EXCESSIVE AMBITIONS[*]

by

Jon Elster

I. Introduction

Amos Tversky once told me about a meeting he had attended with the foremost psychological scholars in the U.S., including Leon Festinger. At one point they were all asked to identify what they saw as the most important current problem in psychology. Festinger's answer was: "Excessive ambitions". In this paper I argue that this is not just the case for psychology, but for the social sciences across the board.[1] I exclude only

---

[*] I am grateful to Chris Achen, George Akerlof, Amar Bhidé, Olivier Blanchard, Ernst Fehr, Dagfinn Føllesdal, Daniel Kahneman, David Laibson, Karl Ove Moene, Pasquale Pasquino, John Roemer, Ariel Rubinstein, Robert Shiller and Gabriele Veneziano for comments on earlier drafts of this paper. I have also benefited from many discussions with Nassim Nicholas Taleb, who argues from somewhat different premises to similar conclusions (Taleb 2005, 2007).

[1] For earlier arguments along the same lines see Elster (2000) and Elster (2007), notably the Introduction and the Conclusion to the latter book. Although some of the arguments that I make are similar to criticisms of mainstream economics inspired by the current financial crisis (see notably Akerlof and Shiller 2009), they were developed well before August 2007.

anthropology, in which the level of ambition often seems *too low*, after it embraced postmodern theory, postcolonial theory, subaltern theory, deconstructionism and the other usual culprits. As I shall paint with broad strokes, some exceptions to my claims will be left out. As my competence in the various social sciences is highly uneven, some of my claims will be based on a deeper understanding of the literature than others. At one point I even commit what is normally considered a deadly sin in scholarship, that of criticizing others on the basis of third-party authorities.

The paper, therefore, is a risky venture, that of offering "outsider criticism". It is unlikely, for instance, that the community of mathematical economists will take radical objections to certain forms of mathematical economics seriously unless they are made by one among themselves. Scholars such as Ariel Rubinstein, Matthew Rabin or Roman Frydman have the credentials that will make insiders listen. As I lack these credentials, there is a risk that my criticism will be dismissed as obscurantist, on a line with the objections made some years ago by the "Perestroika movement" in political science. As it turned my criticism of the mainstream made it difficult to dissociate myself from that movement. Those who fight a two-front war, in my case against soft and hard forms of obscurantism, run the risk that each camp will see them as belonging to the other.

Yet I think the problems are important enough to justify any possible risk to my reputation. As I see it, excessive ambitions cause both *waste* and *harm*. A mind is indeed a terrible thing to waste, and the waste can occur by hypertrophy and atrophy as well as by not developing at all. Cohorts after cohort of students are learning – and many of them subsequently hired to apply or teach – useless theories. Their efforts and talents would have been vastly more useful to society had they been harnessed to more productive purposes.[2] (Needless to say, this comment applies in spades to students in the "soft" humanities, including large parts of anthropology.) Conjecturally, a redirection of their effort might even benefit the concerned individuals themselves.

Society can and indeed should tolerate some waste in the scholarly community. Research, it is often said, is a high-risk activity. Some scholars might by sheer bad luck never strike a rich mine. Others might lose their motivation for research once they are granted tenure. These are inevitably by-products of a system that may otherwise work well. The kinds of waste I have in mind, however, should not be tolerated. For reasons I speculate about below they probably will, though, for the foreseeable future.

Harm is a more serious matter. As one example, investors lost more than 4 billion dollars as a result of the failure of Long Term Capital Management, a hedge fund based on very fragile economic models. This was, of course, only the opening act of a drama that is still unfolding at the time of writing. Intellectual hubris of modelers was surely one of the causes of the subprime crisis and the further collapses it has triggered. The defense of the death penalty on the basis of very controversial statistical arguments, according to which every execution of a murderer prevents several murders from being committed, is another instance of a potentially disastrous overreliance on poorly understood models. (The weakness of these arguments is not of course the only reason to oppose the death penalty.)

In a different realm, the use of psychological "expert" witnesses in child abuse cases has caused large amounts of unjustified suffering, exposed

---

[2] On these lines, see the Letter to the Editor by James Mitchell in *The Economist* for October 11-17 2008: « Imagine what these young people [who were lured into the banking industry] could have done if they had chosen careers in science and medicine ».

among others by Elizabeth Loftus and Robyn Dawes. In a Norwegian case known to me, a conviction resulted because a psychologist testified that the sharp fence posts in a child's drawing of a house surrounded by a fence very likely had a sexual significance. The acquittal on appeal could not undo the harm. Earlier, a generation of psychodynamically trained psychologists caused great harm by telling parents of autistic children the falsehood that the condition was caused by their bad parenting behavior.

One fundamental cause of these disturbing phenomena may be our unwillingness to admit ignorance, and, rather than grasping for knowledge, try to do as well we can given that we are ignorant. Albert Hirschman has said that most Latin American cultures "place considerable value on having *strong opinions* on virtually *everything* from the *outset*".[3] In such societies, to admit ignorance is to admit defeat. But the phenomenon is really much more general. Montaigne said that "Many of this world's abuses are engendered - or to put it more rashly, all of this world's abuses are engendered - by our being schooled to be afraid to admit our ignorance and because we are required to accept anything which we cannot refute."[4] The mind abhors a vacuum.[5]

In developing and illustrating some of these claims I shall proceed as follows. In Section **II** I consider the rational-choice paradigm that has a dominant status in economics and political science and to a smaller extent in sociology. (I use "dominant" in a sociological sense that I explain in Section **V**.) In Section **III** I discuss whether the "behavioral economics revolution"

---

[3] Hirschmann (1986).

[4] Montaigne (1991), p. 1165.

[5] The idea of a need for *cognitive closure* (e.g. Kruglansky and Webster 1986) points in the same direction. The classic study by Neurath (1913) is still worth reading.

can be said to offer a more attractive alternative to rational-choice models. In Section **IV** I address pitfalls and fallacies in statistical data analysis. Since this is an area in which my own competence is particularly weak, I rely on other scholars who combine a high reputation among their peers with deep skepticism of the ways in which statistical analysis is routinely applied. In Section **V** I speculate somewhat inconclusively about the causal mechanisms that sustain the reproduction of these pathologies. I conclude in Section **VI** by sketching a more modest but also, I believe, more robust approach to explanatory efforts in the social sciences.

II. The problems with rational choice theory

There is no doubt in my mind that rational-choice theory in general, and game theory in particular, produced the greatest intellectual revolution in the social sciences since their beginnings. Before rational-choice theory, there were no intellectual tools available to help us understand how people make tradeoffs among the different dimensions of the choices they face. By explaining the consumer's choice in terms of a budget set and convex indifference curves, and the producer's choice by an analogous model, it was possible, for instance, to break out of the sterile structuralism of Marx. Before Schelling, it was hard to make sense of the idea that an agent might rationally decide to destroy some of her assets (e.g. burning her ships). Today, the idea is routinely applied to the analysis of industrial organization.

Or consider the idea, which will be put to use in Section **V**, of a *bad equilibrium*. It has probably always been vaguely understood that societies could be stuck in needlessly bad states, but harder to understand why these persist. Why, for instance, do many languages maintain the discrepancy between written and spoken language that is often the despair of schoolchildren? One reason is probably that among users of the written language a unilateral attempt to move closer to the spoken language would impede their ability to communicate with other users, not because the latter would not understand them, but because the unusual form would distract them from the substance to be communicated.

The ideas of burning one's bridges or of a bad equilibrium are simple and robust. They can be used to explain a good deal of behavior. To my mind, however, the main value of the bulk of rational-choice theory is conceptual rather than explanatory. For the theory to explain behavior, two assumptions must hold. First, the theory must give a determinate prediction in the case at hand. (In the case of the consumer's choice, this assumption holds because the convexity of the indifference curves guarantees a *unique* tangency point with the budget line.) Second, the observed behavior of the agent or agents we are considering must conform to the predictions of the theory. If either of these assumptions fails to hold – if *the theory is indeterminate* or *the agents are irrational* - no explanation will be forthcoming. It would be meaningless to make a quantitative statement about how often both assumptions hold. I feel confident, however, in asserting that in many important cases they do not hold.

