Dilemmas of An Economic Theorist

Presidential Address

Econometric Society, 2004

Ariel Rubinstein

Tel Aviv University

and

Department of Economics, New York University
1. An economic theorist’s motivation

I am going to talk briefly about several pieces of research which I have been involved in during the last few years. However, they are only a means to an end. I would like to use them to illustrate three dilemmas which I have encountered in my work as an economic theorist. They are:

The dilemma of absurd conclusions: Should we abandon a model if it produces absurd conclusions or should we regard simply a model as a very limited set of assumptions which will inevitably fail in some contexts?
The dilemma of responding to reality: Should our models be judged according to experimental results, should they provide the hypothesis for testing or are they simply exercises in logic in which case regularities can be found without them?
The dilemma of relevance: Do we have the right to give advice or to make statements which are meant to have an influence in the real world?

Lurking in the background is one big question which I ask myself obsessively: What are we trying to accomplish as economic theorists? In some sense, we essentially play with toys called “models”. We have the luxury of remaining children for our whole professional lives and are even well paid for it. We get to call ourselves economists and the public naively thinks we are improving the economy, increasing the rate of growth or preventing economic catastrophe. Of course, we can justify our public image by repeating some fancy sounding slogans (which we often do) we believe in them?

I recall a conference I attended in Lumini, France in the summer of 1981 which 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, junior economists were standing off to the side eavesdropping on their conversation. They loudly discussed the relevance of game theory and one of them suggested that they were “making a living”. 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 which only appears to others to be useful?

I did not fulfill any childhood fantasy by becoming a professor. It was never my dream to become an economist. I respect philosophers, teachers, writers and nurses not less 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 advice to government bodies and I am not happy with the idea that I may be acting in the service of fanatic profit maximizers. Fortunately, people seldom ask me what I do. I was once asked for advice about real estate. My very honest answer - that I didn’t have the slightest idea about real estate - was viewed as arrogant.

Perhaps, I am a proud skeptic. However, 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 three parts.

2. The Dilemma of Absurd Conclusions
Consider lonely Adam in the garden of Eden who is taking a crash course in what life is all about. He is endowed with the right to pick a certain stream of apples from the trees. Assume that he cannot store apples from one day to the next. To gain experience, he is given options for exchanging his endowment for other streams of apples. In the case that he does not have the right to pick an apple today, he is ready to surrender the right to pick two apples tomorrow in order to get one today. This assumption is not implausible even for individuals outside the garden of Eden. 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 31 days to one in 30 days.
A rational choice analysis of this situation requires a definition of what a “final consequence” is. I am more than a little confused about the meaning of this concept. 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 any case, we usually adopt the view that a “final consequence” in such a model is a list of days and the quantities of apples to be consumed on each of those days. Thus, for example, eating one apple on April 13th 2071 is a final consequence independent of the day on which I made the decision to eat it.

We now look at the first traumatic event experienced by Adam in Eden. The following is a calibration “theorem” for “time preferences”:

**Proposition:** Assume that Adam:

- Has a preference relation $\succ$ over the set of streams of apple consumption.
- Likes to eat up to 2 apples a day and cannot bear to eat 3 apples a day.
- Is time impatient. He will be delighted to increase his consumption at day $t$ from 0 to 1 or from 1 to 2 apples at the expense of an apple he is promised at a day later than $t$.
- At every day $t$, and given that he has a consumption stream $c$, he holds preferences $\succ_{tc}$ over changes in apple consumption in days following $t$ satisfying $D_1 \succ_{tc} D_2$ iff $c + D_1 \succ c + D_2$.
- Whenever he lacks an apple he prefers to get it right away in exchange for two apples tomorrow. That is, if $c_t = 0$, he is ready to exchange 2 apples from day $t + 1$ for 1 apple at day $t$.

Then, if (poor) Adam owns a stream of 2 apples starting in day 18 for the rest of his life (assuming he does not expect to live more than 120 years), he will be willing to exchange it for 2 apples right away!

The essence of the proof can be seen from the following observation: The stream of two apples per day for $2^1$ days after a delay of $2^0$ day, namely $(0, 2, 2)$, is inferior to $(1, 0, 2)$, and also to $(1, 1, 0)$ and, by impatience, also to $(2, 0, 0)$. Similarly, the stream of two apples per day for $2^2$ days with a delay of $2^1$ days, namely, the stream $(0, 0, 2, 2, 2, 2)$, is inferior to $(0, 2, 0, 2, 0, 0)$ and therefore to $(1, 0, 1, 0, 0, 0)$ and thus to $(2, 0, 0, 0, 0, 0)$. By induction we conclude that he must find the stream of $2^{17}$ days of two apples per day...
with a delay of \(17\) days inferior to receiving 2 apples right away. It is only left to calculate that in 120 years there are no more than \(2^{17}\) days and we are done.

