any wealth w who has a vNM (von Neumann–Morgenstern) preference relation $\succsim_w$ over the set of "wealth changes" derives that preference from $\succsim$ by $L_1 \succsim_w L_2$ if and only if $w + L_1 \succsim w + L_2$.[5] Kahneman and Tversky (1979) have already pointed out that this assumption clashes with unambiguous experimental evidence and, in particular, that there is a dramatic difference between our attitudes toward relative gains and relative losses. Withdrawing from the assumption that a consequence must be the final wealth level and allowing a consequence to be a wealth change avoids Rabin's absurd conclusion.[6] (See Cox and Sadiraj (2001) for an independent, though similar, argument.)

The Second Traumatic Experience

Once Adam has split into a collection of infinite agents, one for each point in time, he has his second traumatic experience. Assume that the first trauma changed his preferences and that he now has less appetite and does not eat more than one apple per day. He has lost his confidence and becomes an extreme example of a hyperbolic discounter who cares only about what happens in the next two days. On the other hand, whenever he compares eating an apple today to eating an apple tomorrow, he prefers to delay the pleasure.

By now, Adam has found Eve. Eve offers Adam one apple. When he is about to eat the apple, she says to him, "Why don't you give me the apple and get an additional one tomorrow?"At this point Adam still does not realize that he might have a conflict between his selves. He is still naive. Each of his selves takes actions as if the others do not exist. Naive Adam will take the bait and never eat the apple. How sad.

[5]Note that nothing in the vNM axioms dictates that consequences should be the final wealth levels rather than wealth changes. When discussing vNM theory, standard textbooks are indeed vague on the interpretation of w. They usually state that the decision maker derives utility from "money," with no discussion of whether "money" is a flow or a final stock.

[6]It allows us to make the plausible assumption that for a wide range of moderate wealth levels w, a decision maker rejects the lottery $0.5[-10] \oplus 0.5[+11]$ (probably applying an instinctual aversion to risk), and were he to start from wealth 0, for example, he would prefer the lottery $0.5[w - 10] \oplus 0.5[w + 11]$ over the sure amount $[w]$ (probably applying an argument that when all prizes are similar, he considers expected gains).

Recovering from the Second Traumatic Experience

Frustrated by Eve, Adam goes to the snake, a successful consultant who has graduated from a course in game theory. The snake tells Adam that he must be more sophisticated about the interaction between his various selves. He explains to Adam that the common assumption made in economics is that the decision maker's behavior must be consistent with a "perfect equilibrium procedure" ("sophisticated behavior" as it is called in the behavioral economics literature). The snake shows Adam that there are only two perfect equilibria for the game between his selves and according to them he should eat the apple on the first or second day. Adam feels relieved.

The Third Traumatic Experience

The snake has already won Adam's trust, but now Adam goes through a third traumatic experience. Adam is told that he can pick one apple every day. What could be simpler than that? Adam plans to pick an apple every day. However, the snake has different advice for Adam. He recommends a "perfect equilibrium": Adam should pick an apple only after an odd number of consecutive days during which he has not done so.

Adam is impressed by the snake's originality but nevertheless verifies that there is no hypothetical history after which one of Adam's selves can find a reason not to follow the snake's advice:

(i) Consider a self after a history in which he is not supposed to pick an apple, that is, after an even number of days during which he did not eat any apples. The self expects to eat an apple a day later. This is better than the alternative in which he does not eat the apple and, according to the equilibrium, neither will the next self (because he will be acting after zero days during which Adam has not eaten any apples).

(ii) Consider a self after a history in which he is supposed to eat an apple, that is, after an odd number of days during which he did not eat apples. According to the equilibrium, the self expects that the next self will not eat an apple. This is better than the alternative in which the self does not eat the apple and, according to the equilibrium, neither does the next self (because he will be acting after an even number of days during which Adam has not eaten any apples).

To conclude, Adam does not find any problem with the snake's advice and eats apples only once every two days.