Let me begin with the issue of indeterminacy. I shall single out two issues: belief formation and decision-making. Since decisions rest on beliefs, they cannot in general be determinate unless the beliefs are. Yet even with determinate beliefs, decisions may be indeterminate.

The question of the indeterminacy of rational beliefs can also be stated as the question of *uncertainty*, in the sense of Knight or Keynes. In some cases, agents may be unable to assess numerical probabilities to the possible outcomes of actions.[6] Cases include brute uncertainty, information-gathering uncertainty, and strategic uncertainty.

Brute uncertainty linked to “fat tails” is a prominent issue in the literature on climate change.[7] I shall use a simpler example, linked to the use of the principle of insufficient reason to resolve uncertainty into risk. Generally speaking, appeal to a uniform distribution without specific evidence is unwarranted, unless one can conduct Bayesian updating that will swamp the initial vagueness. Such appeals are often, made, however, without this justification. Moreover, one has to ask the question: uniform distribution of what? “Even the structure of parameter space is a subjective choice and has a first-order effect […]. Physically, we can equally well use a parameter labeled ‘ice fall rate in clouds’ or its inverse (‘ice residence time in clouds’) and achieve identical simulations. Sampling uniform distributions under each of the two different labels however, yields completely different results.”[8] Finally, Bayesian updating will not work if one is aiming at a moving target. In the Vietnam War, for instance, updating estimates about enemy strength on the basis of enemy sightings would be meaningless if enemy forces are waxing or waning. Although I have no evidence that military decisions in that war were in fact based on Bayesian updating, I believe that the assumption of an unchanging (but unknown) state of the world is often used without sufficient justification.

---

[6] I ignore the even more intractable case of *ignorance*, or “unknown unknowns”, in which we may be unaware of some of the possible outcomes. I suspect, however, that ignorance will turn out to be a major issue in climate change.

[7] Weitzman (2009).

[8] Stainforth et al. (2007.

Uncertainty generated by the unknown and rationally unknowable costs of information-gathering is virtually neglected in the rational-choice literature. An important exception is Sidney Winter, who observed that the idea of reducing satisficing to a form of maximizing creates an infinite regress, since "the choice of a profit-maximizing information structure itself requires information, and it is not apparent how the aspiring profit maximizer acquires this information or what guarantees that he does not pay an excessive price for it."[9] Along the same lines, Leif Johansen characterized the search process as "like going into a big forest to pick mushrooms. One may explore the possibilities in a certain limited region, but at some point one must stop the explorations and start picking because further explorations as to the possibility of finding more and better mushrooms by walking a little bit further would defeat the purpose of the hike. One must decide to stop the explorations on an intuitive basis, i.e. without actually investigating whether further exploration would have yielded better results".[10] When rational belief formation is indeterminate, one does indeed have to rely on intuition. Even assuming that one can *predict* the outcome of intuition as based on heuristics and biases, that prediction does not aspire to the *normative* force of a prediction based on rational-choice theory.

Strategic uncertainty arises when agents have to form beliefs about one other, including beliefs about beliefs etc. In theory, one can short-circuit the looming infinite regress by the notion of an equilibrium set of strategies. These often involve mixed strategies that are only *weakly* optimal against each other, in the sense that an agent can do just as well by adopting one of

[9] Winter (1964), p. 252.

[10] Johansen (1977), p. 144.

the strategies in the mix as her pure strategy. In that case, however, why should an agent believe that the others are playing the mixed equilibrium strategy? Would it not be rational to play it safe and adopt a maximin pure strategy?

One may also try to justify the idea of mixed strategies by appealing to a causal mechanism, as in the following argument put forward to explain the passivity of bystanders in the Kitty Genovese case: "[M]ixed strategies are quite appealing in this context. The people are isolated, and each is trying to guess what others will do. Each is thinking, Perhaps I should call the police … but maybe someone else does … but what if they don't? Each breaks off this process at some point and does the last thing that he thought of in this chain, but we have no good way of predicting what that last thing is. A mixed strategy carries the flavor of this idea of a chain of guesswork being broken by a random point."[11] So far, so good. The authors then go on, however, to commit a simple quantifier fallacy: from the correct premise that for every person there is a probability **p** that he will not act, they reach the false conclusion that there is a **p** such that each person will abstain from acting with probability **p**. Moreover – a second unjustified step – they assume that when all abstain from acting with probability **p**, their choices will form an equilibrium.

In this example, the probabilities completely *lack microfoundations*. The authors give no reason why all subjects should come up with the particular probability that has the property of generating an equilibrium. The number is top-down, invented to close the system, not bottom-up. Along similar lines, Roman Frydman and Michael Goldberg argue that rational-

[11] Dixit and Skeath (2004), p. 416.

expectations macroeconomics is lacking microfoundations.[12] Frydman (personal communication) suggests a striking analogy between rational-expectations economics and the Marxist theory of the ideological superstructure in society. In neither case are the beliefs imputed to social agents supposed to stem from the information available to them, but are simply stipulated to close the system.

Let us assume, however, that people form rational and determinate beliefs and ask how these can enter into the explanation of their *decisions*. For this purpose one has to (i) identify the objective function of the agent and (ii) show that they possess the cognitive capacities to maximize it. Although the second problem is the more serious, I begin with a few comments on the first.

When economists try to estimate the utility function from observed behavior, they often assume a certain functional form for which they estimate the parameters. This seems arbitrary. If they also try out several function forms, they are pretty sure to find one that fits the data. This is curve-fitting, to be accepted *only* if the hypothesized utility function is used to generate other predictions, preferably in the form of "novel facts", over and above those it is supposed to explain.[13]

Some economists are of course aware of this.[14] Others, however, simply assume a utility function, without attempting to justify it or verifying that the results generalize to other functions. Theodore Bergstrom showed, for

---

[12] Frydman and Goldberg (2007).

[13] Alternatively, they could try to prove their results assuming only that the first derivate of the utility function is positive and the second native. For most purposes, however, this assumption may be too weak to generate interesting results.

[14] See for instance Chiappori (1990) for an unusually thoughtful discussion.

instance, that Gary Becker's "rotten-kid theorem" holds only for a restricted set of utility functions.[15] Often, they assume "for simplicity" that utility is linear in money, or separable in its arguments, without telling the reader how many of the conclusions can be expected to hold in the non-simplistic case. *Very generally*, they assume that time discounting is exponential, without addressing the large body of evidence suggesting that people discount the future hyperbolically. Ease of calculation seems to have been the dominant reason for maintaining this demonstrably false idea. Tellingly, hyperbolic discounting did not make much of an inroad in the literature until one substituted quasi-hyperbolic discounting, which has the two advantages of offering computational ease and being a reasonably good approximation to hyperbolic discounting.[16]

Let us assume, however, that in addition to determining what in a given situation would constitute rational beliefs we are able to identify the utility function and the time preferences of the agent or agents in question. Is it possible to derive a unique behavioral prediction? The reason why the answer will often have to be negative is that the framework of rational-choice theory, game theory and decision theory is too narrow. They focus exclusively on the *preferences* and *beliefs* of the agents, while ignoring their *capacities*.

The point I am about to make is embarrassingly simple. I cannot help believing that practitioners of rational-choice theory are aware of it "at some level", as the phrase goes, but that they manage to ignore it most of the time.

[15] Becker (1974); Bergstrom (1989).

[16] By a curious twist, evidence from brain imaging suggests that the quasi-hyperbolic model may actually be a better representation of how the mind works than the hyperbolic form (McClure et al. 2004). But see some qualifications in note 23 below.

(Pascal said that "Ordinary people have the power of not thinking of that about which they do not wish to think". In this respect, economists and political scientists are probably like ordinary people.) The point is this: *how can one impute to the social agents the capacity to make the calculations that occupy many pages of mathematical appendixes in the leading journals* and that can be acquired only through years of professional training? Why should we believe in "as-if" rationality?

