

**UNIVERSITY
OF OSLO**
HEALTH ECONOMICS
RESEARCH PROGRAMME

**Defence of Absurd
Theories in Economics**

Morten Nordberg

Ole Røgeberg

*The Ragnar Frisch Centre for Economic
Research and HERO*

Working Paper 2003: 18



**HE
RO**



Defence of Absurd Theories in Economics*

Morten Nordberg and Ole Røgeberg

October 2003

Health Economics Research programme at the University of Oslo
HERO 2003

Keywords: As-if theory; Economic methodology, welfare economics

JEL Classification: B41, D00

* This paper is part of the HERO program at the University of Oslo. We would like to thank all those who commented on earlier drafts, with a special thanks to Arvid Senhaji for his extensive comments and suggestions.

Correspondence to:

The Ragnar Frisch Centre for Economic Research, Gaustadalleen 21, N-0349 Oslo, Norway

E-mail: morten.nordberg@frisch.uio.no; o.j.rogeberg@frisch.uio.no

Abstract

Theories that involve plainly false and even bizarre assumptions are argued to have an important role in bundling empirical facts in a way that allows these to be understood, handled and used as modules in the construction of mechanisms by economists with human cognitive limits. Absurd theories are subcomponents used in a valid explanatory strategy as long as the mechanisms only derive the implications of the facts summarised. This provides a defence and explanation of many economic theories, but also imposes hard limits on such theorising.

Introduction

Are all economists quite mad? The common and long-standing claim that they believe in the continuing real-world existence of general equilibrium, flawlessly profit-maximising firms and über-rational human logic-machines that calculate the optimal amounts of laziness and the welfare maximising paths of heroin consumption might suggest so. Anyone agreeing that such assumptions poorly approximate the real world faces “perhaps the key methodological problem in contemporary economics” (Backhouse (1999:124)): “[H]ow a model, based on assumptions that are typically not true, can tell us about the real world” (Backhouse (1999:122)).

In what follows we explain why and how clearly absurd theories can tell us about the real world, and what role such theories can validly play within economics. We do this by giving a new justification for what we call the *as-if* methodology – the (vulgarised?) version of Milton Friedman’s methodological views that sees the realism of assumptions as completely irrelevant to how we judge a theory. Furthermore, our defence rests on the same kind of facts of human cognitive limitations that many critics of economics use to argue against our theories. The short version of our claim: Absurd theories in economics can exploit mental models that humans naturally use to bundle established facts in untrue mock-explanations. This allows us to remember, communicate and exploit the bundled facts and to use them as building blocks in sensible theories that deduce the implications of the facts summarised by our absurd theories.

As this summary shows, our defence of absurd theories involves a claim that they represent a distinct kind of theory that should be constructed and evaluated in very different ways from other theories. For this purpose, we suggest three categories of economic theory – As-if theories, mechanisms and pure theory – each discussed in a separate chapter. Our claim is that economic theories are quite sensible as long as the different kinds of work are not confused with each other.

The confusion we wish to remove may be indicated by noting a variety of common views: Graduate level textbooks state that utility functions represent binary preference relations defined over all choice alternatives, and that they involve no assumption concerning altruism, selfishness or other sources of motivation. This “formal” and “official” view clashes with practise, where economists discuss whether it is “rational” to have a preference for, say, giving to charity or whether this presupposes some “warm glow,” etc.¹ Real world data are explained and used to “test” economic models of individual decision making even when the models are based on seemingly arbitrary collections of possible incentives. Sometimes the data will be “average” consumption levels across individuals, with entry-exit decisions explained by our model as marginal adjustments by our hypothetical “representative agent.” A common defence is that we are only interested in prediction, yet if someone attempts a theory without “rationality” or “profit-maximising firms” this is seen as ad-hoc data fitting rather than a theory experiment to be evaluated with tests of predictive power. We need the rationality assumption, it is claimed, as a defence against “anything goes” theories that specify tastes and decision rules “ad-hoc.” Yet it is the rationality assumption that has been repeatedly argued by critics to be vacuous and compatible with any evidence, while “deviations” from rationality and profit-maximisation are often based on empirical work by psychologists such as Kahneman and Tversky and economists such as Herbert Simon.

In chapter 1, we note two strategies that have been used to defend economic theories – denying absurdity and asserting its irrelevance – and briefly define “as-if methodology.” In chapter 2 we present our view of what as-if methodology is and should be, and why it is necessary and highly useful. In chapter 3 we discuss research on mechanisms, and explain why this kind of work (sometimes labelled deductivism) builds on as-if theories more than it provides an alternative to such theories. In chapter 4 we briefly discuss pure theory, and in chapter 5 we summarise the discussion.

1 Some previous attempts to defend economic theories

Two simple strategies for answering the criticism that economic theories are absurd are, first, to deny that they are, and second, to deny that it matters. Both these

¹ Others who have noted this seeming contradiction are Broome (1999) (Chapter 2) and Sen (1982)

strategies have been pursued by economists, and since the as-if defence developed in this paper contains elements of both it is useful to briefly consider them.