The Dilemma of Absurd Conclusions

We have now arrived at the dilemma. We want assumptions to be realistic and to yield only sensible results. Thus, nonsensical conclusions will lead us to reject a model. However, unlike parrots, human beings have the ability to

invent new ways to reason that will clash with any theory. Attempting to escape from the calibration theorem, Adam ran into Eve. Escaping from Eve, he ran into the snake. If we followed the behavioral economics methodology of rejecting a theory if it reaches an absurd conclusion, we would trash expected utility and constant discounting, but then would reject the alternative theories as well. I doubt there is any set of assumptions that does not produce absurd conclusions when applied to circumstances far removed from the context in which they were conceived. So how should we respond to absurd conclusions derived from sensible assumptions?

3. THE DILEMMA OF RESPONSE TO EVIDENCE

The connection between the models in economic theory and reality is tricky. I do not think that many of us take our models seriously enough to view them as platforms for producing accurate predictions in the same way that models in the sciences are viewed. When comparing a model to real data, we hope at best to find some evidence that "something" in reality is close to the model's prediction. Experiments are used to verify assumptions and conclusions. Should we change a model if one of its assumptions is experimentally refuted? Let us consider, for example, the evaluation of assumptions regarding time preferences.

The Case for Hyperbolic Preferences

Recently there has been a trend in "behavioral economics" to replace the traditional discounting formula with a variation of the hyperbolic discounting formula whereby, for each day, the payoffs from that point on are discounted by $1, \beta\delta, \beta\delta^2, \beta\delta^3, \ldots$. This trend has gained popularity despite the problem (mentioned in the previous section) that it involves much more than just changing the scope of the preferences—it introduces time inconsistencies and requires assumptions about the interaction between the different selves.

The hyperbolic discounting literature (see, for example, Laibson (1996)) is based on unequivocal statements such as "Studies of animal and human behavior suggest that discount functions are approximately hyperbolic." Indeed we have reliable evidence (especially because it is confirmed by our own thought experiments) that, for certain decision problems, stationary discounting is inconsistent with the experimental results and that hyperbolic discounting preferences fit the data better. For example, there are more people who prefer an apple today over two apples tomorrow than there are those who prefer two apples in 21 days over one apple in 20 days. So we adopt hyperbolic discounting or, to be more precise, a simple version of this approach characterized by two parameters, β and δ.

The Case Against Hyperbolic Preferences

What if we can easily design experiments that reject the alternative theory as well? Following are the results of an experiment I conducted in 2003 on the audiences of a lecture delivered at the University of British Columbia. Students and faculty were asked to respond on-line to the following problem:

PROBLEM 1: Imagine you have finished a job and have to choose between two payment schemes:
(A) Receiving \$1,000 in 8 months.
(B) Receiving \$500 in 6 months and \$500 in 10 months.
Which scheme would you choose?

Receiving \$1,000 in 8 months is not much different from receiving \$500 at $8-\varepsilon$ and \$500 at $8+\varepsilon$. Thus, a reasonable application of the (hyperbolic) discounting approach in this case would imply that advancing the receipt of \$500 from $t=8$ to $t=6$ has more weight than postponing the receipt of \$500 from $t=8$ to $t=10$. Therefore, we would expect the vast majority of people to choose B. However, 54% of the 354 participants in this experiment chose A.

I believe that the phenomenon we see here is somewhat related to risk aversion: Given two alternatives, there is a strong tendency to choose the one perceived as the "average." In the context of decision making under uncertainty, people tend to prefer the certain expectation of a lottery over the lottery itself. In the context of streams of money, the averaging might be done on the time component. This consideration leads an individual to prefer one installment. Apparently, for a majority of subjects the preference for the average is stronger than the consideration underlying hyperbolic discounting (advancing the receipt of \$500 by two periods is a more significant than the loss from postponing the receipt of the same amount for two periods) which, of course, I do not deny exists.

