The influence of experimental and computational economics: Economics back to the future of social sciences

Giovanna Devetag
University of Trento

Francisco Louçã
ISEG, Lisbon

Preliminary version, not to be quoted
Comments welcome

Abstract

Economics has been a most puzzling science, namely since the neoclassical revolution defined the legitimate procedures for theorisation and quantification. Its epistemology has based on farce: decisive tests are not applied on dare predictions. As a consequence, estimation has finally been replaced by simulation, and empirical tests have been substituted by non-disciplined exercises of comparison of models with reality. Furthermore, the core concepts of economics defy the normally accepted semantics and tend to establish meanings of their own. One of the obvious instances is the notion of rationality, which has been generally equated with the apt use of formal logic or the ability to apply econometric estimation as a rule of thumb for daily life. In that sense, rationality is defined devoid of content, as alien to the construction of significance and reference by reason and social communication.

The contradictory use of simulacra and automata, by John von Neumann and Herbert Simon, was a response to this escape of economic models from reality, suggesting that markets could be conceived of as complex institutions. But most mainstream economists did not understand or did not accept these novelties, and the empirical inquiry or the realistic representation of the action of agents and of their social interaction remained a minor domain of economics, and was essentially ignored by canonical theorizing.

The argument of the current paper is based on a survey and discussion of the twin contributions of experimental and computational economics to these issues. Although mainly arising out of the mainstream, these emergent fields of economics generate challenging heuristics as well as new empirical results that defy orthodoxy.

Their contributions both to the definition of the social meanings of rationality and to the definition of a new brand of inductive economics are discussed.

JEL: B41, B52, C71, C72

1. Introduction and some methodological digressions

Contributions in experimental economics nowadays include all domains of economics, from individual decision making to industrial organization, game theory, finance and macroeconomics. Clearly, any attempt at reviewing such immense body of literature in the following few pages would be beside the point (a must-read for an extensive review of the discipline categorized by subfields is Roth and Kagel (1995)); rather, in the following section we will try to briefly highlight few of the main problems raised by contemporary investigation in experimental economics, focussing especially on studies of individual decision making and rationality on one hand, and on strategic interaction on the other hand.

The first experiments in economics, some of which date back to the works of Thurstone (see Roth, 1995) were concerned, as appears intuitive, with testing ordinal utility theory in simple individual decision making tasks. Later on, the publication of *Theory of Games and Economic Behaviour* by Von Neumann and Morgenstern gave rise to a great amount of experimental work on strategic interaction. The first, famous attempts to verify the empirical soundness of expected utility date back to the famous informal experiments by Allais (1953) and Ellsberg (1961). The field of individual decision making has borrowed more than any other from cognitive psychology, and the extensive and fruitful dialogue between theory and experimentation has given birth to the field of `behavioural economics', which fosters the introduction of psychological principles into economic theory (see, e.g., Loewenstein, 1999).

Game theory too contributed enormously to sky-rocketing economists' interest toward experiments, since it lent itself very naturally to being tested in the lab, as testified also by the fact that some prominent game theorists, such as John Nash, Thomas Schelling and Reinhardt Selten have been (and the last one still is) active experimentalists. Relatedly, the famous experiments by Chamberlin (1948) and afterward Smith (1962) generated an equally influential stream of experimental research on the functioning of diverse market institutions.

Nowadays, experimental economics seems to have been finally acknowledged as a fully legitimate branch of `mainstream' economic science; the assignment in 2002 of the Nobel Memorial Prize to Vernon Smith, who is unanimously considered, together with Charles Plott, the pioneer of the discipline (though certainly not the *firs*t economic experimentalist) appeared to many as the long-anticipated official acceptance of laboratory experiments within the realm of `respectable' methods of inquiry in the economics profession. It seemed all the more surprising to many, since the impossibility to test economic theories by means of experiments had for a long time been one of the most widely accepted articles of faith among economists.