You might have noticed a similarity between the above observation and an argument due to Matt Rabin (see Rabin (2001)). Rabin considers 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 are resolved. Such a decision maker, who rejects, at all levels of wealth (or at a certain interval, such as \([0, 4000]\)), the lottery \(0.5[-10] \oplus 0.5[+11]\), will reject an equal chance to lose a moderate amount like \$100 and to make a large gain like \$64000 when he holds the initial wealth of \$3000. The basic idea is as follows: denoting the vNM utility function by \(u\) we obtain

\[
u(w + 11) - u(w) < u(w) - u(w - 10)\text{ for all } w \in [0, 4000]\text{ and thus the marginal utility function } \mu u(w) \text{ satisfies}
\]

\[
\mu(u(w + 11) - u(w)) / 11 < (10/11)(u(w) - u(w - 10)) / 10 \leq (10/11)\mu(u(w - 10)) \text{ in that domain. In other words, it falls at a faster rate than that of a geometrical sequence.}
\]

How should we respond to such observations? When I initially added Rabin’s argument to the material for my graduate micro-economics course, I added sarcastically: “What conclusion should we derive from this observation? Do we economists take our own findings seriously?” My first instinct was that something is deeply wrong with the model we so commonly use. I felt that the situation is similar to the case in which a set of assumptions yields a contradiction and thus any conclusion can follow. If our model of decision making with time preferences or under uncertainty yields conclusions which are absurd, what is the validity of conclusions which are not?

So how do we proceed? Adam learns (following Strotz) that he should split his personality. He should think about himself as a collection of egos each with a different time point of view. The consequences of an agent’s choice at time \(t\) will be all streams of apples from time \(t\) on. However, the meaning of eating an apple at day 17 will not necessarily be the same at \(t = 0\) as at \(t = 16\). Indeed, when discussing time preference in the world of pairs \((x, t)\) (interpreted as getting \(x\) apples on day \(t\) and assuming \(c = 0\)), it is not uncommon to attribute a sequence of preference relations \((\succeq_t)\) to a single decision maker and not to assume that \((+1\text{apple, right away}) \succeq_t (+2\text{apples, in one day delay})\) iff \((1\text{apple, t}) \succeq (2\text{apples, t + 1})\) for some \(\succeq\).

The same alteration has to be made in the context of decision making under uncertainty.
The absurd conclusion reached by Rabin was not only an outcome of expected utility theory assumptions but also of the assumption that there is a single preference relation \( \succeq \) 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 preference relation \( \succeq_w \) over the set of “wealth changes” derives that preference from \( \succeq \) by \( L_1 \succeq_w L_2 \iff w + L_1 \succeq w + L_2 \). Kahneman and Tversky (1979) have already pointed out that this assumption clashes with clear cut experimental evidence and in particular that there is a dramatic difference between our attitudes towards relative gains and our attitude towards relative losses. In the context of Rabin’s example it might be plausible that for a wide range of moderate wealth levels \( w \) a decision maker rejects the lottery \( 0.5[-10] \oplus 0.5[+11] \) (as he holds 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 \( [w] \) (using an argument of the type that all possibilities are similar and thus I will decide simply by calculating the expectation). 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 actually vague on the interpretation of “\( w \)” - usually they state that the decision maker derives utility from “money”, with no discussion of whether “money” is a flow or a final stock.

Once he was split into a collection of infinite agents, one for each point in time, naive Adam has his second experience (based on Rubinstein (1998)): His first trauma changed his preferences. He is now on a diet and does not eat more than one apple per day. He has also become an extreme example of a hyperbolic discounter who cares only about what happens in the next two days but whenever he compares eating an apple today to eating an apple tomorrow, he prefers the former.

Adam now has his second experience. He finds Eve, who is a very tempting lady. Eve offers Adam one apple. When he is about to eat the apple she tells him, “Why don’t you give me the apple and get another 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 did not exist. Naive Adam will take the bait and never eat the apple. Sad.

Frustrated by Eve, Adam goes to Mr. Snake, a successful consultant who has graduated from a course in game theory. The snake tells Adam that he must be more sophisticated regarding 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 as a “sophisticated” decision maker he should eat the apple on the first or second day. Adam feels relieved.