A clear and well-written denial of the fact that economic theories are absurd in any important sense is Gibbard and Varian (1978). They liken different kinds of models to photographs, realistic drawings and caricatures and state that economic models ask “What would happen if such and such were the case?” (p. 668). The reason we should trust economic models, according to them, is that the assumptions of applied theories are approximations, hypothesized by the economist to be “close enough to the truth for his purposes” (p.669). Simpler, cruder models are attempts to capture tendencies robust to the details of the specification, etc.

As the response from Dancy (1978) makes amply clear, this defence does not convince those who feel that economic theories are mere exercises in mathematics used to prop up ideological views:

Gibbard and Varian’s message has no bearing on that feeling, for those who share it are rejecting neo-classical theory not because it employs models, but because it employs poor models, and pronouncing them poor not because they are only approximations, or highly selective, but because they are not even approximations, and not selective but fictional. (Dancy (1978:677))

A second response, from philosopher Alexander Rosenberg, sees Gibbard and Varian’s account “more like the description of a problem than its solution” (Rosenberg (1978:683)), and concludes by asking

[w]hy it is that economists have continued for so long to use models with false assumptions, and indeed with the very same false assumptions, not just for decades, but for more than a century, even though they purport to be engaged in an attempt to understand the world. (p. 683)

The “our theories are not absurd” strategy fails because critics feel that our theories are so obviously based on false assumptions. To show one example:

People are assumed to want to get as much for themselves as possible, and are assumed to be quite clever in figuring out how best to accomplish this aim. Indeed, an economist who spends a year finding a new solution to a nagging problem – such as the optimal way to search for a job when unemployed – is

quite content to assume that the unemployed have already solved the problem and search accordingly. (Richard Thaler, quoted in Buchanan (2002:187))

Economics, then, does not need a defence that talks about “good approximations” or “capturing the main forces” or “essentials.” We need a defence of theories that do not even approximate reality, with plainly wrong assumptions and major, known determining factors left out.

The probably most widely read paper on the method of economics – Milton Friedman’s essay on “The Methodology of Positive Economics” – has often been read as offering a defence of such absurd theories. His essay and views are more nuanced than his simple, readable style might suggest at first glance, a fact revealed by the large secondary literature that has quarrelled over how he is to be understood: As shown by Stanley (1985), there are at least four different labels that have been slapped on Friedman’s essay by various economists and philosophers of science, some of them mutually contradictory (for instance falsificationist and verificationist)².

The question of what Friedman *really* meant might be interesting for historical reasons, for shedding light on his own contributions to economic theory, etc., but our present concern lies more with the message received than with the message transmitted: According to Hausman (1992:164), “Lee Hansen [...] recalls economists in the 1950s reacting to Friedman’s essay with a sense of *liberation*. They could now get on with the job of exploring and applying their models without bothering with objections to the realism of their assumptions.” A similar understanding of “as-if methodology” seems common today. A “slogan” that captures this view of the methodology is that *realism of assumptions is irrelevant as long as the theory predicts well, since the only goal of economic theories is to predict*. Several quotes can be selected from Friedman’s essay to point in this direction, at times making it seem that lack of realism is a virtue:

Truly important and significant hypotheses will be found to have “assumptions” that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions (in this sense). [...] To be important [...] a hypothesis must be descriptively false in its assumptions [...]. (Friedman (1953:14))

² Hirsch and De Marchi (1984), adds two more and are a good entry point into the discussion of Friedman’s essay.

Whether quotes such as this one give a fair and representative picture of the totality of Friedman's essay is not our concern here, since we are interested in the simplified message that was picked up by other economists. A full and nuanced understanding of some sub-discipline or theory in economics requires specialisation, and since any economist can specialise in only a relatively small number of fields, most arguments and results from one subdiscipline will be used mainly in simplified and even misunderstood ways by other subdisciplines. This suggests that there is limited reason to believe that an in depth study of Friedman's views in light of his own work would help us understand what others of his and later times have understood by his essay. For this reason, we use the term "as-if methodology" rather than "Friedman's methodology of positive economics." This term – *as-if* – is used repeatedly in Friedman's essay and has been picked up by the profession since. It seems, in our experience, often to function as a blanket defence of a model's assumptions: "This is standard as-if methodology," or "obviously this model is not true, it's just meant to be an as-if model."

2 As-if "theories"

It is a common human desire to understand something, to explain why something is the way it is or how something works. Since the as-if method results in theories that – taken "literally" as realist theories making claims about the true nature of people and firms – seem clearly untrue, it becomes easy to judge the as-if method harshly. "Despite its influence," as Hausman (1992:163) notes, "Friedman's essay has not been admired by other writers on economic methodology." Herbert Simon, for instance, took time in a footnote to the published text of his Nobel Lecture to state that "I cannot in this brief space mention, much less discuss, the numerous logical fallacies that can be found in Friedman's 40-page essay" (Simon (1979:4)), often focusing on the "realism" issue in his comments.

In our view, Simon's criticism is like criticising a hammer for not being a saw. As-if theories are not intended to be true, not even approximately. Their function is to take established facts and create a theory/model/story/explanation that implies these facts, thus simplifying our knowledge by folding it into a mock-explanation. The theory/model/story/explanation can be called a mock-explanation because it is untrue and has no intention of being true: Its assumptions may be plainly false, its implied

causal mechanisms clearly irrelevant. It may make absurd assumptions concerning how people reason, how businesses are organised, how businesses determine prices, etc., yet it can still serve a valuable purpose: Bundling facts into building blocks. The mock-explanation allows us to remember, grasp, understand and manipulate the bundled facts, and it allows us to use these as modules (building blocks) when creating higher-level mechanisms built on the facts.