Smith (2002) himself reports the Cambridge economist Joan Robinson affirming that "economists cannot make use of controlled experiments to settle their differences" (quoted in Smith, 2002: 114). She was only one among many to have made such (or a similar) statement. They were of course all wrong in a literary sense; it must be said that Robinson referred precisely to experiments in the macro sense and, in spite of favouring some sort of social engineering, she refused to accept that economists could simply test their models and check the results in real economies, since social life is not a laboratory. Nevertheless, there is something peculiar about experiments in economics which renders them very different, in some fundamental way, from experiments in the natural sciences like physics, biology or chemistry and which prevents many from considering economics an empirical science. Such peculiarity lies, in our view, in the perceived larger gap - as compared to the commonly held perception regarding the realm of the natural sciences - that would exist between phenomena that emerge in the laboratory and what happens in the `real world'. A widely diffused attitude among economists confronted with experimental results that apparently refute some theoretical predictions is that of dismissing the evidence on the - admittedly, rather unconvincing - grounds that individuals operating in the real world and subject to market forces `would get it right'. This objection may at times get instantiated in the form of a critique to the relevance for real world economic decisions of the small financial stakes typically involved in laboratory experiments. As Colin Camerer nicely puts it when speaking about the emergence of other-regarding behaviour in experimental bargaining games ``If I had a dollar for every time an economist claimed that raising the stakes would drive ultimatum behaviour toward self-interest, I'd have a private jet on standby all day" (Camerer, 2003: 60).

Such criticism to the `falsificational' role of experiments is not at all unique to economics; on the contrary, it pervades all experimental sciences and can be subsumed under the well-known Duhem-Quine (D-Q) problem, which states that any experiment jointly tests the specific theory that motivated the experiment in the first place, and the set of auxiliary hypotheses that were required in order to implement the experiment (see, e.g., Smith, 2002). Among the latter are, in the case of economic experiments, the hypotheses that financial rewards were adequate, that subjects were sufficiently motivated, that instructions were clear, that all potentially interfering noisy variables had been controlled for, and so on (see Smith, 2002). Consequently, if the results of the experiment contradict the prediction of the theory, it can be equally safely concluded that the theory was falsified or that one or more of the auxiliary hypotheses were falsified.

And yet, economists seem particularly reluctant to take experimental results seriously, and many of them persistently prefer the second of the two possible explanations of the evidence, quite often with a dismissive attitude toward those scholars – mainly psychologists – who dare to question expected utility theory and related dogmas. Binmore states: "Just as we need to use clean test tubes in chemistry experiments, so we need to get the laboratory conditions right when testing economic theory'' (Binmore, 1999, p. F17). While it's hard to deny the need to perform chemistry experiments using clean tubes, it's not at all clear when exactly the conditions for economic experiments can be considered to be right. Binmore (1999) seems to provide an answer by stating that economic theory can be subject to laboratory testing only when the task at hand is reasonably simple and it is perceived as such by the experimental subjects, when financial stakes are `adequate', and when the time allowed for trial-and-error learning is `sufficient'. But the fact that `adequate' and `sufficient' are put between quotation marks by the author should immediately point at the inherent arbitrariness of such definition. When are financial incentives considered adequate and time for learning sufficient? No matter how high the stakes (conditional on the relatively limited financial possibilities of the majority of experimental labs) and how many repetitions of a single trial subjects perform, there simply is no way to establish when these conditions are met in any particular experiment.

This impediment emerges out of three types of difficulties. First, it seems reasonable to require that time for learning should be longer the more complicated the task; hence, it sounds implausible to establish a rule that does not take differences between experiments into account. Second, there are many more arguments that defenders of economic orthodoxy can (and do) use to continuously downturn the relevance of experimental results, including the defence of existing theories on the grounds that these are `useful approximations' not intended to describe individual behaviour precisely. But when is a given approximation not good enough any more (especially if there is a better approximation available?) And, finally, no matter how much time for learning you give your subjects, there will always be someone objecting that the feedback environment of the laboratory is certainly sparse compared to the real environment in which economic agents operate and learn (Binmore, 1999). Unfortunately, there is no easy way to establish whether such statements are right or wrong. It could as well be argued, for example, that the environment in which real economic agents learn and interact, exactly because is immensely richer than the artificial environment of the lab, is also much more complex and difficult to understand for boundedly rational individuals.

When problems are ill-defined, as most real life problems are, and the number of relevant variables is much higher than our minds are able to grasp, learning is more difficult because feedback can be very ambiguous, and the process of causal attribution necessary for most learning processes to occur can be impeded by the opacity of the relations between actions and consequences which instead are made transparent to subjects in the typical laboratory experiment.