If I am right, then one would expect, following Kahneman and Tversky (1979), that the subjects' choices in the dual problem, which involve losses rather than gains, would be reversed. To strengthen the experimental evidence against hyperbolic discounting, I tested this as well. Students and faculty invited to a lecture at Georgetown University were asked to respond on-line to the following problem:

PROBLEM 2: Imagine you have bought a computer and have to choose between two payment schemes:
(A) Paying \$1,000 in 8 months.
(B) Paying \$500 in 6 months and \$500 in 10 months.
Which scheme would you choose?

While 54% of the subjects chose one installment when they had to choose between payment schemes for earnings in Problem 1, only 39% of the 382 par-

ticipants chose one installment in Problem 2, when they had to choose between payment schemes for losses.

The Dilemma of Response to Evidence

The results of both the experiments are the opposite of what is predicted by the hyperbolic discounting approach. So should we dismiss the hyperbolic discounting model? According to the methodological guidelines implicitly followed by many behavioral economists, the answer is yes.

Of course, there is a tempting alternative—simply to dismiss evidence we do not like. I know personally of one paper (Rubinstein (2003)) that presented the results of several experiments aimed at refuting the hyperbolic discounting theory. An editor of a very prestigious journal,[7] which has published many of the hyperbolic discounting papers, justified his decision to reject the paper as follows: "Ultimately this seems like a critique of the current approach which is right in many ways, but criticisms and extensions of existing research are best sent to more specialized outlets."

Taking a more serious approach, we are faced here with the dilemma of how to respond to experimental evidence. We want our assumptions to reflect reality, but you can put together any combination of reasonable assumptions and be certain that someone will find an experiment to defeat your theory. So how can we proceed given the fact that rejecting assumptions using experimental results is so easy?

4. THE DILEMMA OF MODELLESS REGULARITIES

Models in economic theory are also used to suggest regularities in human behavior and interaction. By regularities I mean phenomena that appear repeatedly in similar environments at different points in time and at different locations. I have the impression that as economic theorists, we hope that regularities will miraculously emerge from the formulas we write leisurely at our desks. Applied economists often feel the need for a model before they mine data for a pattern or regularity. Do we really need economic theory to find these regularities? Would it not be better to go in the opposite direction by observing the real world, whether through empirical or experimental data, to find unexpected regularities? Personally I doubt that we need pre conceived theories to find regularities.

The Traveler's Dilemma

To illustrate the point, let us have a look at a version of the Traveler's Dilemma (owing to Basu (1994)):

Imagine you are one of the players in the following two-player game:

[7]To prevent any misunderstanding, it was the *Quarterly Journal of Economics*.

- Each of the players chooses an amount between \$180 and \$300.
- Both players are paid the lower of the two chosen amounts.
- Five dollars are transferred from the player who chose the larger amount to the player who chose the smaller one.
- In the case that both players choose the same amount, they both receive that amount and no transfer is made.

What is your choice?

The standard game theoretic analysis assumes that the players care only about their final dollar payoff. Since the only Nash equilibrium for the game is for both players to choose 180, the standard application of game theory would explain a regularity in which all players choose 180.

A Regularity Is Found

During the years 2002–2003, I was able to collect large amounts of data from audiences of a public lecture that I delivered at several universities.[8] People who were invited to attend the lecture, most of them students and faculty, were asked to respond to several questions before the lecture on the website *gametheory.tau.ac.il*. One of the questions was the above version of the Traveler's Dilemma.

Figure 1 shows the results for nine universities in six countries: Ben-Gurion University, Tel Aviv University, the Technion (Israel); Tilburg University (Holland); the London School of Economics (United Kingdom); the University of British Columbia and York University (Canada); Georgetown University (United States); and Sabanci (Turkey). The five graphs look quite similar and reveal a regularity in the distributions of something like the following:

180	181–294	295	296–298	299	300
13%	15%	5%	3%	9%	56%

Note that this regularity was found without any preconceived model and I am not aware of any existing game theoretical model that can, in fact, explain it.