2.1 The need for mental models

The philosopher and mathematician Bertrand Russel once argued forcefully that the notion of causes in itself was unnecessary in science and that systems of equations relating different observables were all that we could observe. The word “cause” was “so inextricably bound up with misleading associations as to make its complete extrusion from the philosophical vocabulary desirable” (Russel (1913)).

An analogous claim in economics is that any talk of “unobservables” such as rationality, motives/preferences/desires, beliefs/expectations, etc. are “unscientific” and that we should restrict ourselves to making theories at the purely formal level where our theories relate entirely to observables. Such an argument holds an obvious appeal to the positivist in us, since anything beyond what we can observe would seem to be “pure speculation” or influenced by unsystematic and unrigorous common sense and intuition. The perhaps clearest case in economic theory of trying to follow this line of work in practise is Samuelson’s Revealed Preference Theory (see quotes in Sen (1993:496-497)).

One might ask why Samuelson’s and Russel’s projects failed to gain dominance. One reason, we would suggest, is that they did not take seriously the constraints of human thought: Theories along the above lines would to a large extent be consequences deduced from large collections of loose facts. The problem with this is that people are unable to hold large collections of arbitrary facts in mind, and when they try they immediately start structuring them and creating mental models that account for them (Norman (2002:38-39,66-72)).

The problem is not just to remember the facts, but also to deduce their implications. Humans are inherently good with some kinds of mental models, and a problem involving these can be solved easily even when the – formally same – problem is hard to solve when formulated in a different way. For instance, one

experiment has involved showing people a set of four cards. You are told that the cards have letters on one side and numbers on the other. Your job is to choose which of the visible cards you need to turn around to test the rule “If a card has D on one side, it has a 3 on the other.” The cards shown read (on the visible side) “D,” “F,” “3,” and “7.” The results show that only about five to ten percent were able to do this correctly, and that even people who had taken logic courses tended to get it wrong. BUT: The exact same problem can be put into a context, given a mock-explanation, that enables people to reason validly:

You are a bouncer in a bar, and are enforcing the rule “If a person is drinking beer, he must be eighteen or older.” You may check what people are drinking or how old they are. Which do you have to check: a beer drinker, a Coke drinker, a twenty-five year old, a sixteen-year-old? (Pinker (1999:336)).

According to Pinker, it seems that “people get the answer right when the rule is a contract, an exchange of benefits. In those circumstances, showing that the rule is false is equivalent to finding cheaters.” One hypothesis is that “the mind seems to have a cheater-detector,” and that this would make sense from an evolutionary perspective.

A different example concerns our ability to do valid Bayesian inference. Only 18% (12% in a replication study) were able to find the Bayesian answer to a relatively simple problem in medical diagnosis when presented with conditional and unconditional probabilities. The same information, in a form designed to correspond to how humans process information, increased the proportion who found the Bayesian answer to 76% (Gigerenzer (2000:Chapter 6)).

Our defence of economic theories can now be formulated as a solution to a design problem: We have a complex social reality we wish to understand, and (we presuppose) there are some stable, observable patterns. Instead of trying to take an overwhelming collection of disconnected facts and seeing their consequences, we bundle the facts together using mock-explanations that are absurd and patently false as-if theories not even loosely approximating reality. Since the bundles exploit the kinds of mental models people are inherently good at using, they allow us to combine the bundled facts and use them in our thinking and in the construction of higher-level theories that build on these facts. A well-designed as-if theory thus exploits precisely

those everyday reasoning principles that Russel and Samuelson would remove:
Causes determining events and motives determining actions.

2.2 *The mental models humans are good at*

The ease with which humans construct and grasp teleological explanations when explaining the behaviour of others is well known, e.g. “He painted his house *because* he didn’t want the neighbours to see it as run-down and he *thought* the house was beginning to look shabby.” The theory implicit in such “common-sense,” everyday explanations – that humans make choices on the basis of beliefs to satisfy their desires – sometimes goes by the name of “folk psychology.” Our predilection for such explanations is so strong that many construct them even where they may make little sense: Shown a movie of abstract shapes bouncing around on a screen, people ascribe motives to the shapes, imagining that the little circle is *trying* to get away from the big triangle which is *trying* to catch it (Davies (1997:8)). In Aristotelian physics, rocks fell because they *wanted* to return to the earth whence they came. Functionalist theories in the social sciences sometimes ascribe sophisticated, intentional action to abstract concepts such as “class” or “institution” or “state,” asserting that these “desire” or “aim” to achieve ends that neither the individuals who constructed the institutions or the individuals who are part of them today are aware of. Finally, psychoanalytic theories would postulate obscure and unobservable desires to provide teleological explanations for any and all human behaviour. E.g. if you constantly ended up in serious accidents, you had a subconscious *desire* to punish yourself. If you messed up your life, you were *attempting* to punish your parents. And so on. What these phenomena show is that humans find it easy to make mental models assuming intentionality. Such explanations come easy, and we find them relatively simple to understand.