I shall discuss *four possible answers* to these rhetorical questions. The first – which to my knowledge is virtually never proposed – is to invoke the precedents of Newton's law of gravitation and of quantum mechanics. Early critics of Newton objected to the law of gravitation that it presupposed the metaphysically absurd notion of action at a distance. Eventually, however, everybody accepted the theory because *it worked*, with an amazing degree of precision. The even more incomprehensible theory of quantum mechanics, which involves not only action at a distance but objective indeterminacy, is also accepted because its predictions are verified with nine-decimal accuracy. Similarly, in spite of the general objections to rational-choice theory that I have proposed, one might be willing to accept it if its predictions were verified with comparable many-decimal precision. However, anyone with the slightest acquaintance with economics or political science will dismiss the idea as laughable. Often, scholars are happy if they "get the sign right".

The second and most frequent defense of the explanatory relevance of rational-choice theory would appeal to a causal mechanism capable of *simulating rationality*. Just as economists are fond of arguing that self-interest can simulate altruism, they often claim that non-intentional mechanisms can simulate intentional optimizing. These mechanisms will generate behavior with (say) utility-maximizing *consequences* even though the agents are incapable of deriving it from utility-maximizing *intentions*. Generally speaking, there are two mechanisms that might be capable of this feat: *reinforcement* and *selection*.[17] The former works by causing given behavioral units to optimize, the latter by eliminating non-optimizing units. As defenders of rational-choice theory rarely if ever appeal to reinforcement, and since the mechanism doesn't simulate optimality very well in any case[18], I shall ignore it.

The only relevant selection mechanism is social or economic selection. Natural selection has of course produced the kind of rough-and-ready and cognitively undemanding rationality that serves us well in everyday life. As an example, consider the Norwegian proverb: "Don't cross the river to the other bank when you go to fetch water". An organism that engaged in such wasteful behavior would be quickly eliminated. There is no reason to believe, however, that natural selection could produce the highly sophisticated strategic behaviors that the models predict. Evolutionary game theory may have some uses, but that of sustaining the models is not one of them.

Models of "economic natural selection" do have some empirical relevance. The writings of Richard Nelson and Sidney Winter, in particular, shed some qualitative light on economic development.[19] Yet they do not provide the sought-for simulation of rationality, for several reasons. First, as

---

[17] Skinner (1981).

[18] Herrnstein and Prelec (1992).

[19] Nelson and Winter (1982). A largely neglected article by Nelson, Winter and Schuette (1976) remains worth reading for the discussion of methodological issues related to those I address in Section **V** below.

with simulations and agent-based modeling in general, it is often hard to know the extent to which the results are artifacts of the assumptions. Second, and more important, these results do not show optimizing behavior. In a population of firms that evolve by innovation and imitation there is always a substantial proportion of non-optimizing firms. Since firms are adapting to a rapidly changing environment, they are (as in some cases of Bayesian updating) aiming at a moving target. In any case, there is no hope whatsoever that the simulations could mimic the models *all the way down to the mathematical appendices*. Third, and even more important, bankruptcy-driven or takeover-driven elimination of inefficient agents could never generate optimizing behavior in non-market societies or in non-market sectors in market societies. I conclude that *appeal to selection is pure hand-waving*.

A third possible defense (suggested to me by a reader of a draft of the present article) is that although boundedly rational agents are liable to make *mistakes*, these will cancel each other out in the aggregate. If we required each person in a group to carry out calculations of the order of difficulty, say, of multiplying 49 and 73 in at most 30 seconds, we would expect there to be some mistakes, but also that these would be normally distributed around the correct answer. For some purposes, this fact might justify the rationality assumption. When, however, the correct answer requires solving differential equations or carrying out other complicated operations, *there is no reason to expect answers or guesses to be normally distributed around the correct answer*. It seems to me that the burden of proof is on those who might claim that they will.

Finally, the often-cited example offered by Milton Friedman of the expert billiard player whose experience helps him figure out the angles, even though he would be utterly incapable of solving the relevant equations, suggests a fourth possibility, that of learning by trial and error.[20] To my knowledge, there have been few attempts to transform this metaphor into a theory. One article concludes that "individual learning methods can reliably identify reasonable search rules only if the consumer is able to spend absurdly large amounts of time searching for a good rule".[21] This conclusion may be related to the fact that *people cannot be experts across the board*.

Rational-choice models may be of interest on three distinct grounds. First, they may help us explain, predict or shape behavior. Although simple and robust models may do this in a rough-and-ready sense, the sophisticated models that are the pride of the profession do not. Second, they may have an aesthetic appeal. It can be intrinsically satisfying to figure out the actions and interactions of ideally rational agents who have never existed and never will. Refining the equilibrium concept in game theory is an example. Third, the models may some have mathematical value or spur mathematical investigations. Googling, I found links for instance between the work of the Field medal winner Pierre-Louis Lions and work in demand theory.

My claim is that much work in economics and political science is devoid of empirical, aesthetic or mathematical interest, which means that it has *no value at all*. I cannot make any quantitative assessment of the proportion of work in leading journals that fall in this category. I am firmly convinced, however, that the proportion is non-negligible and important enough to constitute something of a scandal. I also believe, more tentatively,

---

[20] Friedman (1953).

[21] Allen and Carroll (2001).

that the proportion may be higher in the leading journals than in the non-leading ones. I return to that issue in Section **V**.

A reader of an earlier draft of this paper commented that it is unfair or superficial to make these critical claims without citing and criticizing specific instances, in the way David Freedman did for the abuse of statistics (see Section **IV**). Although my competence to assess deductive models is certainly less than Freedman's competence to assess statistical models, I believe I could make a convincing argument that this or that article by an eminent economist published in a leading journal is nothing more than a piece of science fiction. I would only have to exhibit the assumptions and the deductive apparatus and point out their obvious lack of realism. The reason I do not believe I have to it is simply that I do not think any competent economist would actually contest the point that the models lack realism. Defenders of the theory will, I imagine, resort to as-if justifications rather than claim that the models are literally correct, or approximately so. I have argued that the as-if version of theory lacks solid foundations.

III. Is behavioral economics the solution?

In addition to the problem of indeterminacy, rational-choice theory faces the problem of *irrationality*. People do not behave as the theory says they will or should. Rational-choice theory seems incapable, for instance of explaining voting, addiction, precommitment, revenge, self-deception, and many other observed phenomena.

The response to these anomalies has been twofold. On the one hand, economists have tried to show that these behaviors are, in fact, rational. Despite some successes, most of the attempts have been tortuous, simplistic, tautological, or otherwise flawed. On the other hand, a common response to the anomalies has been to say that "You can't beat something with nothing". Although the existence of irrational behavior was not denied, it was for a long time seen as a residual category, not as a positive phenomenon with specific implications. The idea of satisficing as an alternative to maximizing never struck deep roots in the profession because it was largely descriptive, with neither prescriptive nor predictive implications.[22]

This state of affairs changed in 1974-1975, with the publication of the first major article by Daniel Kahneman and Amos Tversky on choice under uncertainty and George Ainslie's resurrection of R. H. Strotz's theory of hyperbolic time discounting.[23] In the 35 years that followed, the research program of behavioral/experimental economics has unearthed a vast number of positive mechanisms generating "predictably irrational" behavior.[24] The program, in other words, has predictive but not prescriptive implications.[25]

Although it would be impossible to attempt a complete statement of the irrationality-generating mechanisms, I shall try to produce a representative shortlist.[26] If we go by the literature, the two most important ones are

---

[22] In Section **VI** below I argue that in addition to *prescription* and *prediction* we should consider *explanation* (or *retrodiction*) as a separate category.

[23] Kahneman and Tversky (1974) ; Ainslie (1975); Strotz (1956). Let me note here that I do not count mathematical analyses of choice with hyperbolic or quasi-hyperbolic discounting as falling under the heading of behavioral economics. Although these models depart from the rationality assumptions of standard models, they share the – highly unrealistic – assumption of those models concerning the computational capacity of social agents.

[24] I take this phrase from the title of Ariely (2008).

[25] Policy makers are often unhappy with these models. They do not want to assume that the economic agents are irrational, because doing so might lead them to embrace paternalism. I shall not pursue this issue, however.