Further Insights Are Found

Finding an explanation of the regularity in the distributions of responses in a case like the virtual Traveler's Dilemma is likely to involve the search for a recurring distribution of more fundamental psychological traits. For that we need to have a better psychological understanding of the meaning of each of the responses, rather than a fancy model.

[8]In the lecture, titled "John Nash, Beautiful Mind and Game Theory," I critically introduced the basic ideas of game theory, spoke about my personal encounter with John Nash, and discussed a bit about the book and the movie.

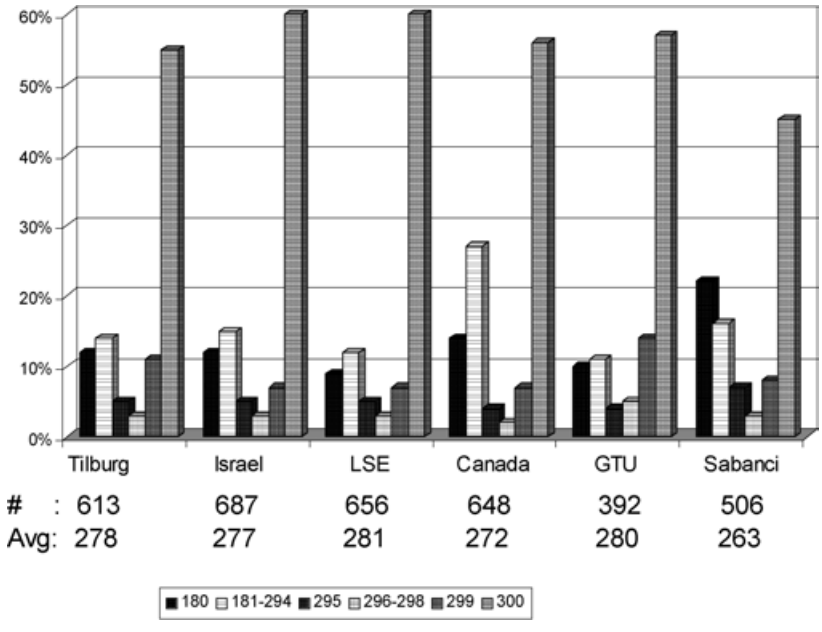

FIGURE 1.

The players who chose 180 are probably aware of the game theoretical prediction. On average, they would do badly playing against a player chosen randomly from the respondents. These players can claim to be the "victims" of game theory. The subjects whose answers were in the range 295–299 clearly exhibit strategic reasoning. The answer 300 seems to be an instinctive response in this context and the responses in the range 181–294 appear to be the result of random choice.

To support this interpretation, I gathered data on the subjects' response times (see Rubinstein (2004)). Response time is a very noisy variable due to differences in server speeds, differences in cognitive abilities among subjects, etc. Nevertheless, when the sample is large enough, as this one was, we can obtain a reliable picture (which is confirmed by the fact that the relationship between the distributions is similar at all the various locations). Table I shows the median response times in seconds and Figure 2 shows the cumulative distributions of the response times among 2,985 subjects[9] for the four ranges {180}, {181, . . . , 294}, {295, . . . , 299}, and {300}.

Remarkably, the response 300 and the responses in the range 181–294 are the quickest. Apparently 300 is indeed the instinctive response and responses in the range 181–294 are the result of "random" choice without a clear rationale. The responses in the range 295–299, which imply greater cognitive efforts,

[9]Time response was not recorded for the first two audiences to whom the lecture was delivered.

TABLE I

$n = 2{,}985$	%	Median Response Time
180	13%	87 s
181–294	14%	70 s
295–299	17%	96 s
300	55%	72 s

indeed take the most time. The "victims" of game theory who chose 180 are somewhere in between. The shape of their distribution seems to indicate that some of the subjects calculated the equilibrium (a cognitive operation) and that some of them were already familiar with the game.