The snake has already won Adam’s trust, but now Adam has his third experience. Adam is told that he can pick one free apple every day. What could be simpler than that? Adam plans to pick an apple every day. However, the snake has other plans for Adam. He recommends a “perfect equilibrium” to Adam: Adam should pick an apple only after an odd number of consecutive days in which he has not done so. Adam is impressed by the snake’s originality. Recall that Adam has been split into a sequence of agents/selves. None of Adam’s selves can find a reason not to follow the snake’s advice. A self attached to a day in which he is not supposed to pick an apple, expects to get an apple a day later which is better for him than eating the apple now and not eating it (as the equilibrium suggests) on the second day (since this will follow 0 days of not eating apples). On a day he is supposed to eat the apple, not eating the apple will result in him not eating the apple on the following day as well (as it will have been an odd number of days since he ate an apple) and this is worse for Adam than eating the apple right away. Adam does not find any problem with the snake’s advice and eats apples only once every two days.

What conclusion should we draw from the calibration theorem and the snake’s trick? Rabin and Thaler, referring to Rabin’s calibration theorem, attack expected utility theory. They say they feel “much like the customer in the pet shop, beating at a dead parrot”. But, if we were following the methodology of at least some behavioral economists, we would trash not only expected utility and constant discounting preferences but hyperbolic discounting as well.

We have now arrived at the **Dilemma of Absurd Conclusions**. 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 of reasoning that will confound any theory. I doubt if there is any set of assumptions which will not produce absurd conclusions when we apply them to
circumstances which are far removed from the context they were originally intended for. So, how should we respond to absurd conclusions derived from sensible assumptions?

3. The Dilemma of Response to Reality
The connection between our models and reality is tricky. I don’t think we take our theoretical models seriously enough as to consider them to be a platform for producing verified predictions in the same way the sciences do. When comparing a model to real data, we hope at best to find some evidence that “something” in reality is correlated with a prediction of the model. A theoretical model in economics is judged by the plausibility of both its assumptions and its conclusions. Experiments are used to verify its assumptions and its conclusions are often compared with empirical and experimental data. Should we change the model if one of its assumptions is refuted? Why do we need a model to evaluate the plausibility of conclusions?

Let us consider, just as an example, the question of evaluating the plausibility of assumptions regarding time 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 above, 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)), states quite unequivocally that: “Studies of animal and human behavior suggest that discount functions are approximately hyperbolic”. Such statements are based on virtual and real human experiments. There are more people who prefer an apple today over two apples tomorrow than there are who prefer 2 apples in 21 days over 1 apple in 20 days. Thus, since we want the assumptions of our models to agree with experimental findings, then reliable data which rejects the assumptions leads us to look for new assumptions which
better fit the data. In our case we have reliable evidence (especially as it is confirmed by own thought experiments) that for certain decision problems stationary discounting is inconsistent with the experimental results while hyperbolic discounting preferences tend to fit the data much better. So, we adopt hyperbolic discounting or, to be more precise, a simple version of this approach characterized by two parameters, $\beta$ and $\delta$. But what if we can easily design experiments which reject the alternative theory as well?

Following are the results of an experiment I conducted recently on the audiences of a lecture I delivered at two universities. Students and faculty at the University of British Columbia were asked to respond online to Problem 1:

**Problem 1**

*Imagine you have finished a job and have to choose between two payment schemes:*

*A) Receiving $1000 in 8 months time.*

*B) Receiving $500 in 6 months and $500 in 10 months.*

*What scheme would you choose?*

Students and faculty invited to a lecture at Georgetown University were asked to respond online to Problem 2:

**Problem 2**

*Imagine you have bought a computer and you have to choose between two payment schemes:*

*A) Paying $1000 in 8 months time.*

*B) Paying $500 in 6 months and $500 in 10 months.*

*What scheme would you choose?*

A “present value calculation” with $\delta < 1$ (or positive interest rate) leads of course to preferring $A$ in the first problem and $B$ in the second. One way to think about this is that receiving $1000 in 8 months time is not much different from receiving $500 at $8 - \epsilon$ and
$500 at 8 + \epsilon$. 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 $B$ in Problem 1 and $A$ in Problem 2. Here are the “survey” results:

<table>
<thead>
<tr>
<th>Problem</th>
<th>University</th>
<th>#</th>
<th>8</th>
<th>6/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Receipt</td>
<td>U. British Columbia</td>
<td>354</td>
<td>54%</td>
<td>46%</td>
</tr>
<tr>
<td>2-Payment</td>
<td>Georgetown U.</td>
<td>382</td>
<td>39%</td>
<td>61%</td>
</tr>
</tbody>
</table>

The survey results are the opposite of what is predicted by the standard economic approach. In fact, a majority of subjects chose one payment when they had to choose between ”gains” and an even larger majority chose two installments when they had to choose between ”losses”.