[26] The following books contain a total of 344 articles detailing the more important mechanisms : Kahneman, Slovic and Tversky (1982), Bell, Raiffa and Tversky (1988), Loewenstein and Elster (1992), Kahneman, Diener and Schwartz (1999), Kahneman and

probably loss aversion and hyperbolic discounting. In my view emotions are at least equally important, although so far they have proved less tractable for experimental purposes. Among other mechanisms the following may be cited:

- the sunk-cost fallacy and the planning fallacy (especially deadly in conjunction)
- the tendency of unusual events to trigger stronger emotional reactions (an implication of "norm theory")
- the cold-hot and hot-cold empathy gaps
- trade-off aversion and ambiguity aversion
- anchoring in the elicitation of beliefs and preferences
- the representativeness and availability heuristics
- the conjunction and disjunction fallacies
- the certainty effect and the pseudo-certainty effect
- choice bracketing, framing, and mental accounting
- cases when "less is more" and "more is less"
- sensitiveness to changes from a reference point rather than to absolute levels
- status quo bias and the importance of default options
- meliorizing rather than maximizing
- motivated reasoning and self-serving biases in judgment
- flaws of expert judgments and of expert predictions
- self-signaling and magical thinking

Tversky (2000), Connolly, Arkes and Hammond (2000), Gilovich, Griffin and Kahneman (2002), Brocas and Carillo (2003, 2004), Camerer, Loewenstein and Rabin (2004), Lichtenstein and Slovic (2006) and Loewenstein (2007). (Because of overlap the number of distinct articles is somewhat smaller; also, not all the articles can be classified as belonging to behavioral economics.) The volumes include but do not do full justice to the impressive cumulative work of Ernst Fehr and his associates. Camerer (2003) provides a full survey of behavioral game theory. Finally, the first textbook of behavioral economics was just published (Wilkinson 2008).

- non-consequentialist and reason-based choice
- overconfidence and the illusion of control
- spurious pattern-finding

I offer this list mainly to underline the fact that unlike rational-choice economics, behavioral economics does not rest on a unified theory. Rather, it consists of a bunch of theories or mechanisms that are not deductively linked among themselves. Human behavior seems to be guided by a number of *unrelated quirks* rather than by consistent maximization of utility. In fact, there are so many quirks that one suspects that for any observed behavior, there would be a quirk that fits it. Many mainstream economists seem to shy away from behavioral economics because they think it invites ad-hoc and ex-post explanations. Whereas I shall defend ex-post explanations, I certainly do not want to defend ad-hoc-ness. I believe, however, that behavioral economics can avoid this danger by imposing standard explanatory requirements.

Also, what triggers one quirk rather than another may be small or seemingly irrelevant variations in experimental protocols or settings. Partly for this reason, perhaps, there are relatively few applications of behavioral economics outside the laboratory. There are, to be sure, some examples. Colin Camerer's survey of "Prospect theory in the wild" is a rare, perhaps unique systematic survey.[27] Linda Babcock, George Loewenstein and their collaborators offer an elegant combination of experiments and field research to demonstrate the importance of self-serving conceptions of fairness and

---

[27] Camerer (2000).

their impact on bargaining failures.[28] Arguably, loss aversion and choice bracketing combine to explain the puzzling fact that stocks historically yield much higher returns than bonds.[29] The existence of Christmas clubs and other real-life precommitment devices are plausibly explained by assuming that people discount future rewards hyperbolically rather than exponentially, *and that they know they do so*.[30] The success of "libertarian paternalism" in causing people to increase their savings is due to a clever exploitation of three mechanisms: loss aversion, hyperbolic discounting, and the tendency to prefer the default option.[31]

These and other examples notwithstanding, it seems fair to say that the successes of behavioral economics, like those of traditional psychology, are mainly found in the laboratory. Behaviors "in the wild" are usually cited as illustrations, not as explananda. One may easily find historical or contemporary cases that are *consistent* with mechanisms such as the sunk-cost fallacy[32], the planning fallacy[33], spurious pattern-finding[34] or magical thinking[35]. It is harder to show that the observed behaviors were in fact *caused by* these mechanisms.

---

[28] Babcock, Wang and Loewenstein (1992), Babcock and Loewenstein (1997).

[29] Benartzi and Thaler (1995). The « myopia » in their title is to be read as choice bracketing (for which see Read, Loewenstein and Rabin 1999).

[30] Thaler and Shefrin (1981). The italicized phrase points to an important qualification of some findings in behavioral economics. If people are capable of dealing rationally with their irrational propensities, they may limit the damage they would otherwise suffer.

[31] Thaler and Sunstein (2008).

[32] Arkes and Blumer (1985).

[33] Buehler, Griffin and Ross (2002).

[34] Feller (1968), p. 160.

[35] Rozin and Nemeroff (2002).

Let me illustrate this difficulty through a study by Daniel Kahneman and Dale Miller on “norm theory”, in which they assert a

> correlation between the perception of abnormality of an event and the intensity of the affective reaction to it, whether the affective reaction be one of regret, horror, or outrage. This correlation can have consequences that violate other rules of justice. An example that attracted international attention a few years ago was the bombing of a synagogue in Paris, in which some people who happened to be walking their dogs near the building were killed in the blast. Condemning the incident, a government official singled out the tragedy of the “innocent passers-by.” The official's embarrassing comment, with its apparent (surely unintended) implication that the other victims were not innocent, merely reflects a general intuition: The death of a person who was not an intended target *is* more poignant than the death of a target.[36]

The statement by the “government official” – it was in fact Raymond Barre, the Prime Minister at the time – is indeed consistent with the proposed explanation in terms of norm theory. It is also, however, consistent with an explanation in terms of an anti-Semitic prejudice. The considerable amount of evidence suggesting that Barre had an anti-Semitic bias includes his strong defense of Maurice Papon and a directive he signed in 1977 (later struck down by the Conseil d’État) that effectively cancelled anti-racist legislation from 1972. Moreover, Barre’s actual comment was somewhat less innocuous than in the paraphrase of Kahneman and Miller. He referred to “the odious attack that intended to strike Jews on the way to the synagogue and that struck innocent French citizens crossing the street”.

[36] Kahneman and Miller (1986), p. 146.

(*l'attentat odieux qui voulait frapper des israélites qui se rendaient à la synagogue et qui a frappé des Français innocents qui traversaient la rue*). In fact, the Jews in question were French too. In my view, this phrasing supports an explanation in terms of anti-Semitism. Although Barre may not have "intended" the implications that the Jewish victims were not innocent and that they were not French, many studies show that prejudice often operates at an unconscious level.[37] The claim that the official's comment "merely reflects a general intuition" *may* be correct, but Kahneman and Miller do not show that it is more plausible than alternative explanations.

It is an open question whether experiments provide sufficient evidence for the claims of behavioral economics, or whether validation outside the laboratory is essential. Some claims against the relevance of experiments are readily dismissed. In particular, the fact that most findings can be replicated with large monetary stakes refutes the objection that the experiments only involve trivial amounts of money. Other objections may be more worrisome:

- In the post-Milgram era, scholars are prohibited from conducting experiments with high-stake *emotional* charges. Extrapolating from behavioral expressions of the positive affect subjects feel when given candy or when discovering that the pay phone already has a coin in it may not be justified.
- At the other end of the emotional spectrum, extrapolations from the behavioral expressions of negative affect generated by unfair behavior in an Ultimatum Game or a Trust Game are not necessarily justified.
- Although the great care taken in many experiments to ensure subject-subject and experimenter-subject anonymity can be justified by the need to isolate intrinsic motives from socially induced ones, the infrequency of "anonymity in the wild" makes it hard to interpret the findings.

[37] See for instance Nosek, Banaji and Greenwald (2002).

- When the experimenter asks a subject how he *would* react if another subject behaved unfairly rather than observing how he *does* react to the same behavior, the answer may not be valid.[38]
- It is virtually impossible to recreate, inside the laboratory, the ongoing open-ended interactions that shape much of social behavior.[39] (It is easy to *model* them as iterated games, but the models suffer from the problems discussed in the previous Section.)