The time response data add meaning to the results. Choices associated with a long time response are likely to be the outcome of a more intensive use of a cognitive process whereas a more instinctive process might be responsible for short time response. The distinction between fast intuitive operations and slow cognitive operations is related to the psychologists' distinction between systems 1 and 2 (see, for example, Stanovich and West (2000) and Kahneman (2003)). However, note that we had no model in mind before looking at the data and we are still a long way from explaining the stable distribution of responses across different populations.

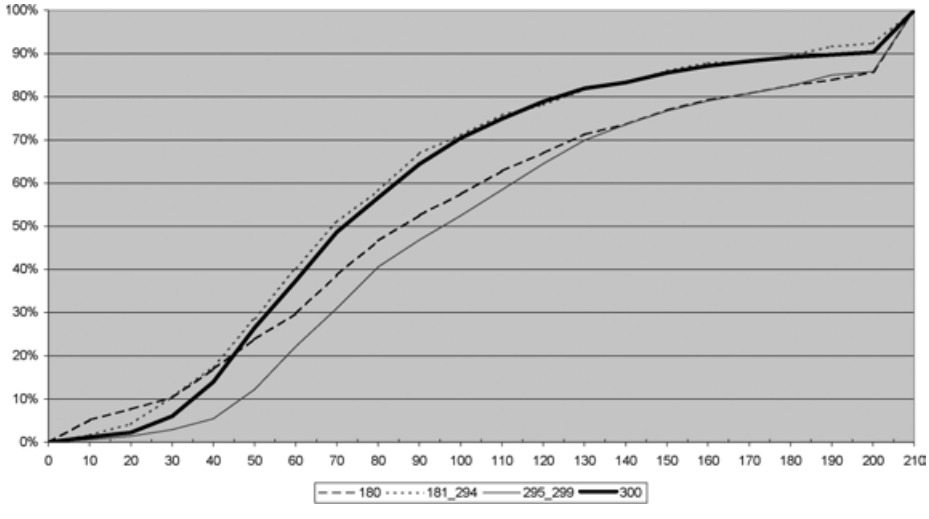

FIGURE 2.

The Dilemma of Modelless Regularities

We have now arrived at the dilemma of modelless regularities. We would like a model to produce interesting conclusions that are consistent with observed regularities so we can claim that the model provides an explanation of those regularities, but are complicated theoretical models really necessary to find interesting regularities?

5. THE DILEMMA OF RELEVANCE

It is true that I would like to change the world. I want people to listen to me, but as an economic theorist, do I have anything to say to them?

One of my earliest interests as an economic theorist was in bargaining theory. There were two reasons for this: First and foremost, bargaining theory involves the construction of models that are simple but nevertheless rich in results that have attractive interpretations. Indeed, the possibility of deriving meaningful statements through the manipulation of mathematical symbols was something that attracted me to economics in the first place. Second, as a child I frequented the open air markets in West Jerusalem and later the Bazaar in the Old City of Jerusalem, and as a result, bargaining had an exotic appeal for me. I came to prefer bargaining theory over auction theory, because auctions were associated with the rich whereas bargaining was associated with the common people. However, I never imagined that bargaining theory would make me a better bargainer. When people approached me later in life for advice in negotiating the purchase of an apartment or to join a team planning strategy for political negotiations, I declined. I told them that as an economic theorist I had nothing to contribute. I did not say that I lacked commonsense or life experience that might be useful in such negotiations, but rather that my professional knowledge was of no use in these matters. This response was sufficient to deter them. Decision makers are usually looking for professional advice, rather than advice based on commonsense. They believe, and perhaps rightly so, that they have at least as much commonsense as assertive professional economists.

Nevertheless, I am a teacher of microeconomics. I am a part of the "machine" that I suspect is influencing students to think in a way that I do not particularly like.