I believe that the phenomenon we are observing here is related to the findings of Kahneman and Tversky (1979) in the context of decision making under uncertainty. People tend to prefer the average, certain expectation of a lottery when the lottery involves only gains and tend to prefer the lottery itself when it involves only losses. In the context of streams of money the averaging is done on the time component and leads one to prefer one installment in the case of receipts and the multiple installments in the case of payments.

So should we dismiss the hyperbolic discounting model? According to the methodological principles implicitly followed by some behavioral economics, the answer is yes. Of course, there is an alternative, to simply dismiss evidence we don’t like. I know of one paper which presented the results of several experiments aimed at refuting the hyperbolic discounting theory. The editor of a very prestigious journal, which has published many of the hyperbolic discounting papers, commented 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.”
Another senior economist was generous enough to quote the paper in the footnote of a survey of time preferences but dismissed the material as frame dependant. As if there is an experiment which isn't frame dependent!

Taking a more serious approach, we are faced here with one part of the dilemma of response to reality. We want our assumptions to reflect reality, but you can spell out a combination of reasonable assumptions and someone will find an experiment to defeat your theory. So how can we find a balance between our desire for reasonable assumptions and the fact that rejecting assumptions using experimental results is so easy?

Theoretical Economic models are also used to suggest regularities in human behavior and interaction: By regularities I mean similar phenomena which are repeated in similar social scenarios at different points in time and at different locations. Do we need economic theory to find these regularities? Somehow, we hope that real life regularities will miraculously emerge from the formulas we write leisurely at our desks. Wouldn’t it be better to go in the opposite direction: examining the real world, whether through empirical or experimental data, to find unexpected regularities? My limited personal experience creates doubt in my mind as to the need for theories to find regularities.

To illustrate, let us have a look at the traveler’s dilemma (due to Kaushik Basu):

Imagine you are one of the players in the following two-player game:
- 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?
Assuming that the respondents care only about their final dollar payoff, the only equilibrium strategy in this game is the choice of 180. Thus, the standard game theoretic analysis points to the unique prediction that all participants in the game will choose 180. I am not familiar with any game theoretical model that would predict a distribution of responses like the following:

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>181-294</td>
<td>295</td>
<td>296-8</td>
<td>299</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>13%</td>
<td>15%</td>
<td>5%</td>
<td>3%</td>
<td>9%</td>
<td>56%</td>
<td></td>
</tr>
</tbody>
</table>

During the past two years, I have had the opportunity to collect large amounts of data from audiences of a public lecture titled “John Nash, Beautiful Mind and Game Theory”, which I have delivered at nine universities. In the lecture, I spoke about my personal encounter with John Nash, critically introduced the basic ideas of Game Theory and spoke a little bit about the book and the movie. People who planned to attend the lecture (mostly students and faculty) were asked to respond to several questions via the site gametheory.tau.ac.il before the lecture.

Here are the results for 9 universities in 5 countries: Beer Sheva, Tel Aviv, Technion (Israel), Tilburg University (Holland), the London School of Economics (UK), University of British Columbia and York Univeristy (Canada), Georgetown University (USA) and Sabanci (Turkey), where I delivered the lecture:
The six graphs look quite similar revealing a regularity which I have no explanation to.

Any hope to explain such a distribution will require a better psychological understanding of the meaning of each of the responses. The group of players who chose 180 seem to be playing according to the game theoretical prediction, they do extremely badly and can consider themselves to be “victims” of game theory. The subjects whose answers were in the range 295-9 clearly exhibit strategic reasoning. 300 seems to be an instinctive response in this context and the responses in the range 181-294 appear to be the results of random choice.