Having said that, we think experimental economics so far has only benefited from such intense criticism, and it will continue to benefit even more in the future. Partly in response to the objections above outlined, new experiments have been conducted, experimental techniques have refined a great deal, new, more radical departures from experimental `orthodoxy', like brain imaging experiments and anthropological studies on small-scale societies around the world, are becoming increasingly popular, and truly unexpected phenomena have been discovered, which in turn have stimulated both theorizing and further experimentation. At the same time, the accumulation of evidence has made it increasingly difficult to defend existing theories where these do not work, and has strengthened the belief in the accuracy of theories that do seem to work in the lab.

More radically, we think that the insistence on a theory-testing view of experiments is highly misleading and does not make justice of the importance of experimentation in economics. Maybe too much ink has been wasted debating on whether obscure and highly formal theories that belong to mathematics more than economics could be said to have been falsified or not by experimental results, while we believe that a much more *inductive* approach should inspire both economists who run experiments and economists who build new theories to interpret their results (the two of course sometimes coincide!). Experiments can be wonderful tools to inductively explore human behaviour in economic contexts and to discover new, interesting behavioural regularities. This applies to individual decision-making and even more so to situations of collective interaction like games and markets. For example, the mere discovery that individuals are not Bayesian sounds much less interesting in itself than the unpacking of the specific heuristics (like, e.g., representativeness, availability and the like) that people use when forming probabilistic judgments[1].

Another good example is the behaviour of experimental subjects in coordination games, where theory is simply silent due to the impossibility of discriminating among multiple equilibria. In such cases, experiments have been useful to explore the context-inspired coordination patterns among groups of individuals. Experiments on coordination games have produced other surprising results, such as the emergence of almost perfect coordination observed in market entry-games arising out of simple behavioural heuristics combined with heterogeneity.

This is acknowledged, not surprisingly, by some of the most heterodox of the orthodox, such as William Brock (2002), who states:
"Some versions of standard economics do not much look for patterns. Instead, they are a type of deductive study that is often highly formal involving proofs and lemmas. That is not the approach to science that most scientists take; this type of standard economics approach belongs to *mathematical philosophy* more than it does to science. It may be important, but it is not science. As some famous sage stated: `To do science is to find patterns and explain them' "(Brock, 2002: 31, emphasis added).

[1] By the way, this was the initial spirit with which Kahneman and Tversky initiated their famous 'Heuristics-and-Biases' research program, which has been subsequently highly criticized on the grounds that it would depict humans as dumb ill-informed decision makers who fall victims of all sorts of cognitive illusions and framing effects. In reality, Kahneman and Tversky explicitly referred to an analogy with studies on perception, where cognitive illusions are of scientific interest only insofar as they can provide clues to understand the functioning of human vision and are not relevant in themselves.

Indeed, this statement echoes previous considerations on the exhaustion of economics as an elegant corpus of deductive theories, with a rigorous aesthetic discipline imposed by some bourbakist chien-de-garde. Frank Hahn, who endeavoured the representation of general equilibrium with Kenneth Arrow, recognised with sorry that "the task we set ourselves after the last war, to deduce all that was required from a number of axioms, has almost been completed, and while not worthless has only made a small contribution to our understanding" (Hahn, 1994: 258).

In our view, experiments can represent an extremely useful alternative tool to find patterns and try to explain them. Together with agent-based simulations, they can foster a much more inductive, bottom up approach to economics which can bring the discipline much closer to being a science than it has ever been in the last century or so. In particular, once the notion of bounded rationality, for a long time advanced by Herb Simon as a useful concept to inspire empirically-driven models of human behaviour and invariably neglected by most mainstream economists, is taken seriously, experiments appear as the ideal methodological tools (though not the only ones) to give bounded rationality a more concrete and operational meaning. Furthermore, experiments, together with simulations, can help disentangle the complex relations between the micro-motives and the macro-behaviour of economic systems (as Thomas Schelling (1978) had long ago pointed out in a wonderfully inspiring book), once fictitious simplifications such as that of the `representative agent' are abandoned.

2. Individual preferences and rationality

Among the most basic tenets of economics is the notion of preferences, with its declination in terms of time preferences and risk preferences. Utility theory simply derives from such primitives as a convenient mathematical formulation of preferences, and people's attitudes toward different degrees of risk are conveniently rendered through the addition of a risk parameter to the utility function.