The Layoff Survey

In 2004, I conducted a survey among six groups of Israeli students. The students were told that the questionnaire was not an exam and that there were no "right" answers. The core of the questionnaire was as follows:

Q-TABLE (Translated from Hebrew): Assume that you are a vice president of ILJK company. The company provides extermination services and employs a certain number of permanent administrative workers and 196 nonpermanent

TABLE II

Number of Workers Who Will Continue to Be Employed	Expected Annual Profit (Millions of Shekels)
0 (All the workers to be laid off)	Loss of 8
50 (146 workers to be laid off)	Profit of 1
65 (131 workers to be laid off)	Profit of 1.5
100 (96 workers to be laid off)	Profit of 2
144 (52 workers to be laid off)	Profit of 1.6
170 (26 workers to be laid off)	Profit of 1
196 (No layoffs)	Profit of 0.4

workers who are sent out on extermination jobs. The company was founded five years ago and is owned by three families. The work requires only a low level of skill, with each worker requiring only one week of training. All the company's employees have been with the company for three to five years. The company pays its workers more than minimum wage. A worker's salary includes payment for overtime, which varies from 4,000 to 5,000 shekels per month.[10] The company makes sure to provide its employees with all the benefits required by law. Until recently, the company was making large profits. As a result of the continuing recession, there has been a significant drop in profits, although the company is still in the black. You will be attending a meeting of the management in which a decision will be made regarding the layoff of some of the workers. ILJK's finance department has prepared scenarios of annual profits shown in the table (Table II). Complete the following:

I recommend continuing to employ _____ of the 196 workers presently employed by the company.

The full results of this experiment appear in Rubinstein (2006).[11] Six groups of students were approached by e-mail and asked to respond to a series of questions via the Internet. The groups comprised undergraduate students in the departments of economics, law, mathematics, and philosophy at Tel Aviv University; MBA students at Tel Aviv University; and undergraduates in economics at the Hebrew University of Jerusalem. I will refer to the six groups using the abbreviations EconTAU, Law, Math, Phil, MBA, and EconHU.

Table III presents the responses of 764 students (who answered 100 or more[12]) to the Q-Table. The differences between the groups are striking. The Econ students both at the Hebrew University and Tel Aviv University are much more pronounced profit maximizers than the students in the other groups.

[10]The minimum wage in Israel was about 3,300 shekels at the time of the experiment.

[11]In addition to a more complete presentation of the results, Rubinstein (2006) also reports the results of the survey among several thousand readers of an Israeli daily newspaper and among Ph.D. students at Harvard.

[12]For a discussion of the 5% who chose a number less than 100, see Rubinstein (2006).

TABLE III

	EconHu	EconTAU	MBA	Law	Math	Phil
Q-Table	$n = 94$	$n = 130$	$n = 172$	$n = 216$	$n = 64$	$n = 88$
100	**49%**	**45%**	**33%**	**27%**	**16%**	**13%**
144	**33%**	**31%**	**29%**	**36%**	**36%**	**19%**
170	**7%**	**9%**	**23%**	**18%**	**25%**	**25%**
196	**6%**	**13%**	**12%**	**13%**	**11%**	**36%**
Other	4%	2%	3%	6%	13%	7%
Average	127	133	142	144	151	165

Almost half of the Econ students chose the profit-maximizing alternative, as compared to only 13–16% of the Phil and Math students. The MBA and Law students are somewhere in between. The response of "no layoffs" was given by only a small population of respondents (6–15%) in five of the six groups. The philosophers were the only exception: 36% of them chose to ignore the profit-maximizing target. A major surprise (at least for me) was the fact that the MBA students responded differently than the Econ students. I think that this has to do with the way in which MBA programs are taught. Perhaps the study of cases triggers more comprehensive thinking about real-life problems than the study of formal models, which conceals the need to balance between conflicting considerations.