In conclusion, it is impossible to deny the intellectual excitement generated by experimental economics and the opportunities it offers for testing and refining behavioral hypotheses. The reliability of most findings is not in question. The large, and largely unresolved issue concerns their validity. What seems to be needed at this stage is a much more active and systematic interaction between experiments and case studies, as in the work by Linda Babcock and George Loewenstein cited above or in the work on labor markets by Ernst Fehr and his associates.[40]

---

[38] A weakness of some early studies by Ernst Fehr and his associates (e.g. Fehr and Fischbacher 2003, 2004) is that subjects do not respond to an actual choice by other parties, but to a range of hypothetical choices. In later work (Falk, Fehr and Fischbacher 2005, p. 2022) they show that the latter (the "strategy method") is in fact not valid. This suggests an intrinsic difficulty in experimental economics. Consider an Ultimatum Game. On the one hand, the determination of responses to proposals requires that the latter exhibit sufficient variation. On the other hand, proposals that will predictably trigger mutually detrimental rejections will rarely be made. This problem could be overcome by having subjects respond to computer-generated proposals at any level, as long as they thought they were dealing with a real person. Given the anonymity of the experiments, this would be easy to achieve. A norm against this practice seems to be emerging in the behavioral economics community, however, because the experiments would cease to be reliable if the practice became known in the student populations from which most subjects are taken.

[39] See, however, Gächter, Renner and Sefton (2008) for an attempt to attenuate this problem.

[40] Se notably Fehr, Goette and Zehnder (2009).

IV. Data analysis

I am now about to stick my neck out and discuss matters I know little about. My reason for doing so is that the little I do know suggests that there is something seriously wrong with the way social scientists use statistical analysis. In my comments below I draw on writings by scholars who are both highly recognized by their peers as eminent statisticians and are deeply skeptical about the way statistics is used in "normal social science". My views have been particularly influenced by the writings of the late David Freedman.[41] I use him and others, notably Chris Achen, as *authorities*, in the sense that I cite their views without fully having assimilated their reasons for holding them. If I had a first-hand understanding of the issues, I wouldn't need to use them as crutches. My excuse for this unscholarly practice is, once again, that I believe I understand enough to suspect that we may be observing wasteful and spurious research on a large scale.

As I understand data analysis, it has an almost infinite number of potential temptations, pitfalls and fallacies. Let me cite a few: data snooping (shopping around for independent variables until one gets a good fit), curve-fitting (shopping around for a functional form that yields a good fit), arbitrariness in the measurement of independent or dependent variables, sample heterogeneity, the exclusion or inclusion of "outliers", selection biases, the choice of the proper level of significance, the choice between one-tailed and two-tailed tests, the use of lagged variables, the problem of

---

[41] See notably Freedman (1991, 2005, 2006) and Freedman, Pisani and Purves (2007). In addition I have benefited from Achen (1982) and Abelson (1995). The brief comments in Ch. 6 of Bhidé (2008) are also very much in line with the views sketched here. From a somewhat different perspective, Ziliak and McCloskey (2008) is also relevant. It is worth while stressing that these writings reflect *insider criticism* of a kind that is rare within rational-choice theory.

distinguishing correlation from causation, and that of identifying the direction of causation.

These problems – of which I have cited only some of the best known - are too numerous and varied to be fully covered by a textbook exposition, even at an advanced level. There are certain general lessons, such as testing for "robustness", but even then the number and variety of tests to run is a matter of judgment and experience. Scholars simply have to learn by trial and error until they know what tends to work. Data analysis is not a science, nor – as is sometimes asserted – an art, but a *craft*. In the words of Chris Achen, it is guided by "informal norms" shared by elite scholars rather than by formal rules that can be mechanically applied.[42] To learn the craft properly, a practitioner has to work through hundreds, perhaps thousands of applications. Citing Achen again, "wise investigators know far more about true variability across observations and samples than any statistical calculation can tell them".[43] Similarly, the "process of testing and eliminating counterhypotheses is a subtle skill that cannot be reduced to rote".[44] For all but exceptionally gifted scholars, the acquisition of "wisdom" is a task that is so time-consuming and demanding that it excludes the acquisition of substantive knowledge in any broad field of empirical inquiry.

At the same time, substantive knowledge is often indispensable. Among the various problems I enumerated above, the crucial one of distinguishing causal from spurious correlations may require deep familiarity with the field

---

[42] Achen (1982), p. 7.

[43] *Ibid*., p. 40.

[44] *Ibid*., p. 52. a

in question, in order to know which among the indefinitely many possible variables one should include as controls in the regression equations. To take a simple example, a person unfamiliar with geometry might try to estimate the area of rectangles as a function of their perimeter. Drawing 20 typical rectangles and doing the regression, he finds a correlation coefficient of 0.98.[45] In a similar example, he might try to estimate the surface area of randomly selected cylinders and cones as a function of their radius and height, and find a significant relationship.[46] In both cases the correlations would be spurious and non-predictive. In these examples, to be sure, the correct understanding is a matter of logic, not of causality. They serve only to illustrate the point that in the absence of substantive knowledge – whether mathematical or causal – the mechanical search for correlations can produce nonsense.

I suggest that a non-negligible part of empirical social science consists of *half-understood statistical theory applied to half-assimilated empirical material*. To substantiate this assertion, I first refer to David Freedman's detailed analyses of six articles (four of which are reprinted in his *Statistical Models*) published in leading academic journals: four from *American Political Science Review*, one from *Quarterly Journal of Economics*, and one from *American Sociological Review*[47]. The number of mistakes and confusions that he finds – some of them so elementary that even I could understand them – is staggering. It would be tempting to dismiss his criticism by responding that "substandard work exists everywhere". Yet,

---

[45] Freedman, Pisani and Purves (2007), pp. 211-13.

[46] Bhidé (2008), p. 241-42.

[47] The four articles reproduced in Freedman (2005) are discussed quite thoroughly, the remaining two more briefly in Freedman (1991), p. 301-2.

commenting on a subset of three of the articles, Freedman writes that they « may not be the best of their kind, but they are far from the worst. Indeed, one was later awarded a prize for the best article published in *American Political Science Review* in 1988 ».[48] If a substandard article can not only pass peer review in the leading journal of the profession but also be deemed "best of the year", one has to wonder about the quality of the field as a whole.

Next, I refer to his comments on how to avoid data snooping. From my limited exposure to the literature, I have concluded that even when scholars try to be honest and not rig the cards in their favor, they may unconsciously favor definitions and measurements that favor the hypothesis they want to be true. To keep this tendency in check, the scholar could use either *replication* or *cross validation*. The former, according to Freedman, is "commonplace in the physical and health sciences, rare in the social sciences".[49] The latter takes the following form: "you put half the data in cold storage, and look at it only after deciding which models to fit. This isn't as good as real replication but it's much better than nothing. Cross validation is standard in some fields, not in others."[50] As far as I can gather, it is *not* standard in the social sciences. It is not recommended in textbooks nor required by journal editors. An alternative form of self-binding – probably too utopian to be seriously considered – would be to post the hypothesis to be tested on the

[48] Freedman (1991), p. 301.

[49] Freedman (2005), p. 64.

[50] *Ibid*.

Internet ahead of testing it. The alcohol researcher Kettil Bruun apparently used a procedure of this general kind.[51]

In a recent work on *The Cult of Statistical Significance* Stephen Ziliak and Deirdre McCloskey denounce the mindless search for the magical 5% significance level, at the expense of substantive significance. Their point is not new.[52] I cite their book here only because of two remarks they cite from prominent economists.[53] We are told that Orly Ashenfelter « said that he 'basically agreed' with our criticism of statistical significance but then added that 'Young people have to have careers' and so the abuse should continue ». Similarly, James Heckman « told us recently that he didn't bother to teach [the difference between substantive and statistical significance] because his students at a leading graduate school were 'too stupid' to do anything but the 5 percent routine ». Whether or not these reports of oral remarks are accurate, they resonate with the passages I quoted from Chris Achen. In fact, when I had the occasion to cite the first passage to Achen, his reaction was that because of the difficulty of assimilating the informal norms of the profession, "a 5% significance level at least puts some limit on the amount of self-deception one can employ".[54]

Yet even if mechanical application of regression models makes them less vulnerable to abuse, one might still question their usefulness. In one of his more provocative statements, David Freedman asserted that in his view

[51] Ole-Jørgen Skog (personal communication).

[52] See for instance Achen (1982), pp. 46-51.