To support this interpretation I found it useful to gather more data. For 7 out of the 9 lecture audiences I also recorded the subjects’ response time. It is true that response time is a very noisy variable due to differences in server speeds, differences among subjects in the speed with which they read and think, etc. Nevertheless, when the sample is large enough, as this one was, we should get a reliable picture. The following table and figure summarize the results for 2985 subjective:

<table>
<thead>
<tr>
<th></th>
<th>Tilburg</th>
<th>Israel</th>
<th>LSE</th>
<th>UBC</th>
<th>GTU</th>
<th>Sabanci</th>
<th>York</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>613</td>
<td>687</td>
<td>656</td>
<td>368</td>
<td>392</td>
<td>506</td>
<td>280</td>
</tr>
<tr>
<td>Avg</td>
<td>277.8</td>
<td>277.8</td>
<td>280.7</td>
<td>276.1</td>
<td>280</td>
<td>262.8</td>
<td>266.1</td>
</tr>
<tr>
<td></td>
<td>Nash Lectures</td>
<td>Median Response Time (MRT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>---------------</td>
<td>----------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2985</td>
<td>77s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>13%</td>
<td>87s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>181-294</td>
<td>14%</td>
<td>70s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>295-299</td>
<td>17%</td>
<td>96s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>55%</td>
<td>72s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is a clear pattern in the responses: the response 300 and the results in the range 181-294 (which seem to be random without a clear rationale) are the quickest. Apparently 300 is indeed the instinctive response. The results 181-294 seem to be the result of a "random" process. The responses in the range which require more cognitive efforts, i.e. 295-299, indeed take the most time. The "victims" of game theory are somewhere in between. The shape of their graph seems to indicate that some of them calculated the equilibrium (a cognitive operation) and that some of them were already familiar with the game. Thus, the time response data adds meaning to the various results, but we are still very far from explaining the stable distribution of responses.
across different populations.

So, we have arrived at the second part of the dilemma of response to reality. We want our models to produce interesting conclusions which are consistent with observed regularities. However, finding interesting regularities can be done very satisfactory without solving complicated models but rather by looking at data directly, even without having any model in mind.

4. The Dilemma of Relevance

Yes, I want to change the world just like everyone else. I want people to listen to me. But as an economic theorist, do I have anything to tell them?

One of my earliest interests in economic theory was bargaining theory. There were two sources for my interest: First and most important, it allows for the construction of models which are simple but nevertheless rich in results which have attractive interpretations. The possibility of deriving meaningful statements through the manipulation of mathematical symbols was something which 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 image for me. I came to prefer bargaining theory over auction theory since auctions were associated with the rich while bargaining was associated with the ordinary people in the markets of Jerusalem. But, I have never dreamed of becoming 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, politely telling them that as an economic theorist I had nothing to contribute. I did not say that I lacked common sense or life experience which might be useful in negotiations, but rather that professional knowledge was of no help in these matters. This response was sufficient to deter them. Decision makers usually look for an assertive advisor and not one who is offering common sense; they believe, perhaps rightly so, that they have at least as much of that as do professional economists.
But I am a micro economics teacher and as a lecturer I am a part of a big “machine”
which I suspect is brainwashing students to think in a way which I do not particularly like.
A few months ago I conducted a survey among six groups of Israeli students. The
students were approached by E-mail and had to respond to a series of questions on the
web (for a demo see http://gametheory.tau.ac.il/expEconEng/ ). The six groups were
comprised of undergraduate students in the departments of Economics, Law,
Mathematics and Philosophy at Tel Aviv University, MBA students at Tel Aviv University
and economics undergraduates at the Hebrew University of Jerusalem. I will refer to the
six groups using the abbreviations Econ-TAU, Law, Math, Phil, MBA, Econ-HU. The
students were explicitly told that the questionnaire was not an exam and that there were
no “right” answers. The core of the questionnaire was the following question:

**Q1-Table (translated from Hebrew)**

Assume that you are vice president of ILJK company. The company provides
extermination services and employs permanent administrative workers and 196
non-permanent workers who are sent out on extermination jobs. The company was
founded 5 years ago and is owned by three families. The work requires only a low level
of skills: each worker requires only one week of training. All the company’s employees
have been with the company for between three to five years. The company pays its
workers more than minimum wage. A worker’s salary includes payment for overtime
which amounts to 4,000 to 5,000 shekels per month (comment: the minimum wage in
Israel was about 3,335 IS at the time of the experiment). 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 its profits although the company is still in
the black. You attend 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
the following forecast of annual profits:
<table>
<thead>
<tr>
<th>Number of workers who will continue to be employed</th>
<th>Expected annual profit in millions of IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (all the workers will be laid off)</td>
<td>Loss of 8</td>
</tr>
<tr>
<td>50 (146 workers will be laid off)</td>
<td>Profit of 1</td>
</tr>
<tr>
<td>65 (131 workers will be laid off)</td>
<td>Profit of 1.5</td>
</tr>
<tr>
<td>100 (96 workers will be laid off)</td>
<td>Profit of 2</td>
</tr>
<tr>
<td>144 (52 workers will be laid off)</td>
<td>Profit of 1.6</td>
</tr>
<tr>
<td>170 (26 workers will be laid off)</td>
<td>Profit of 1</td>
</tr>
<tr>
<td>196 (no layoffs)</td>
<td>Profit of 0.4</td>
</tr>
</tbody>
</table>

I recommend continuing to employ ______ of the 196 workers in the company.