A variant of the problem, Q-Formula, was identical to the Q-Table except that the table was replaced with the following statement: "The employment of x workers will result in annual profits (in millions of shekels) equal to $2\sqrt{x} - 0.1x - 8$." Note that this profit function yields similar values to those presented in the table and has an identical maximum at $x = 100$.

In the Law and Phil groups, all subjects received the Q-Table version. Subjects in the other four groups, who have more mathematical background, were randomly given either the Q-Table or the Q-Formula.

A total of 298 subjects responded to the Q-Formula. Here there were no major differences between the four groups. A vast majority (around 75%) of subjects in all groups maximized profits, although many of them were aware of the existence of a trade-off (as is evident from the fact that many of those who chose 100 revealed in a subsequent question that they believe that a real vice president would fire a smaller number of workers than that required to maximize profits). Thus, presenting the problem formally, as we do in economics, seems to obscure the real-life complexity of the situation for most students (including math students).

The interpretation of the results cannot be separated from one's personal views regarding the behavior of economic agents in such a situation. If you believe that the managers of a company are obligated morally or legally to maximize profits, then you should probably praise economics for how well it

indoctrinates its students and be disappointed that so many of them still do not maximize profits. On the other hand, if you approach the results with the belief that managers should also take into account the welfare of the workers, particularly when the economy is in recession and unemployment is high, then you probably feel uncomfortable with the results.

Of course, it is possible that the differences between the two groups of economics undergraduates and the other groups is due to selection bias rather than indoctrination. However, the fact that the responses of the economists differed from those of the lawyers and MBA students, and not just from those of the philosophers and mathematicians, makes this possibility less likely. The uniformity in the responses to the Q-Formula appears to provide support for the indoctrination hypothesis also.

Perhaps there is no connection between the responses and the choices that would be made in real life. However, if there is no connection, does not that mean that what a student learns in economics will have no influence on his behavior and we should be revising our curriculum? Overall, I am left with the impression that in the best case, the formal exercises we assign to our students make the study of economics less interesting; in the worst case, they contribute to shaping a rather unpleasant "economic man."

The Jungle Model

Guilt feelings probably motivated me in Piccione and Rubinstein (2003). This is the only paper I have ever been involved with that was motivated by real-life problems.

We constructed a model that we called The Jungle. Whereas in an exchange economy, transactions are made with the mutual consent of two parties, in the jungle it is sufficient that one agent, who happens to be the stronger of the two, is interested in the transaction. The model is meant to be similar to the exchange economy model with the exception that there is no ownership and agents do not come into the model with an initial endowment. Formally, the vector of initial endowments is replaced with a power relationship in this model.

After spelling out the model and the definition of a jungle equilibrium, examples are brought to illustrate the richness of the model. Several propositions are proved: existence, uniqueness, and the First Fundamental Welfare Theorem (under some smoothness assumptions, the jungle equilibrium is efficient). Finally, an analogy to the Second Fundamental Welfare Theorem is discussed and it is shown that every jungle equilibrium allocation is also supported by equilibrium prices such that the stronger are also the richer. One could interpret this statement to mean that power and wealth go hand-in-hand.

When I present this model in public lectures, I ask the audience to imagine that they are attending the first lecture of a course at the University of the Jungle designed to introduce the principles of economics, and to show how the

visible iron hand produces order out of chaos and results in the efficient allocation of available resources without the interference of a government. We argued in the paper that the greed on which the market economy is based is analogous to the strength to take advantage of the weak in the jungle economy. The market economy encourages people to produce more, thus increasing society's resources, whereas the jungle economy encourages people to develop their strength, thus facilitating society's expansionist ambitions.

I view the jungle model as a rhetorical exercise designed to sow (more) doubts for economics students in their study of models of competitive markets. The idea was to build a model that is as close as possible to the standard exchange economy, using terminology that is familiar to any economics student, and to conduct the same type of analysis found in any microeconomics textbook on competitive equilibrium. A standard economics course impresses students with its elegance and clarity. We tried to create a model of the jungle that does the same.