[53] Their way of quoting is indirect, and perhaps of dubious propriety. In the main text the authors are referred to only as « one eminent econometrician » (p. 89) and « a famous econometrician » (p. 111). In the Index under « Ashenfelter » and « Heckman » we find page references to and paraphrases of these quotations.

[54] For witty and perspective insights on this point see also Abelson (1995), p. 56.

the truth of the matter lies somewhere between the following: "regression sometimes works in the hands of skilful practitioners, but it isn't suitable for routine use" and "regression might work, but it hasn't yet". [55]

I do not have a well-founded opinion about the proportion of "skillful" practitioners who absorb the informal norms and so are able to avoid the temptations, pitfalls, and fallacies. My suspicion is that they constitute a small elite. I do not feel confident that this suspicion is founded, but the importance of the claim (if true) seems to justify my making it. (Think of Pascal's wager.)

V. Explaining excessive ambitions

If many applications of rational-choice theory and statistical theory are wasteful or harmful, why do they persist? There cannot be a simple answer to this question. I shall briefly mention a couple of simplistic ones, and then develop two that, although incomplete, may suggest more promising directions.

Some scholars believe, almost as an axiom, that "social science" can or must become a *science*, on the model of the natural sciences. While talk about "physics-envy" would be too strong, there may be an unconscious desire to emulate the most prestigious scientific disciplines. The highly sophisticated theories hold out the promise of satisfying this desire. At a more mundane level, the high prestige of mathematical social science goes together with very high salaries. A young scholar with a talent for mathematics may easily be lured by the prospect of rising to the top of the profession. At the same time, the normal workings of self-deception may

[55] Freedman (1991), p. 292.

prevent him from understanding that he is weaving a web of froth. Finally, self-selection may cause recruitment into these professions to be biased in favor of scholars who are subject to the relevant forms of hypertrophy and atrophy to begin with.

Although these explanatory suggestions may, in a given case, have some force, they lack the necessary sociological dimension. The persistence over time of pseudo-science at a large scale is a collective phenomenon, which must be sustained by mechanisms of social interaction. I shall discuss two possible mechanisms: *mind-binding* and *pluralistic ignorance*. Although I shall conclude that neither is fully adequate by itself, they may perhaps supplement each other in ways that will enable some progress to be made.

In an important article, Gerry Mackie discusses the puzzling phenomena of the foot-binding of Chinese women and the female genital mutilations practiced in parts of Africa.[56] Limiting myself to the former practice, I shall cite two salient features of Mackie's analysis. First, foot-binding persisted as a *bad equilibrium*. Given that no parents would let their son marry a women did not have her feet bound, it was in the interest of the parents of girls to adhere to the practice. Although crippling and horribly painful, the practice was sustained by the fact that no family had an incentive to deviate unilaterally. Second, the practice stopped, over the span of a few decades, by *successful collective action*. Because people came to perceive that the practice made everybody worse off than they could be, groups of parents came together to pledge in public that they would not bind the feet of their daughters nor marry their sons to women whose feet were bound.

---

[56] Mackie (1996).

Before turning to the social sciences, let me suggest an application of this idea to theoretical particle physics.[57] Within this field string theory is now dominant, in the sociological sense that it is virtually impossible (in the U.S.) for someone not working within that paradigm to be hired as assistant professor at a major research university. At the same time, string theory is not dominant in the *scientific* sense, as shown by the fact that it has not been awarded a single Nobel Prize, mainly because it has not generated confirmed predictions that are not also consequences of rival theories. One would think that from a scientific point of view, a department that contained a mix of string theorists and other theorists would be healthier than one in which all the particle theorists subscribed to string theory. This is, for instance, the view of Gabriele Veneziano (personal communication), himself a co-inventor of string theory. Yet in the given state of affairs, the dominance of string theory persists as a bad equilibrium. For American students to be *marriageable*, that is, capable of being hired as particle theorists by a high-prestige department, they *must* work in string theory.

In theoretical physics, sociological and scientific domination diverge. In economics, they seem to coincide. Consider first scientific domination. Many of the economists who have received the Alfred Nobel Memorial Prize for Economic Science work within the paradigms of rational choice theory and statistical modeling. Yet is a noteworthy fact that *not a single one of them has been awarded the prize for confirmed empirical predictions*. By an ironic contrast, on the one occasion it *was* awarded on that basis it went

---

[57] The following draws on Smolin (2006). As far as I know, string theory may, unlike rational-choice theory, be the correct theory of the part of the universe it pretends to explain. I cite string theory only to illustrate a sociological idea.

to Daniel Kahneman for his work in behavioral economics, notably for the discovery of loss aversion.

Let me mention a couple of other ironies. Earlier, I mentioned the disastrous performance of Long Term Capital Management. Two of the founders of this fund, Robert C. Merton and Myron Scholes, had received the Prize one year before the crash. Consider also the work of George Akerlof on asymmetric information, and notably his model of "the market for lemons" for which he received the Prize in 2001.[58] The model is firmly within the rational-choice tradition. In a situation that is essentially identical to the market for lemons, "The Winner's Curse", experiments show that people consistently fail to act rationally.[59] Contrary to Akerlof's predictions, people *buy lemons*. The irony is that much of Akerlof's current work is firmly within behavioral economics.

Consider next sociological domination. My personal observation of the American academic situation strongly suggests to me that departments of economics and, increasingly, political science are caught in a bad equilibrium. The mind-binding to which they subject their students is due, at least in part, to the perceived need to produce marriageable – hirable - candidates. It may also be due in part to the value signals sent by the Nobel Prize Committee and similar institutions, such as the committee that awards the John Bates Clark Medal. Although the latter has shown some openness, by awarding the Medal to Steven Levitt and Matthew Rabin, it has also chosen to award scholars within "science-fiction economics".

[58] Akerlof (1970).

[59] For a survey, see Charness and Levin (2009).

Mind-binding is obviously different from foot-binding. Even before the latter was abolished, it was widely perceived as absurd and perverse, a perception that eventually led to its abolition. By contrast, most chairs of departments of economics and political science probably do not perceive themselves as being in a bad equilibrium. To the extent that individuals scholars are seized with occasional doubts, as (being human) they can hardly fail to be, a look at what their colleagues are doing may assuage their worries or at least prevent them from speaking up. This remark brings me to the second interaction-based mechanism, *pluralistic ignorance*.

This idea dates from 1835, when Hans Christian Anderson published his tale about the "Emperor's New Clothes". It was given a more theoretical formulation five years later, in the second volume of Tocqueville's *Democracy in America*, and then rediscovered by Floyd Allport in 1924. In an extreme case, pluralistic ignorance obtains when no member of a community believes a certain proposition or espouses a certain value, but each believes that everybody else holds the belief or the value. For a game-theoretic example, we may imagine a case of collective action in which all participants have Assurance-Game preferences but each believes that all others have Prisoner's-Dilemma preferences. In the more common case it obtains when only a few members hold the belief or the value in question, but most of them believe that most others do.

In the game-theoretic example, when people act on their false beliefs about the preferences of others, the observed actions will confirm their beliefs. Each will make the non-cooperative choice as his best response to the non-cooperative behavior that his false belief makes him expect from others. They are trapped in a bad equilibrium, not (as in the mind-binding case) in a Nash equilibrium but in a rationalizable one.[60] What is needed to escape the bad equilibrium is not collective action, but improved information.

In the case of economic and statistical models, pluralistic ignorance would obtain if each scholar, although secretly worried about the procedures, kept quiet because of the perception that his colleagues are firmly convinced of their validity. There are several mechanisms that might be at work here. From my own experience I know very well how a scholar's confidence in his own judgment can be undermined by the fact that the majority thinks differently. *How could all these people, who are certainly smarter than I am, be so wrong?* Also, even with unshakeable self-confidence a scholar might worry that speaking up might cause ostracism and career obstacles.

Once again, I do not think this model offers a full explanation. Although it is impossible to tell the proportion of practitioners in the relevant disciplines that harbor secret doubts, they are likely to be a minority rather than a majority. My suggestion is that mind-binding and pluralistic ignorance may interact to produce the phenomena I have been trying to describe. Individual-level mechanisms such as the desire for rigor, prestige and reward, or the susceptibility to wishful thinking, self-deception and dissonance reduction, may also play a role. Given the general tenor of the present essay, however, it would obviously be absurd if I pretended to be able to make a more ambitious claim. The sociology of economics is not likely to come up with firmer answers than economics itself.