**Observation 1: There were sharp differences between groups in confronting the dilemma of profit maximizing vs. worker layoffs.**

The following table presents the 764 responses (of 100 or more) to question 1-Table:

<table>
<thead>
<tr>
<th>Q1-Table</th>
<th>EconHu</th>
<th>EconTA</th>
<th>MBA</th>
<th>Law</th>
<th>Math</th>
<th>Phil</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>94</td>
<td>130</td>
<td>172</td>
<td>216</td>
<td>64</td>
<td>88</td>
<td>764</td>
</tr>
<tr>
<td>100</td>
<td>49%</td>
<td>45%</td>
<td>33%</td>
<td>27%</td>
<td>16%</td>
<td>13%</td>
<td>31%</td>
</tr>
<tr>
<td>144</td>
<td>33%</td>
<td>31%</td>
<td>29%</td>
<td>36%</td>
<td>36%</td>
<td>19%</td>
<td>31%</td>
</tr>
<tr>
<td>170</td>
<td>7%</td>
<td>9%</td>
<td>23%</td>
<td>18%</td>
<td>25%</td>
<td>25%</td>
<td>18%</td>
</tr>
<tr>
<td>196</td>
<td>6%</td>
<td>13%</td>
<td>12%</td>
<td>13%</td>
<td>11%</td>
<td>36%</td>
<td>15%</td>
</tr>
<tr>
<td>other</td>
<td>4%</td>
<td>2%</td>
<td>3%</td>
<td>6%</td>
<td>13%</td>
<td>7%</td>
<td>5%</td>
</tr>
<tr>
<td>Average</td>
<td>127</td>
<td>133</td>
<td>142</td>
<td>144</td>
<td>151</td>
<td>165</td>
<td>143</td>
</tr>
</tbody>
</table>

The differences between the groups are striking. The economics students both at the Hebrew University and Tel Aviv University are much more pronounced profit maximizers than the students in the other groups. 45-49% 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 number of respondents (ranging from 6-15%) in five of the six groups; the only exception was the philosophers - 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. My conjecture is that this has to do with the way in which the MBA program is taught. The study of cases triggers more comprehensive thinking about real life problems. Study using formal exercises conceals the need to balance between conflicting considerations.

Following their response to question 1, all subjects had to indicate what do they thought would be the choice of a real vice president?

**Observation 2:** There were almost no differences between groups as to what the subjects thought a real vice president would do.

In Law and Phil all subjects received the version 1-Table (presented above). The other four groups, who were better trained mathematically, were randomly allocated two versions of the question. Q1-formula was identical to question 1-Table with the only difference that the table was replaced with:

“The Finance Department has prepared a forecast of profits according to which the employment of x workers will result in annual profits of (in millions of shekels):

\[
2\sqrt{x} - 0.1x - 8
\]

This profit function yields similar values to those presented in the table. Its maximum is at \(x = 100\). Note that Q1 explicitly emphasized that with no layoffs, profits will be still positive.

**Observation 3:** The formula distortion

The following table summarizes the 298 answers of 100 or more:

<table>
<thead>
<tr>
<th></th>
<th>EconHu</th>
<th>EconTA</th>
<th>MBA</th>
<th>Math</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>55</td>
<td>74</td>
<td>125</td>
<td>44</td>
<td>298</td>
</tr>
<tr>
<td>100</td>
<td>74%</td>
<td>77%</td>
<td>73%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td>101-195</td>
<td>10%</td>
<td>9%</td>
<td>11%</td>
<td>15%</td>
<td>11%</td>
</tr>
<tr>
<td>196</td>
<td>16%</td>
<td>14%</td>
<td>15%</td>
<td>10%</td>
<td>14%</td>
</tr>
<tr>
<td>Average</td>
<td>120</td>
<td>117</td>
<td>120</td>
<td>116</td>
<td>118</td>
</tr>
</tbody>
</table>

There are no major differences between the groups. A vast majority of subjects in all groups maximized profits though 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 said that they believe that a real vice president would fire less than the number required to maximize profits). Thus, presenting the problem formally, as we do in economics, seems to conceal the real life complexity of the situation from most students (including Math students).