Similar arguments to the above would hold for causality, where our every-day concept seems to relate to experiences of pushing, pulling, striking, squeezing, throwing and other such physical manipulations. If we are presented with a story where teleological and “everyday causality” are the ingredients, the story will often “click” and trigger a feeling of understanding. This feeling carries with it both problems and benefits.

One problem with this is that we tend to be satisfied once we get the feeling of understanding. As argued in Trout (2002), however, the historical record clearly shows that this feeling is no trustworthy criterion of truth.

A second problem is that we have difficulty taking evidence into account in the absence of a good “story,” and we have difficulty taking a lack of evidence into account in the presence of a good “story.” Dawes (1999) gives examples of this as concerns probabilistic reasoning in general, and Dawes (1996) shows how good but false stories survive falsification even in science: Despite having been shown to be worthless in a number of empirical studies, a large number of psychologists persist using Rorschach tests (“What does this ink-blot look like to you?”) and keep on believing that a person’s response (“It looks like my father...”) is psychologically revealing (p.146-154). Another example concerns the extreme persistence of beliefs in “expert psychological judgment,” even when an overwhelming share of studies fail to find even weak tendencies in favour of such judgment (p. 75). A further example is the article by Milton Friedman (1992) called “Do Old Fallacies Ever Die?” Friedman notes the ignorance many show for regression towards the mean even though this phenomenon has been known to science for roughly a hundred years. If we are right, the problem should be seen as one of design: The explanation involved has yet to be phrased in a way that “fits” the thinking of educated economists.

The above remarks might seem very critical of human abilities, but it is important to note that there is an upside as well: Teleological and simple “common sense causality” explanations are grasped easily and we “see” what they imply. They can therefore be seen as resources that can be exploited, and as-if theories can be seen as one way of doing this. For instance: Consider the “selfish gene” formulation of darwinian evolution. As its creator Richard Dawkins has written, “[t]he metaphor of the intelligent gene reckoning up how best to ensure its own survival [...] is a powerful and illuminating one.” At the same time, Dawkins is well aware that it is a metaphor, and that – taken at face value – the story is false.

From 13 years’ experience of teaching it, I know that a main problem with the ‘selfish-gene survival machine’ way of looking at natural selection is [that some are] carried away, and allow hypothetical genes cognitive wisdom and foresight in planning their ‘strategy.’ At least three out of twelve misunderstandings of

kin selection [...] are directly attributable to this basic error. (Dawkins (1999:15))

2.3 *Exploiting teleological reasoning*

In economics, the teleological principle occurs in (at least) two guises: As rationality in consumers, and as profit maximisation in firms. For simplicity, we will concentrate on rationality, the case of profit maximisation is quite analogous.

The first thing to note is that rationality is an empirically empty concept. A different term for rationality is “the maximisation assumption” that “[f]or all decision makers there is something they maximize” (Boland (1981:1034)). You might establish that they do not maximise X, Y or Z, but this proves only that they do not maximise X, Y or Z. They *might* still be maximising something else, e.g. A, B or C. Rogeberg (2003) uses the extended utility approach of Gary Becker to argue that “even time inconsistencies and what has been called weakened capacity for rational choice” can be included in a rational choice theory (p. 307).

That this point still needed to be made by Boland could be seen as dejecting: Hutchison (1938:115) made this point back in the thirties by stating that, to

render [the “maximum principle”] not obviously false [economists] have had to steadily widen it, and thus to diminish its empirical content. [...] Fewer and fewer, if any, types of economic conduct remained which were not subsumed under it, and almost none were excluded or could falsify it. Its empirical content, therefore, simultaneously grew smaller and smaller. To say that a piece of economic conduct was “rational” came to mean little, if anything, more than that it was a piece of economic conduct.

The second thing to note is that, from our perspective, this non-falsifiability of rationality could be seen as a benefit rather than a problem: Since any behaviour can be explained as rational, we can take *any* observable relationship that remains stable in the face of various manipulations of relevance, and summarise the relationship by the use of a story. With the story in mind, we can use the established relationships in our reasoning, we can communicate them, we can see their implications by using them as building blocks in larger theories. The story can well be unrealistic, as long as it is consistent with the facts. If well done, it provides a Eureka experience as when a number of seemingly unrelated facts that are difficult to understand and recall and use

suddenly snap into focus as part of an understandable, intentional pattern. In other words: Rationality provides us with a universal fact-summarising machine. It is compatible with any facts, and can therefore have universal usefulness in this regard. This answers the question asked by, amongst others, Dancy (1978:679) as to why and how absurd theories can make true statements about the world.

2.4 Reinterpreting "De Gustibus Non Est Disputandum"

Our suggestion that as-if theories are stories constructed after the fact and designed to reproduce and capture as many known empirical facts as possible, may not be that far from the intention of past economists: Stigler and Becker (1977), for instance, in their famous "De Gustibus Non Est Disputandum" paper, note that it is easy to throw out seeming anomalies, behaviour or observations that seem incompatible with rational choice. "If we could answer these [...] to your satisfaction, you would quickly produce a dozen more" (p.76). They state that they interpret their Latin title as stating "that tastes neither change capriciously nor differ importantly between people." We thus have two viewpoints, one claiming that the analysis stops when we find that tastes differ, and the other claiming that tastes do not differ and the analysis only stops when we find that this is the case. According to Stigler and Becker, the choice between these two

must ultimately be made on the basis of their comparative analytical productivities. On the conventional view of inscrutable, often capricious tastes, one drops the discussion as soon as the behavior of tastes becomes important – and turns his energies to other problems. On our view one searches, often long and frustratingly, for the subtle forms that prices and incomes take in explaining differences among men and periods. If the latter approach yields more useful results, it is the proper choice. (p.76)