---

[60] Bernheim (1984).

Turning from explanation to prediction, it is not easy to see how the scientific community can move away from the bad equilibrium. The guild of economists and high-tech political scientists forms an almost impregnable bastion of skilled professionals, who do not believe in anyone's credentials but their own. If prestigious scholars from within the dominant tradition were to act as whistle-blowers, as the small child in Andersen's tale, they could perhaps make a difference. This might, however, require them to denounce their own past achievements and thereby risk destroying the very basis for their reputation. If the current economic crisis makes a dent in the hubris of the modelers, as seems likely, it may reappear as soon as the economy itself recovers. Keynes shook the complacency of the economic profession for a generation or so, but his impact eventually wore off.

Let me conclude on this point by exploring a conjecture alluded to earlier: we may learn more about the world by reading medium-prestige journals than by reading high-prestige and low-prestige journals. The last may owe their low prestige in the profession to their esoteric (e.g. neo-Marxist, post-Keynesian or neo-Austrian) character or to general sloppiness. As I have argued, the leading journals owe their prestige to the mathematical arguments that, although mostly trivial by the standards of mathematics, seem forbiddingly impressive to those who do not master them. As I have also argued, many of the articles published in these journals are in fact worthless. To learn something about the world, one should read the leading journals in specialized fields, such as *Industrial Relations* or *Journal of Development Economics*. Scholars publishing in these journals are kept honest by some kind of Reality Principle that high-flying theorists feel free to ignore.

VI. Towards a more modest and more robust social science.

As I do not want to repeat at length what I have written elsewhere[61], I shall only make a few brief comments about how I think the social sciences *should* develop, as distinct from my pessimistic views about how they *will* develop.

One crucial step is to *replace the aim of prediction with that of retrodiction*, and the concomitant move of *replacing general laws with mechanisms*. Retrodiction – explaining the past - is a perfectly respectable intellectual enterprise, because hypotheses about the past no less than predictions about the future can be falsified. Given an explanandum E and a hypothesis H, the scholar has to generate additional implications of E and see whether they do indeed obtain. If they do, they provide support for H. In addition, the scholar has to play the devil's advocate and think of the most plausible rival explanations of E, derive additional implications from them, and show that these do *not* obtain. These trivial statements are equally true whether E lies in the past or in the future. There is of course a difference, in that when generating additional implications from H the scholar who studies the past is easily tempted to choose only the ones he already knows to obtain. To overcome the temptation, he should try to generate *novel facts*, not already known to obtain.[62] Alternatively, as explained above, he should put half the data into cold storage and *not peek* until he has fitted his explanation to the other half.

In my terminology, mechanisms are *frequently occurring and easily recognizable causal patterns that are triggered under generally unknown*

---

[61] Elster (1999, Ch.I and 2007, *passim*).

[62] « Does the model predict new phenomena ? » (Freedman 1991, p. 293). See Abelson (1995), p. 184-87 for a good example.

conditions or with indeterminate consequences. Since this bare statement may be close to unintelligible, let me offer three examples, the first two inspired by Tocqueville's writings. [63]

If a King offers tax exemptions to the nobility but not to the bourgeoisie, the latter may react by envy towards the former or by anger towards the King. Even if we cannot predict which of the two reactions will occur, whichever of them does occur can be explained by the King's behavior.

If a King enacts repressive measures, they may make the subjects less likely to rebel (because the measures heighten their fear) and also more likely to rebel (because the measures increase their hatred).[64] The net effect is in general unpredictable, but if in a given case we observe that repression causes rebellion, we can conclude that the second effect dominated the first.

As a third example, let me cite La Fontaine's dictum that "Each believes easily what he fears and what he hopes". Since the hope that a given state of affairs will obtain is equivalent to the fear that it will not obtain, what determines whether, in a given situation, the first or the second belief will be triggered? In the recent movements of the stock market there may have been a point when agents switched from unjustified optimism to unjustified pessimism, but the psychology of the switch seems ill-understood.

---

[63] See Elster (2009).

[64] See for instance the cartoon in the London *Observer* on January 4 2009, showing a young boy on a heap of rubble in Gaza, watching Israeli bombers and asking himself « Is this going to make me more or less likely to fire rockets at Israel when I grow up ? » The social sciences are unlikely to provide an answer, even at the aggregate level.

To illustrate the ambiguity and indeterminacy of mechanisms, consider also the gambler's fallacy and its nameless converse. The purchase of earthquake insurance increases sharply after an earthquake, but then falls steadily as memory fades.[65] Like gamblers who make the mistake of believing that red is more likely to come up again if it has come up several times in a row, the purchasers form their beliefs by using the *availability heuristic*. Their judgment about the likelihood of an event is shaped by the ease with which it can be brought to mind, recent events being more easily available than earlier ones. Conversely, people living in areas that are subject to frequent floods sometimes (act as if they) believe that a flood is less likely to occur in year **n+1** if one has occurred in year **n.**[66] Like gamblers who make the mistake believing that red is less likely to come up again if it has come up several times in a row, they form their beliefs by relying on the *representativeness heuristics*. They believe, or act as if they believe, that a short sequence of events is likely to be representative of a longer sequence of which it is embedded. In casinos, players are equally vulnerable to either effect.[67]

In laboratory experiments it is sometimes possible to isolate sufficient conditions for a specific causal mechanism to be triggered. In those cases, we can appeal to a *law* and offer *predictions*. Outside the laboratory, where these conditions rarely obtain, retrodiction and appeal to mechanisms is usually the best we can do. In my opinion, the future of social science – or at least the hope for social science - lies in the cumulative generation of

---

[65] Slovic, Fischoff and Lichtenstein (1982), p. 465.

[66] Kunreuther (1976).

[67] Wagenaar (1988), p. 13.

mechanisms and their application to individual cases. As the program implies that we should cease looking for *laws*, it is a modest one. As the generation of mechanisms is cumulative and irreversible, it is a robust one.

Modest and robust statistical analysis also has a place in the social sciences. I tend to agree with the following statement: “Where the medians and means (and basic cross-tabulations) don’t persuade, the argument probably isn’t worth making”.[68] Statistical analysis should indeed be seen as “principled argument”, in Abelson’s phrase. It crucially turns on *substantive causal knowledge* of the field in question together with the *imagination* to concoct testable implications that can establish “novel facts”.

More generally, some often cited words by Keynes still bear repeating: “If economists could manage to get themselves thought of as humble, competent people on a level with dentists, that would be splendid.”[69] The competence of economists may not be in question, but their humility is. *Or perhaps humility properly conceived is part of competence*. To cite Pascal again, “The last step that Reason takes is to recognize that there is an infinity of things that lie beyond it. Reason is a poor thing indeed if it does not succeed in knowing that.”

---

[68] Bhidé (2008), p. 244.

[69] Keynes (1931), p. 373.

REFERENCES

Abelson, R. P. (1995), *Statistics as Principled Argument*, Hillsdale, NJ: Lawrence Erlbaum.

Achen, C. (1982), *Interpreting and Using Regression*, Beverly Hills: Sage.

Ainslie, G. (1975), "Specious reward", *Psychological Bulletin* 82, 463-96.

Akerlof, G. (1970), "The market for lemons", *Quarterly Journal of Economics* 84, 488-500.

Akerlof, G. and Shiller, R. (2009), *Animal Spirits*, Princeton University Press.

Allen, T. and Carroll, C. (2001), "Individual learning about consumption", *Macroeconomic Dynamics* 5, 255-71.

Ariely, D. (2008), *Predictably Irrational*, New York: Harper

Arkes, H. and Blumer, C. (1985), « The psychology of sunk cost », *Organizational Behavior and Human Decision Processes* 35, 124-40.

Babcock, C. Wang, X. and Loewenstein G. (1992), "Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias", *Quarterly Journal of Economics* 111, 1-20.

Babcock, L. and Loewenstein, G. (1997), "Explaining bargaining impasse: the role of self-serving biases", *Journal of Economic Perspectives* 11, 109-26.