Our view of the results cannot be separated from our personal evaluation of the behavior of economic agents in such a situation. If you believe that the manager of a company is obligated morally or legally to maximize profits, then you should probably hail economics for its achievement in educating its students so well. On the other hand, one might approach the results with the belief that a manager should also take into account the welfare of his workers, particularly when the economy is in recession and unemployment is high, but then one feels uncomfortable about 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 and not a result of indoctrination. The fact that the economists are different from the lawyers and MBA students and not only from the philosophers and mathematicians makes this possibility more doubtful. The minimal differences in the responses to questions 1-formula appear to also somewhat support the indoctrination hypothesis.

And, it may be that there is no connection between the responses of subjects in such a questionnaire and the choices they would make in practice. But if there is no connection, are we saying that what a student learns in economics will have no influence on his future behavior? And if there is such a connection, shouldn’t we be revising our curriculum?

Overall, I am left with the suspicion that in the best case the formal exercises we give our students, make the study of economics less interesting; in the worst case, they contribute to the shaping of a rather unpleasant "economic man". I find it difficult to say that the way I teach economics does not effect the world in a direction I am not happy
Those guilt feelings led me to a paper I coauthored with my friend Michele Piccione from the London School of Economics. This may be provocative but this was my only paper which was motivated by real world problems and was intended to have an impact on the way that economists think.

In the paper we presented a model we call the “Jungle”: $< N, K, w, \{\succeq^i\}_{i\in I}, \{X^i\}_{i\in I}, S >$. The first two ingredients of the model are a set of agents $\{1, \ldots, N\}$ and a set of commodities $\{1, \ldots, K\}$. The agents can distribute a bundle $w = (w_1, \ldots, w_K) \gg 0$ among themselves. Each agent $i$ is characterized by a preference relation $\succeq^i$ on the set of bundles $R^i$ and by a consumption set $X^i \subseteq R^i$ specifying the bounds on individual’s ability to consume (or the amount he can possess). The set $X^i$ is compact and convex and satisfies free disposal. The preferences of each agent satisfy strong monotonicity, continuity and strict convexity. The original element in the model: a strength relation $S$. We assume that $S$ is a linear ordering and thus, without loss of generality, assume that $1S2, 2S3, \ldots, (N - 1)SN$. The statement $iSj$ is interpreted as “$i$ is stronger than $j$” and has the meaning that $i$ can take from $j$ anything that $j$ has. 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 stronger than another is interested in the transaction. The notion of power in our model is quite narrow. If we were after a more realistic model, we would probably introduce power in a more elaborate way and would allow, for example, the possibility that an attacker could lose in a confrontation, and that the meaning of strength is a higher probability of winning a fight. Overall, 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 to the model with an initial endowment. The vector of initial endowments is replaced in this model with the power relationship.

By the way, the jungle system is very popular outside the jungle as well. When we talk about power we don’t necessarily mean physical power. Power is often more subtle. In the male-female “market”, for example, charm and attraction play a key role in obtaining a mate. In an office parking lot, social conventions such as seniority determine the
allocation of parking spaces. The power of persuasion enables certain individuals to convince others to take actions which are against their own interest.

A **feasible allocation** is a vector of non-negative bundles \( z = (z_0, z_1, z_2, \ldots, z_N) \) such that \( z_0 \in \mathbb{R}_+, z_i \in X_i \) for \( i = 1, \ldots, N \), and \( \sum_{i=0}^{N} z_i = w \). The vector \( z_i \) is the bundle held by agent \( i \) ("held" not "owned" since there is no legal system here) and \( z_0 \) is the collection of goods that were abandoned and from which anybody can take what he likes.

A **jungle equilibrium** is a feasible allocation \( z \) such that there are no agents \( i \) and \( j \), \( i \neq j \), and a bundle \( y_i \leq z_i + z_0 \) or \( y_i \leq z_i + z_j \) such that \( y_i \in X_i \) and \( y_i \succ_i z_i \). In other words, a jungle equilibrium is an allocation such that no agent can assemble a preferred bundle from the bundles that are freely available or held by himself and an agent who is weaker than him.

Several propositions are proved:

1. **Existence**: A jungle equilibrium exists.
2. **Uniqueness**: Under smoothness assumptions on the preferences and consumption sets of the agents, a jungle equilibrium is unique and has the additional property that no agent can assemble a more desirable bundle even from goods that he assembles from all weaker agents or are freely available.
3. **“First Fundamental Welfare Theorem”**: Under the smoothness assumptions the jungle equilibrium is efficient.