Note especially the term *useful* results, which can be interpreted as a statement along our lines: It is useful to search out subtle and maybe even bizarre and clearly untrue rational choice theories because it is easier to summarise behaviour patterns in such terms than it is to describe them in detail and try to use them on that level. If the as-if theory also touches on a few explanatory variables that we believe *actually* are important, it becomes even easier to recall and relate to our *everyday beliefs* about *what is actually going on*, and even easier to exploit using the "intuitive

understanding software” that seems to be running in our minds. For this reason, it becomes acceptable to selectively provide stories that provide “intuition” for theories, even when other, equally strong intuitions go against these same theories. For a theory intended to be true, however, this would have been selective, biased evidence.

This would also explain why many economists find it acceptable to sit in their offices and think up the incentives that would explain real human choices, rather than beginning with qualitative interviews, psychometric questionnaires or psychological research.

The above view of as-if theory also provides a new perspective on the resistance many economists have shown towards the work of researchers such as Kahneman-Tversky. The work in cognitive psychology can be (unfairly?) summarised as follows: A number of researchers have identified a number of interesting heuristics – rules of thumb that humans appear to unconsciously rely on in their everyday thinking - and biases reliably triggered in various situations. Let us accept that they have found a long list of decision rules and information processing procedures. We might then compare two types of theories that bundles the same facts of human behaviour. One using the “realistic” heuristics and results from psychology, the other using an absurd rational choice theory. The “realistic” one might then be a much more complicated theory to understand, communicate and reason with than the rational choice explanation, *even though* the kind of *reasoning* assumed by the cognitive psychologists might well be simpler than the kind assumed by the economist.³ That economists and psychologists should bundle the same facts in different ways is related to their differing purposes: The psychologist wants to explain human choices in terms of empirically derived sub-components. The economist wants to understand market outcomes, interaction effects, etc. and to derive these from human behaviour. The economist can therefore be content with a plainly false theory of human behaviour as long as it bundles facts in a way suiting his purposes.

One problem with the theories from cognitive psychology is thus that they fail to supply the economist with the mental model of human behaviour that these same

³ Note: More easy to understand and use by *economists*, not by lay people (who might see the rational choice theories as completely inexplicable). As Trout (2002) notes, “What counts as simple or familiar is theory-dependent, and not surprisingly, verdicts in particular cases are controversial. [...] A finding unfamiliar to the lay public may convey a feeling to the scientist that the pieces of a theoretical puzzle have just fallen into place” (p. 214, footnote).

theories have noted that humans need. The response of Machina (1999) to the Kahneman-Tversky inspired criticisms of rational choice presented in McFadden (1999), seem to illustrate this: Rational choice theory is seen as a complete model of human behaviour, while Machina sees the psychology-based alternative as a loose list of facts tied together mainly by deviating from standard rational choice theory.

2.5 *Economists need not care what makes people tick*

If we are right, the critics and supporters of Rational Choice would both do well to remember that economists are not psychologists. Economists are in the business of studying the outcome of the behaviour of interacting agents. Their demand is for useful *representations* of individual *behaviour* - conceptual robots acting consistent with observed behaviour of real people, not accurate representations of unique individuals. For such purposes, then, rational choice theory might well seem to be the best option available.

Keeping this in mind also helps us make sense of other puzzling things in the economics literature. As for instance why the most mainstream use of finds from cognitive theory is to embed them in an absurd rational choice model to provide new rational choice stories. For instance, Laibson (2001) is an interesting and well-written paper that relies on psychological research to create an absurd theory that can be summarised as follows: There is a large amount of psychological evidence that rats and people easily become physiologically conditioned to crave drugs in the presence of environmental cues. In other words: If you always smoke a cigarette after clearing your desk, clearing the desk will make you crave cigarettes. Laibson shows that this can generate interesting behaviour patterns in a rational toy person who is fully aware of these effects, lives in some weirdly unreal world where there are only two cues – which follow a known probability distribution such that one but only one is present at any time – and where the individual designs the choice-plan that maximises the expected sum of a mathematical function characterising all he cares about in arithmetic quantities. The psychological evidence he presents, on the other hand, consists of people and animals becoming conditioned *without being aware of it* to any of a *huge number of easily manipulable environmental cues*, and behaving in other respects in *direct contradiction* to his model (his model asserts that drug users should want to avoid other drug users, since they present cues with a negative direct effect on utility). In other words: The theory is plainly false if understood as a theory of why

humans do what they do, but Laibson feels that it is able to bundle certain facts not easily bundled into other rational addiction theories.

Using terminology from a recent article from outside of economics by Leplin (1986), we can phrase our point differently: Economists are not methodological realists when it comes to their as-if theories – this kind of work is not guided by a belief that unobservable entities such as preferences and beliefs and choice rule refer to something in the real world.