Becker, G. (1974), "A theory of social interaction", *Journal of Political Economy* 82, 1063-94.

Bell, D., Raiffa, H. and Tversky, A., eds. (1988), *Decision Making*, Cambridge University Press.

Benartzi, S. and Thaler, R. (1995), "Myopic loss aversion and the equity premium puzzle", *Quarterly Journal of Economics* 110, 7-92.

Bergstrom, T. (1989), "A fresh look at the rotten kid theorem", *Journal of Political Economy* 97, 1138-59.

Bernheim, D. (1984), "Rationalizable strategic behavior", *Econometrica* 52, 1007-28.

Bhidé, A. (2008), *The Venturesome Economy*, Princeton University Press.

Brocas, I. and Carrillo, J., eds. (2003, 2004), *The Psychology of Economic Decisions*, vols. I and II, Oxford University Press.

Buehler, R., Griffin, D, and Ross, M. (2002), "Inside the planning fallacy", in Gilovich, Griffin and Kahneman, eds., pp. 250-70.

Camerer, C. (2000), "Prospect theory in the wild", in Kahneman and Tversky, eds., pp. 288-300.

Camerer, C. (2003), *Behavioral Game Theory*, New York: Russell Sage.

Camerer, C., Loewenstein, G. and Rabin, M., eds. (2004), *Advances in Behavioral Economics*, New York: Russell Sage.

Charness, G. and Levin, D. (2009), "The origin of the winner's curse", *American Economic Journal: Microeconomics* 1, 207-36.

Chiappori, P.-A. (1990), "La théorie du consommateur est-elle refutable?", *Revue Economique* 41, 1021-45.

Connolly, T. Arkes, H. and Hammond, K., eds. (2000), *Judgment and Decision Making*, Cambridge University Press

Dixit, A. and Skeath, S. (2004), *Games of Strategy*, 2nd ed., New York: Norton

Elster, J. (1999), *Alchemies of the Mind*, Cambridge University Press.

Elster, J. (2000), "Rational-choice history: A case of excessive ambition?", *American Political Science Review* 94, 685-95.

Elster, J. (2007), *Explaining Social Behavior*, Cambridge University Press.

Elster, J. (2009), *Alexis de Tocqueville: The First Social Scientist*, Cambridge University Press.

Fehr, E. and Fischbacher, U. (2003), "The nature of human altruism", *Nature* 425 785-91.

Fehr, E. and Fischbacher, U. (2004), "Social norms and human cooperation", *Trends in Cognitive Sciences* 8 185-90

Fehr, E., Goette, L. and Zehnder, C. (2009), "A behavioral account of the labor market", *Annual Review of Economics* 1, 000-000.

Falk, A, Fehr, E. and Fischbacher, U. (2005), « Driving forces behind informal sanctions », *Econometrica* 73, 2017-30.

Feller, W. (1968), *An Introduction to Probability Theory and its Applications*, vol.1, New York: Wiley.

Freedman, D. (1991), "Statistical models and shoe leather", *Sociological Methodology* 21, 291-313.

Freedman, D. (2005), *Statistical Models,* Cambridge University Press.

Freedman, D. (2006), "Statistical models for causation", *Evaluation Review* 30, 691-713.

Freedman, D. Pisani, R. and Purves (2007), *Statistics*, New York Norton

Friedman, M. (1953), *Essays in positive economics*, University of Chicago Press.

Frydman, R, and Goldberg, M. D. (2007), *Imperfect knowledge Economics*, Princeton University Press.

Gächter, S., Renner, E. and Sefton, M. (2008), "The long-run benefits of punishment", *Science* 322, 1510.

Gilovich, T., Griffin, D. and Kahneman, D., eds. (2002), *Heuristics and Biases*, Cambridge University Press

Herrnstein, R. and Prelec, D. (1992), "Melioration", in Loewenstein and Elster, eds. pp. 235-63.

Hirschman, A. O. (1986), "On democracy in Latin America", *New York Review of Books* 33 (6).

Johansen, L. (1977), *Lectures on Macroeconomic Planning*, Amsterdam: North-Holland

Kahneman, D. and Tversky, A. (1974), "Judgment under uncertainty", *Science* 185, 1124-31.

Kahneman, D., Slovic, P. and Tversky, A., eds. (1982), *Judgment under Uncertainty*, Cambridge University Press.

Kahneman. D. and Miller, D. (1986), "Norm theory", *Psychological Review* 93, 136-53.

Kahneman, D., Diener, E. and Schwartz, N., eds. (1999), *Well-Being*, New York: Russell Sage.

Kahneman, D. and Tversky, A., eds., (2000), *Choices, Values, and Frames*, Cambridge University Press

Keynes, J. M. (1931), *Essays in Persuasion*, London: Macmillan.

Kruglanski, A. and Webster, D. (1996), « Motivated closing of the mind : 'seizing' and 'freezing' », *Journal of Personality and Social Psychology* 103, 263-83.

Kunreuther, H. (1976), "Limited knowledge and insurance protection", *Public Policy* 24, 227-61.

Lichtenstein, S. and Slovic, P., eds. (2006), *The Construction of Preferences*, Cambridge University Press.

Loewenstein, G. (2007), *Exotic Preferences*, Oxford University Press.

Loewenstein, G. and Elster, J., eds. (1992), *Choice over Time*, New York: Russell Sage

Mackie, G. (1996), "Ending footbinding and infibulation: A convention account", *American Sociological Review* 61, 999-1017.

McClure, S. et al. (2004), "Separate neural systems evaluate immediate and delayed monetary rewards", *Science* 306, 503-7.

Montaigne, M. de (1991), *Essays*, Harmondsworth : Penguin Books.

Nelson, R. and Winter, S. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Mass.: Harvard University Press.

Nelson, R., Winter, S, and Schuette, H. (1976), "Technical change in an evolutionary model", *Quarterly Journal of Economics* 90, 90-118.

Neurath, O. (1913), "Die verirrten des Cartesius and das Auxiliarmotiv", translated in his *Philosophical Papers* vol.I, Dordrecht: Reidel 2004.

Nosek, B., Banaji, M. and Greenwald, A. (2002), "Harvesting implicit group attitudes and beliefs from a demonstration website", *Group Dynamics* 6 101-15.

Read, D., Loewenstein, G. and Rabin, M. (1999), "Choice bracketing", *Journal of Risk and Uncertainty* 19, 171-97.

Rozin, P. and Nemeroff, C. (2002), "Sympathetic magical thinking", in Gilovich, Griffin and Kahneman, eds., pp. 201-16.

Skinner, B. F. (1981), "Selection by consequences", *Science* 213, 501-4

Slovic, P., Fischoff, B. and Lichtenstein, S. (1982), "Facts versus fears: Understanding perceived risk", in Kahneman, Slovic and Tversky, eds., pp. 464-89.

Smolin, L. (2006), *The Trouble with Physics*, Boston: Houghton Mifflin

Stainforth, D. et al. (2007), « Confidence, uncertainty and decision-support relevance in climate predictions", *Philosophical Transactions of the Royal Society* 365, 2145-61.

Strotz, R. H. (1956), « Myopia and inconsistency in dynamic utility maximization », *Review of Economic Studies* 23, 165-80.

Taleb, N. (2005), *Fooled by Randomness*, New York: Random House.

Taleb, N. (2007), The Black Swan, New York: Random House.

Thaler, R. and Shefrin, H. (1981), "An economic theory of self-control", *Journal of Political Economy* 89, 392-406.

Thaler, R. and Sunstein, C. (1008), *Nudge*, New Haven: Conn.: Yale University Press.

Wagenaar, W. (1988), *Paradoxes of Gambling Behavior*, Hillsdale, NJ: Lawrence Erlbaum.

Weitzman, M. (2009), « On modeling and interpreting the economics of catastrophic climate change », *The Review of Economics and Statistics* 91, 1-19.

Wilkinson, N. (2008), *An Introduction to Behavioral Economics*, New York: Palgrave Macmillan.

Winter, S. (1964), « Economic ‘natural selection’ and the theory of the firm », *Yale Economic Essays* 4, 225-72.

Ziliak, S. and McCloskey, D. (2008), *The Cult of Statistical Significance*, Ann Arbor : University of Michigan Press 2008