The Dilemma of Relevance

This brings me to the fourth dilemma. I believe that as an economic theorist, I have very little to say about the real world and that there are very few models in economic theory that can be used to provide serious advice. However, economic theory has real effects. I cannot ignore the fact that our work as teachers and researchers influences students' minds and does so in a way with which I am not comfortable. Can we find a way to be relevant without being charlatans?

6. CONCLUDING WORDS

It is time to sum up. How do I relate to these four dilemmas?

As economic theorists, we organize our thoughts using what we call models. The word "model" sounds more scientific than "fable" or "fairy tale" although I do not see much difference between them. The author of a fable draws a parallel to a situation in real life. He 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 in 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. 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 that have been stripped of most of their real-life characteristics. However, in a good model, as in a good fable, something significant remains.

Like us, the teller of fables confronts the dilemma of absurd conclusions, because the logic of his story may also lead to absurd conclusions.

Like us, the teller of fables confronts the dilemma of response to evidence. He wants to maintain a connection between his fable and what he observes; there is a fine line between an amusing fantasy and a fable with a message.

Like us, the teller of fables is frustrated by the dilemma of fableless regularity when he realizes that sometimes his fables are not needed to obtain insightful observations.

Like us, the teller of fables confronts the dilemma of relevance. He wants to influence the world, but knows that his fable is only a theoretical argument.

As in the case of fables, absurd conclusions reveal contexts in which the model produces unreasonable results, but this may not necessarily make the model uninteresting.

As in the case of fables, models in economic theory are derived from observations of the real world, but are not meant to be testable.

As in the case of fables, models have limited scope.

As in the case of a good fable, a good model can have an enormous influence on the real world, not by providing advice or by predicting the future, but rather by influencing culture.[13]

Yes, I do think we are simply the tellers of fables, but is that not wonderful?

School of Economics, Tel Aviv University, Tel Aviv 69978, Israel; and Dept. of Economics, New York University, New York, NY 10003, U.S.A.; rariel@post.tau.ac.il.

Manuscript received June, 2005; final revision received March, 2006.

REFERENCES

BASU, K. (1994): "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory," *American Economic Review*, 84, 391–395.[873]

COX, J. C., AND V. SADIRAJ (2001): "Risk Aversion and Expected-Utility Theory: Coherence for Small- and Large-Stakes Gambles," *Games and Economic Behavior*, forthcoming.[869]

KAHNEMAN, D. (2003): "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 93, 1449–1475.[876]

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–292.[869,872]

LAIBSON, D. (1996): "Hyperbolic Discount Functions, Undersaving, and Savings Plans," Working Paper 5635, NBER.[871]

PICCIONE, M., AND A. RUBINSTEIN (2003): "Equilibrium in the Jungle," Mimeo.[880]

RABIN, M. (2000): "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281–1292.[868]

RABIN, M., AND R. THALER (2001): "Anomalies: Risk Aversion," *Journal of Economic Perspectives*, 15, 219–232.[868]

[13]I use the term "culture" in the sense of an accepted collection of ideas and conventions that influence the way people think and behave.

RUBINSTEIN, A. (1998): *Modeling Bounded Rationality*. Cambridge, MA: MIT Press.[867]

——— (2001): "Comments on the Risk and Time Preferences in Economics," Mimeo.[867]

——— (2003): "Economics and Psychology? The Case of Hyperbolic Discounting," *International Economic Review*, 44, 1207–1216.[873]

——— (2004): "Instinctive and Cognitive Reasoning: A Study of Response Times," Mimeo.[875]

——— (2006): "A Skeptic's Comment on the Studies of Economics," *Economic Journal*, 116, C1–C9.[878]

SAVAGE, L. J. (1972): *The Foundations of Statistics* (Second Ed.). New York: Dover. [867]

STANOVICH, K. E., AND R. F. WEST (2000): "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences*, 23, 645–665.[876]

STROTZ, R. H. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165–180.[868]