I claim that at the level of methodology, if not epistemology, a distinction must be drawn among theoretical entities between those treated realistically and those proposed as mere possibilities to be exploited for their heuristic and explanatory potential. This means, once again, that questions about the existence and properties of entities of the former sort are regarded as substantive questions which continuing research is directed at answering; our interest in the latter is directed at the empirical adequacy and heuristic utility of theories in which they appear. (p.49)

One piece of supporting evidence from this claim is the experience of economist (or psychoeconomist) Alan Nelson, whose research program, according to Rosenberg, takes the study of choices seriously by actually specifying and trying to operationalise such variables as beliefs and desires.

Nelson asks why traditional economists find the research program of psychoeconomics wholly without interest. [...] Given the evident possibility of the line of inquiry Nelson calls psychoeconomics, it is puzzling to him that the ruling orthodoxy in economics rules out any contribution from psychology. He wonders what it is about this line of inquiry that economists find so repellent. (Rosenberg (1992:134)

Our answer to this question is that economists realise– at some level – that this is misguided, that this is not what they want to do, that this is not the point of their as-if theories. While Rosenberg (1992) is fully aware of the possibility and coherence of positions such as ours, he believes such an answer “incurs considerable costs, costs so great that few economists are willing to pay them” (p. 160). The reason is that such theories no longer provide explanations, and that they become useless for welfare analysis. We agree with this, but would argue that the choice is simple: False but

empirically relevant “theories” useful as a building block in further theories on the one hand, and absurd, poorly justified and unfalsifiable theories of formalised folk psychology taken to extremes on the other.

By deciding to fully step over to our position, economists would also avoid the common retreat they sometimes end up making to the “official” textbook versions of utility functions when someone presses them. “We take choice-patterns as given,” the economists respond, “we just represent whatever observable patterns there are using our formal framework.” Yet, in practise, they soon go back to discussing issues such as whether it can be “rational” to vote since no single individual has any influence except in rare cases. If our view of as-if theory is accepted, this kind of retreat to sterile formalism followed by slinking back into folk psychology and rational choice could be seen as the result of economists not having had any clear, good justification for common practises. They have not been able to take a stand and say that ridiculous as-if theories are perfectly sensible given their aims. They have not been able to give up the silly hope that they can achieve true knowledge about the inner workings and welfare of their fellow humans by setting up a system of equations supported by loose, motivating stories and solving for max of U. Hopefully, this may now change.

To summarise: It is possible to exploit our knowledge of human cognition to bundle observable, stable patterns (the existence of which it seems hard to doubt with any strong degree of faith) inside mock-explanations – absurd theories that “explain” these as sensible, purposeful behaviour – thus enabling us to summarise, communicate and reason with the facts.

2.6 *Remembering that as-if theories are not really theories*

While our defence of absurd theories defends much of what economists do, it does not imply that anything goes: One important fact that follows from the above is that we need to remember and take seriously the fact that our rational choice models are tools, summarising fictions, toy models or, quite simply, just *another way of stating what we see*. Any part of the story not related directly to the observable relationships must be remembered for being just this. We are confused if we start believing that we have *shown* in any realist sense that it is *true* that people *actually are* rational in the sense described by our as-if theory when we observe that real people act in the way our toy models “predict.” We have created a memorable fiction

to summarise empirical relationships, and if competently done this fiction will of course summarise empirical relationships since it *is* nothing but a summary of empirical relationships.

The problem is that there are no clear reminders of this fact, no methodological principles to stop economists from taking the rational choice theories literally. Since as-if theories exploit folk psychology and surround their mathematical formulation with stories that provide “intuition” and feelings of “insight” and “understanding” that tend to give people a mistaken and misguided faith in the truth of these theories, this is a very real danger. The result is misunderstandings both within and outside the profession.

An interesting (and quite innocent) example is Schoemaker (1982), who (in the context of Expected Utility (EU) theory - discusses ideas quite similar to our version of “as-if” methodology under the title postdictive theory, where the

essential premise of the postdictive EU view is that all *observed* human behavior is optimal (in the EU sense), *provided* it is modeled in the appropriate manner. Seeming suboptimalities are explained, ex post facto, by introducing new considerations (e.g. costs, dimensions, constraints, etc.) that account for the anomalies so as to make them optimal. (p. 539)

Schoemaker even notes that an advantage of optimality principles is that they allow a “parsimonious summary of empirical knowledge” (p. 540), yet he differs from us in stating that “A major limitation of the postdictive view is that ex post empirical models may have limited refutation power regarding the corresponding theoretical ex ante model” (p. 540). In other words, he sees what we would call as-if theories as attempts to find *true* theories, and thus sees it as a problem when economists do not try to falsify their theories but instead go on a verificationist search for confirming evidence. From our perspective, as-if theories *should* be “tested” in a verificationist manner, since this involves searching for and identifying *as many empirical facts as we can possibly find to summarise with one and the same as-if theory*.

From our perspective, too, it is not at all puzzling why plainly false assumptions allow us to make true statements about the world: The plainly false assumptions have been carefully chosen to yield the conclusions that we knew – before constructing the theory – were true.