To compare the model with the standard exchange model, let’s consider an example: Consider a market with \( N \) indivisible houses in which each individual can possess only one house and has strict preferences over the set of houses. In the jungle the individuals are endowed with power. The jungle equilibrium is unique. Successively, agents are asked to pick the best house that has not been chosen earlier. Relative power determines the order of the queue. In the market the individuals are endowed with an initial house. David Gale constructed a competitive equilibrium for this simple market. You look for a “top cycle” of agents who can permute their houses so that each will get his first choice. You assign the highest price to these houses, delete these
agents and their houses from your list and move on. The priorities in the market are determined by the existence of complementariness in preferences.

With this example one can demonstrate the second fundamental welfare theorem for the jungle. Recall that the question that the second welfare theorem comes to answer is as follows: given that there exists an efficient allocation, is there an allocation of ownership which yields that allocation as a competitive equilibrium outcome of the exchange economy with this initial allocation. The analogous question for the jungle is: for any given efficient allocation, is there a power relation which yields that allocation as a result of an equilibrium in the jungle with this power relationship. In our example, any efficient allocation is supported as a jungle equilibrium for any power relationship which completes the relation $jSi$ if “$i$ envies $j$” (this is possible since the relation $i$ envies $j$ is acyclic for any efficient allocation).

Another interpretation of the second welfare theorem is that the supporting market equilibrium prices allow the comparison of wealth between agents. In our example, (and things get more complicated beyond this example) every jungle equilibrium allocation is supported by equilibrium prices such that the stronger are also the richer. One might like to interpret this statement as saying that power and wealth go hand in hand.

When I present the 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. Making an analogy, I argued that the market economy accepts the natural desire of people to be richer, to have more. In the same way, the jungle economy accepts people’s desire to use their strength to take advantage of those weaker than them. The market economy encourages people to produce more, the jungle economy encourages people to develop their power, thus facilitating society’s expansionist desires.

I do not view our jungle model as a model of the real jungle. Rather, it was a rhetorical exercise. The whole idea was to build a model which 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. Standard economic courses impress students with their elegance and clarity. We have tried to do the same with the model of the jungle. This exercise is directed at economics students with the goal of creating more question marks in their minds when they study models of competitive markets.

This brings me to the dilemma of Relevance. I believe that as an economic theorist I have very little to say which is of relevance in the real world and I do believe that there are very few models in economic theory (and the more elaborate ones not among them) that could be used to provide serious advice. But I cannot hide behind the view that there are pure theoretical musings. I cannot ignore the feeling that our work as teachers and researchers influences students’ minds in a direction I am not happy with. Can I find a way to be relevant or am I just another charlatan?

5. Concluding Words

It’s time to sum up the discussion. How do I relate to these three dilemmas?

As economic theorists, we organize our thoughts using what we call “models.” 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. He has some moral he wishes to impart to the reader. The fable is an imaginary situation, somewhere between fantasy and reality. Any fable can be dismissed as being “unrealistic” or simplistic. But this is also a fable’s advantage. Being between fantasy and reality gets rid 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 we can use 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 few themes and elucidates them. We perform thought
exercises which are only loosely connected to reality and which have been stripped of most of their real life characteristics. However, in a good model, as in a good fable, something significant remains from our perception of the real world.

True, we have become accustomed to writing these fables in formal language. The main advantage of this is the introduction of discipline. The teller of this type of fable cannot adopt an ending which does not follow from his assumptions. The ending has to be based on reasonable assumptions that are spelled out before presenting the moral of the story. The reader who understands this formal language can objectively consider statements that begin with expressions like "Necessarily," "This shows" or "It follows from our assumptions." On the other hand, formal language makes what we do seem scientific. Furthermore, the formal language narrows the fable's audience to those trained to understand formal models. We have escaped the ambiguity of words by creating a secret language that only we can understand.

Like us, the teller of fables confronts the dilemma of Absurd Conclusions. The logic of his story may lead to absurd conclusions as well. Like us, the teller of fables confronts the dilemma of Responding to Reality. He wants to maintain a connection between his fable and what he observes. There is a fine line between a fantasy without content and a fable with a message. 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.

Absurd conclusions reveal contexts in which we find the model or the fable to be reasonable and may not necessarily make the model or the fable uninteresting. Fable and models in economic theory are derived from observations of the world but they are not meant to be testable. A good fable and 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, that is, the collection of ideas and conventions which people believe in and which influence the way they reason and act.
Yes, I do believe we are simply tellers of fables. But, isn’t it wonderful?
References

Cox, J.C. and V. Sadiraj (2001) "Risk Aversion and Expected-Utility Theory: Coherence for Small- and Large- Stakes Gambles", (mimeo),