Economists are so used to modelling individual behaviour using these categories that most of them simply take them for granted as fundamental `first principles' from which everything else derives. It is curious in this respect to consider how economists, in their continuous effort to render economics a history-free discipline, seem totally unaware of the historical origin of many of the notions that they so carelessly utilize everyday as their habitual modelling tools. The notion of utility itself has a long history dating back at least to Bentham, where the term was used to refer to the – supposedly – fundamental drivers of human behaviour, *pleasure* and *pain*. Then, as it is well known, such hedonic flavour completely disappeared in the modern conception of utility as numerical expression of well-defined preferences, before being rediscovered in recent years thanks to the work of Kahneman and Tversky (1979) and Kahneman and his collaborators (e.g., Kahneman, Wakker, Sarin, 1997).

Preferences then became the fundamental building blocks to describe the behaviour of economic agents, with the only constraint given by an internal consistency requirement, i.e., the need for preferences to be coherent and relatively stable over time and over contexts.

Finally, although in principle no restrictions are imposed as regards the arguments of the utility function of a particular individual (which may contain reference to one's own as well to other individuals' well being), the notion of utility maximization has increasingly been equated *de facto* with the maximization of one's own material interest.

However, as Albert Hirschman points out, it was only recently in the history of thought that the notion of interest acquired a special relevance as an innate driver of human behaviour, replacing "the two categories that had dominated the analysis of human motivation since Plato, namely, the *passions* on the one hand, and *reason* on the other" (Hirschman, 1977, p. 43). Quite interestingly, the adoption of interest as a new paradigm to interpret human motivation had been accepted with intellectual enthusiasm mainly because it seemed to confer behaviour a greater predictability, given that passions were too volatile and reason appeared ineffectual (Hirschman, 1977).

However, decades of research in experimental economics and behavioural decision theory have seriously undermined the usefulness of such mental constructs as self-interest and preferences to model human behaviour in a coherent manner. What experiments have repeatedly shown is that behaviour - as measured and judged on the basis of adherence to an underlying set of coherent and stable preferences that the individual aims at satisfying given available information and budget constraints- is so sensitive to theoretically irrelevant factors such as context, labels, emotional states, to put into question the adequacy of using such mental categories to model decision making in a way that may prove descriptively sound. In the following we briefly report a (necessarily incomplete) list of findings: ??? A), B) …

Perhaps no piece of evidence is so simply and yet powerfully representative of humans' behavioural inconsistency as the evidence on preference reversal. At the same time, perhaps no other piece of evidence has been so disturbing for economists, who in fact have tried hard to eradicate the phenomenon, without quite succeeding. The first experiments on preference reversal were run by the psychologists Lichtenstein and Slovic (1971) and involved the two (perfectly equivalent, according to standard theory) tasks of choosing between two lotteries and pricing those same two lotteries. The lotteries were designed so that one offered a larger probability of winning a smaller amount of money, and the other offered a lower probability of winning a larger sum. The experiments show that individuals typically choose the first lottery in the `choosing' task while pricing the second lottery higher in the `pricing' task. Psychologists labelled this finding as a violation of the principle of `procedure invariance', which states that alternative but logically equivalent methods to elicit one's preferences should yield identical results. Psychological theories also offered a simple explanation for the finding, namely that subjects focus more on the money dimension when asked to price a lottery, while they pay more attention to the probability dimension when choosing. The experiments were then replicated by Grether and Plott (1979) in the hope to make the effect disappear once subjects would be given the right incentives and tighter control. But the violation persisted, and subsequent studies have demonstrated its robustness to varying experimental designs (see Camerer, 1995, for an extensive review).

Another fruitful area of research for experimental and behavioural economics is inter-temporal choice. Frederick, Loewenstein and O'Donoghue (2002) offer a comprehensive survey of the experimental evidence on time preference, together with an assessment of the state-of-the-art in models of inter-temporal decisions that try to incorporate some of the behavioural findings. Here we will only report a small subset of the evidence that they discuss. Time preference is a research area where the largest part of the experiments has been conducted by psychologists, although economists have replicated in some occasions using real monetary incentives. However, economists have produced a wide range of behavioural models of inter-temporal choice that attempt at explaining observed anomalies in savings and consumption decisions (Laibson, 1998), in principal-agent relationships (O'Donoghue and Rabin, 1999) and many others (see Frederick et al., 2002, and references therein).