3 Mechanisms

As-if theories do not exhaust the set of all economic theories, and economics would be a poor science if it did. A second type of theories aims to uncover and describe causal mechanisms. In this case, it is the *true* explanation underlying a phenomenon we are interested in, and this explanation will often consist of deductions from assumptions that are well-established approximations. Philosopher Daniel Hausman has labelled such bottom-up theorising deductivism, and argues that it is a valid and useful method of analysis for a social science given the complexity of social phenomena. Whereas Hausman (1989) seems to view this kind of theory as incompatible with the instrumentalist or “predictionistic” as-if method, we want to argue that the two are complementary. In this, we build on Woodward (2002), whose abstract neatly summarises the main point:

Mechanisms consist of parts, the behavior of which conforms to generalizations that are invariant under interventions, and which are modular in the sense that it is possible in principle to change the behavior of one part independently of the others.

Using this view of explanations from mechanisms we can sketch out a “hierarchy” of empirical work, starting with the establishment of empirical facts (E.g. when the price of a good increases, consumers buy less of it). As-if theories bundle such empirical facts together into a mock-explanation, and the as-if theories (rational consumers, profit maximising firms, etc.) are used as building blocks – modules – in theories that relate these to each other in such a way that they can reproduce the phenomenon we wish to explain. In Woodward’s words

My suggestion is that at least part of what it means to say that we have identified the mechanism responsible [for a phenomenon] and that we have correctly segregated that mechanism into components is that we have exhibited [the phenomenon explained] as the consequence of components that are independently changeable in the way just described. (p.S375)

In our words: We can only explain Y in terms of something else, e.g. X, Z, ... These explanatory factors must then be taken as given. Since our mechanism should explain something in the world, the building blocks need to refer to the world and to capture something we have reason to believe in, either because it is “a priori” true (if

such a thing exists) or because it is well supported by evidence. They have to be stable, in the sense that they represent something we know and believe will also be the case in the future (at least within certain limits). When the summarised facts refer to the responses of firms and people, we will often find that a simple way of summarising them is to summarise them in the form of a rational choice or profit maximisation story. We then have observable facts bundled into absurd theories used as building blocks of a mechanism intended to capture an actual process operating in society. Such a strategy is OK *as long as we do not start deriving the implications of the non-factual parts of our as-if models*. In other words: We must be careful not to treat our armchair assumptions of preferences, expectations, choice rule, information, etc. as true of real world individuals.

It is necessary to emphasise this point repeatedly, since it is violated quite often in practise. At the same time, it is simple enough that it seems difficult to deny: A very basic empiricist view is that we need some form of empirical evidence to support claims about the real world. As-if theories are created by seeing what real people do and then tinkering with the preferences, information, beliefs, etc. of a flexible model until our constructed rational individual “acts like” a real person in some regard. Unless we have evidence that people *are* rational, and *do* possess the *ad-hoc assumed* preferences, information, etc., then there is *no reason to trust the implications of these assumptions*. This should be especially clear if we recall that rationality is empirically empty, so that we would have been able to make a rational choice theory for *any conceivable behaviour*. To use a simple reductio ad absurdum: “My rational toy person, constructed ad-hoc by me in my office based on rough generalisations and bizarre claims about human motivation and reasoning, would receive a higher utility if allowed to maim himself with knives. Ergo, I should not be disturbed by the suicidal behaviour of my son.”

Note also that the above discussion of mechanisms could support a wide variety of mechanism-theories: Some economists might create mechanisms in verbal theories because their assumptions would be poorly captured by formal models, others might create formal models in order to clarify their mechanism, check their reasoning and get their point across – using simple and clearly wrong assumptions intended to be *qualitatively* representative – and yet others might believe that there are stable, quantitatively specifiable relationships that can be empirically estimated using

structural modelling or calibration techniques, to be used as components in larger models intended for precise predictions.

4 Pure theory

Our explanation of as-if theory relies on there being something known about how humans and firms behave that we can summarise. This leaves open the question of how “pure” theory with no known empirical applications, such as General Equilibrium theory or complex game theoretical solution concepts, should be understood. We have no single answer to this question, only some brief suggestions:

Some might be seen as stocking up the profession with further *as-if* tales to be used for summarising as-yet-not-identified empirical facts.

Some might be attempts to summarise facts that are believed by the economist involved to exist, but not yet documented in a manner that persuades others.

Some might be confused attempts to correctly specify the rational choice problem faced by real people in some context and derive the “true” explanation, the “normatively” correct thing people “should” do, or the actual “welfare effects” of policies. One might, of course, believe that people act in their best interest, or that they *tend* to do so, that they *learn* to do so over time, that *most* of them do so, that they *usually* do so, that they do so *except in certain identified contexts*, etc. Or one might believe that we have no feasible alternative preferable to consumer sovereignty. There might be valid arguments and evidence to support such beliefs, our point here is completely neutral towards alternative views on the issue. We restrict ourselves to noting that none of these claims can be either established or falsified by ad-hoc theoretical models fitted to empirical observations in a flexible framework that could generate any conceivable behaviour.

Other such theories might be a form of mathematical fiction, enjoyed by some but not all economists, presenting stories of conceptual people in strange hypothetical worlds who act – in certain aspects – similarly to real people. Some economists might find such work aesthetically pleasing and preferable to reality with its poor data, inference problems, etc.

Though no exhaustive taxonomy of possible views of pure theory, the above suggestions show that such work may be seen as valuable even though it has no apparent connection to reality.

5 Summary and conclusion

It is tempting to see the above as an as-if theory of as-if theory, where as-if theory is explained as the rational response to cognitive limitations in humans, a way to bundle facts and hypotheses in a form that exploits folk psychology. In addition, our “theory” of as-if theory seems useful, in that it makes many common practises in economics seem meaningful and sensible, reminds us of obvious limits such theorising has, and clarifies when economics becomes truly absurd and how to avoid it (e.g. when it starts taking welfare-analysis based on as-if theories seriously). Ideally, there should be something in the name of our theoretical representations of individuals and firms that reminded us that these as-if theoretical constructs were not meant to be taken literally, e.g. utility-robots, proxy-people, math-men. Since economic terminology is well-established, it seems futile to try and change it. Hopefully, providing a story to help us remember it may be a good second-best. To summarise our article as a slogan: A good as-if theory is no theory at all, just a shorthand summary of facts. We assume that consumers are rational, self-interested super-computers because we take seriously the fact that we ourselves are not.

References

- Backhouse, R. E. (1999). "Book review: Models and Reality in Economics. By Steven Rappaport." The Manchester School **67**(1): 122-124.
- Boland, L. A. (1981). "On the Futility of Criticizing the Neoclassical Maximization Hypothesis." American Economic Review **71**(5): 1031-1036.
- Broome, J. (1999). Ethics out of Economics. Cambridge.
- Buchanan, M. (2002). Nexus. New York, London, W.W. Norton & Company, Inc.
- Dancy, R. M. (1978). "Model Behavior." The Journal of Philosophy **75**(11): 677-679.
- Davies, J. B. (1997). The myth of addiction. Amsterdam., Harwood Academic Press.
- Dawes, R. M. (1996). House of Cards. New York, London, Toronto, Sydney, Singapore, The Free Press.
- Dawes, R. M. (1999). "A message from psychologists to economists: mere predictability doesn't matter like it should (without a good story appended to it)." Journal of Economic Behavior and Organization **39**: 29-40.
- Dawkins, R. (1999). The Extended Phenotype. The Long Reach of the Gene. Oxford New York, Oxford University Press.
- Friedman, M. (1953). The Methodology of Positive Economics. Essays in Positive Economics. Chicago and London, The University of Chicago Press: 3-43.
- Friedman, M. (1992). "Do Old Fallacies Ever Die?" Journal of Economic Literature **30**(4): 2129-2132.
- Gibbard, A. and H. R. Varian (1978). "Economic Models." The Journal of Philosophy **75**(11): 664-677.
- Gigerenzer, G. (2000). Adaptive Thinking. New York, Oxford University Press.
- Hausman, D. M. (1989). "Economic Methodology in a Nutshell." Journal of Economic Perspectives **3**(2): 115-127.
- Hausman, D. M. (1992). The inexact and separate science of economics. Cambridge., Cambridge University Press.
- Hirsch, A. and N. De Marchi (1984). "Methodology: A Comment on Frazer and Boland, I." The American Economic Review **74**(4): 782-788.
- Hutchison, T. W. (1938). The Significance and Basic Postulates of Economic Theory. London, MacMillan and Co., Limited.
- Laibson, D. (2001). "A Cue-Theory of Consumption." Quarterly Journal of Economics **116**(1): 81-119.
- Leplin, J. (1986). "Methodological Realism and Scientific Rationality." Philosophy of Science **53**(1): 31-51.
- Machina, M. J. (1999). "A Challenge To The "Econoclasts": A Commentary on "Rationality for Economists?"" Journal of Risk and Uncertainty **19**(1-3): 107-108.
- McFadden, D. (1999). "Rationality for Economists?" Journal of Risk and Uncertainty **19**(1-3): 73-105.
- Norman, D. A. (2002). The Design of Everyday Things. New York, Basic Books.
- Pinker, S. (1999). How the Mind Works. New York, London, W.W. Norton & Company.
- Rogeberg, O. (2003). "Preferences, Rationality and Welfare in Becker's Extended Utility Approach." Rationality and Society **15**(3): 283-323.
- Rosenberg, A. (1978). "The Puzzle of Economic Modeling." The Journal of Philosophy **75**(11): 679-683.
- Rosenberg, A. (1992). Economics - Mathematical Politics or Science of Diminishing Returns. Chicago and London, The University of Chicago Press.

- Russel, B. (1913). "On the Notion of Causes." Proceedings of the Aristotelian Society **13**.
- Schoemaker, P. J. H. (1982). "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations." Journal of Economic Literature **20**(2): 529-563.
- Sen, A. (1982). Rational Fools: A Critique of the Behavioural Foundations of Economical Theory. Choice, Welfare and Measurement, Blackwell: 84-106.
- Sen, A. (1993). "Internal Consistency of Choice." Econometrica **61**(3): 495-521.
- Simon, H. A. (1979). "Rational Decision making in Business Organizations." The American Economic Review **69**(4): 493-513.
- Stanley, T. D. (1985). "Positive Economics and its Instrumental Defence." Economica **52**: 304-319.
- Stigler, G. J. and G. S. Becker (1977). "De Gustibus Non Est Disputandum." American Economic Review **67**(2): 76-90.
- Trout, J. D. (2002). "Scientific Explanation and the Sense of Understanding." Philosophy of Science **69**: 212-233.
- Woodward, J. (2002). "What Is a Mechanism? A counterfactual Account." Philosophy of Science **69**: S366-S377.