As it is well known, standard economics solves the inter-temporal choice problem by assuming it away through the adoption of the discounted utility model (DU) (Samuelson, 1937). Under DU, all the motivations that drive decisions involving consumption that will occur at different moments in the future are subsumed under a single discounting parameter. It is difficult to underestimate the elegance and mathematical convenience of such simple formulation; however, and quite intuitively, experiments have highlighted numerous violations of the model. Furthermore, such violations cannot by any means be considered "mistakes", since there is no compelling argument to consider the DU model normatively correct (Frederick et al., 2002). For example, there is psychological evidence that individuals do not discount the future at a constant rate; rather, implicit discount rates tend to decline over time (e.g., Thaler, 1981; Benzion, Rapoport, Ygil, 1989). This evidence has been incorporated in models of inter-temporal decisions that involve hyperbolic instead of exponential discounting, which have been mainly applied to explain a vast array of stylized facts in consumption and savings behaviours (see, e.g., Laibson, 1994, 1997). However, further evidence (e.g., Rubinstein, 2000) shows that hyperbolic discounting can be as easily violated in experiments as exponential discounting, suggesting the need to incorporate changes in models that go beyond the simple substitution of a given functional form with another.

Further experiments have shown that individuals tend to prefer sequences of outcomes that show improvement over time (like, e.g., increasing wage profiles) to alternative sequences that allow strictly higher amounts of consumption in every period but show declining rates (Hsee, Abelson and Salovey, 1991). Moreover, gains seem to be discounted at a higher rate than losses, and large outcomes are discounted differently than small ones (e.g., Thaler, 1981).

Among the models that try to incorporate some of the evidence, to name a few, are models of habit formation (Abel, 1990; Jermann, 1998; Boldrin, Christiano and Fisher, 2001), models implying game-theoretic interactions between multiple-selves - aimed at modelling self-control and procrastination problems, which are assumed to be at the core of many empirically tested examples of time inconsistency (e.g., Elster, 1985) - models of referent-dependent utility which incorporate some ideas from Prospect Theory (e.g., Loewenstein and Prelec, 1992), models that incorporate anticipatory emotions such as anxiety, suspense and "savouring" (Loewenstein, 1987; Caplin and Leahy, 2001).

While most of these models generally offer a better account of the evidence compared to the standard discounted utility framework, nonetheless they are usually able to explain only part of the observed violations. Rather, the experimental evidence seems to suggest that multiple motives lie at the base of inter-temporal decisions. In addition, the empirical correlation between different measures of time preference, and between time preferences for different choice domains seems rather weak.

Loewenstein, Weber, Flory, Manuck and Muldoon (2001) try to `unpack' the concept of time preference by explicit reference to its constituent motives, which they identify in *impulsivity*, *compulsivity*, and *inhibition.* The first component is responsible for spontaneous, unplanned acts, the second for careful forward-looking decisions and the third component for the taming of impulsive behaviour. The authors argue that such dimensions can be measured quite reliably in individuals, and seem to predict different types of behaviours quite well (so, for example, various types of impulsive and visceral behaviours result positively correlated with impulsivity and negatively correlated with inhibition).

All in all, research on time preference has conclusively demonstrated the total inadequacy of the DU model and the need to search for radically different (and, possibly, empirically-driven) `first principles' to describe human choices over time.

A related domain of research on preferences has to do with the measurement of their degree of coherence and arbitrariness. The questions that these types of experiments try to answer are the following: "to what extent do individuals posses a coherent set of stable preferences that pre-exist at the moment of choice?" "To what extent do preferences reflect some (subjective) fundamental valuation?" "Are people aware of their own preferences, even after they have accumulated experience?" The answers coming from experiments seem largely negative for all these questions. A considerable body of evidence has accumulated in the last twenty years about the extreme sensitivity of preferences (as revealed by observed choices) to normatively irrelevant factors, such as for example the contextual elements of choice. The influence of context may take the form of framing effects, reference-dependence, choice-menu effects, and the like (see, e.g., Devetag, 1999, Mellers et al., 1998, for a survey). The evidence so far accumulated strongly highlights the inadequacy of the traditional economic depiction of individuals as characterized by clearly identifiable, complete and coherent preference sets.

In a more recent experimental work, Ariely, Loewenstein and Prelec (2003) conduct experiments aimed at testing whether preferences that subjects exhibit can be said to be related to some "fundamental" valuation or, on the contrary, exhibit a certain amount of arbitrariness. To this purpose, they borrow an experimental device used by Tversky and Kahneman (1974) and expose experimental subjects to a randomly generated numerical value (the "anchor") before asking subjects to place a monetary value on some sets of consumer goods or "hedonic" experiences (such as listening for a certain amount of seconds to a highly unpleasant noise). In one of such experiments, a group of student subjects was shown a series of consumer products (e.g., wine bottles and PC accessories), each accompanied by a brief description but without any mention of its market price. Subsequently, subjects were asked to state whether they would buy each of the products for a price equal to the last two digits of their social security number (the random "anchor" value). Finally, subjects were asked to state the maximum price they would be willing to pay (their WTP, or "willingness to pay") to purchase any of the goods. The "anchor" value was manipulated so as to make it transparent to subjects that it was clearly randomly drawn and had no connection whatsoever with the value of the products. Incentives were controlled through the well-known Becker-De-groot-Marschak procedure, which determined, through the use of a random device, whether the product would actually be sold to the subject on the basis of the first or second question asked. Subjects knew the procedure and knew that they could buy at most one product, so in principle they had a monetary incentive to reveal their "true" subjective evaluations.

Despite the explicit transparency to subjects of the random nature of the "anchor", the results show a highly significant anchoring effect in subjects' evaluations of the goods, in that subjects with social security numbers above the group median stated values for products that were from 57 to 107 higher than values stated by subjects with below-median social security numbers, and the difference is even more striking when comparing the average evaluation of the top quintile ($56) with that of the bottom quintile ($16). The authors refer to this effect as "arbitrariness" of preferences with respect to absolute evaluations. However, what emerges from the data is subjects' coherence in *relative* evaluations: for example, most subjects coherently valued the highly rated wine more than the poorly rated wine, and generally seemed to know the relative ranking of products within each category.

Such baseline experiment was replicated involving valuations of pain experiences instead of consumer goods to eliminate the possible influence of previous purchase episodes on subjects' evaluations; furthermore, the experiments allowed subjects to have a sample of the pain experience before making the evaluation. Finally, in one of the replications subjects' WTA ("willingness to accept", or the minimum price that they would be willing to receive to experience the pain for a certain time) was elicited using a multi-person auction, with the purpose of testing whether market forces would make the anchoring effect disappear.

Their general results highlight the robustness of anchoring effects and their resistance to experience and to market forces. Overall, these data suggest that there is a strong element of arbitrariness in individuals' absolute evaluations, although people subsequently show "local" coherence when adjusting their judgments in response to variations in a neighbourhood of the initial, arbitrary "anchor". Hence, to the extent that people seem to respond consistently to changes, such as for example, changes in price, they behave "as if" they were revealing a coherent set of stable and well-known "true" preferences. However, to the extent that such supposedly true preferences appear to be correlated with completely random "sunspot" variables, revealed preference theory needs a radical reassessment.

3. Social preferences

A related area of recent research in experimental economics deals with strategic interaction as modelled by game theory (von Neumann and Morgenstern, 1944). A large subset of experimental work in this area has to do with the study of altruism and other-regarding preferences, with the purpose of testing the traditional view of homo oeconomicus as motivated solely by self-interest. As Colin Camerer (2003) noted, results in this area of experimental investigation typically surprise *only* economists. In fact, what most of these studies show is that individuals show altruism to some extent, exhibit cooperative behaviours in a variety of contexts, assign high importance to fairness and equality considerations, and show consistent degrees of both positive and negative reciprocation (reciprocating altruistic acts and punishing unfair acts, even when these acts come at a cost to themselves). All in all, hadn't it been for decades of brain-washing with the "self-interest" assumption of canonical economic theorizing, these findings would hardly surprise *anyone*, as any economist who has ever attempted to explain the results of these studies to non-economists will be ready to acknowledge. Experimental research on social preferences exhibits maybe to the highest degree the benefits of the cumulative progress in knowledge that *series* of experiments permit to achieve through the interplay of simple theory-testing motivations (such as the emergence of the subgame-perfect equilibrium in the ultimatum game), intense (D-Q)-type criticism, and further careful experimentation that builds upon previous studies refining features of the experimental design or varying the subject population. By now, the evidence on altruism and reciprocation is immense, and has resisted various attempts at demolishing it to restore the re-assuring canonical image of the self-interested economic agent (for a detailed and updated critical survey, see Camerer, 2003).

The early research on social preferences elected for decades the famous Prisoner's Dilemma as the prototype-game to study violations of self-interested behaviour and the emergence of cooperation in simple settings. And the Prisoner's Dilemma has a long history of experimental testing, since its conception, as Mirowski aptly narrated. As Merrill Flood and Melvin Drecher invented the PD in January 1950, their essential motivation was "to expose the way that social relationships mattered for economic outcomes" (Mirowski, 2002: 356; also Morgan, 2001) and, in particular, to challenge Nash's concept of equilibria for non-cooperative games, since it was based on dubious dummies for rational behavior. It was because of that strategy that the payoff matrix was designed in order to contrast the Nash equilibrium with the alternative equilibrium emerging out of maximization of payoffs from cooperative strategies. The curious anedocte is that Flood invited two colleagues at Rand, Armen Alchian and John Williams, an economist and a mathematician, to play the game 100 times for pennies, asking both to note their observations on the successive events.

The fact is that Alchian the economist chose cooperation 68 times out of 100, and Williams the mathematician chose it 78 times. Their notes illustrate the asymmetry of expectations: Alchian predicted defection whereas Williams expected cooperation (Mirowski, ibid,: 357-9). But the most impressive fact was the difficulty of communication, as a product of the inability of learning in the universe of skewed payoffs: social life was indeed limited by 'rationality'. Nash reacted to this experiment suggesting that the world would be perfect if interaction was eliminated from the game: ""It is really striking, however, how inefficient AA and JW are in obtaining rewards. One would have thought them more rational. If this experiment were conducted with various different players rotating the competition and with no information given to a player of what choices the others have been making until the end of all trials, then the experimental results would have been quite different, for this modification would remove the interaction between the trials" (qu. in Mirowski, ibid.: 359). Certainly, with no interaction the problem of learning would be dismissible.

Either because of these challenges or because of the PD's narrowness in explaining individual behavior in social interaction, the fact is that more recently increasing attention has been given to the ultimatum game and to the trust game, which are both considered as more suitable than the PD to discriminate among alternative explanations of the results (for example, in case of defection in the PD it is not possible to distinguish between a truly rational and self-interested individual and a conditional cooperator with pessimistic beliefs about her opponent). As is well known, the ultimatum game is a simple sequential bargaining game with two players, a Proposer and a Responder. The Proposer has a sum of money, which she has to decide how to allocate between herself and the Respondent.

The Responder may accept the allocation proposed, in which case the money is split among the two accordingly, or reject it, in which case neither receives anything. The game-theoretic prediction prescribes that the responder, being only interested in the maximization of his monetary payoff, will accept any positive offer, however small. The Proposer, being able to anticipate the responder's behaviour, will propose an extremely unfair allocation, giving only the minimum possible to the Responder and keeping the rest for herself. Clearly, things go very differently in reality. In the typical one-shot ultimatum game with a condition of total anonymity, offers by the Proposers vary from 40% to 50% of the entire pie, and quite often offers below 20% are rejected. A variation of the game called the dictator game, in which the Responder has no influence on the final allocation, offers from the Proposers are generally lower than in the ultimatum game, suggesting that the Proposer's behaviour in the ultimatum game is not entirely due to altruism but, to some extent, also to strategic considerations (i.e., fear of rejection); however, offers in the dictator game show a much higher variability than in the ultimatum game, and seem to be much more sensitive to the contextual features of the experimental design (such as, for example, the use of "double blind" procedures that prevent even the experimenter from knowing the dictator's allocation decision, or having the dictator receive a brief description of the recipient before a decision is made).

Raising the stakes to sums that amount to one-month salary of the experimental subjects changes the results very little, as does allowing for repetition (although both offers and rejection rates tend to decline over time).

In the trust game (Berg, Dickhaut and McCabe, 1995) a subject (player 1) receives a sum T; she can decide to invest the entire sum or part of it by giving it to player 2, which then receives by the experimenter K times whatever he has received from player 1 (K>1); player 2 then must decide how much (if anything) give back to player 1. Game theory would prescribe that player 2 will keep all the money for himself; player 1, anticipating this behaviour, should give nothing to player 2. In the lab, subjects typically give half of their endowment to player 2, who in turn give back to player 1 a sum more or less equal to the initial investment. The game nicely captures situations of pure *trust*, whereby a player can decide to take a risky action (investing the money rather keeping it) in the hope that the other player will reciprocate trust, even at a cost to himself. The degree of trust is implicitly measured by the amount invested by player 1, and the degree of trustworthiness by the amount player 2 gives back. The fact that in a situation of total anonymity and one-shot interaction, most subjects decide to trust the other (unknown) player is a result that is slightly more surprising than behaviour in the ultimatum game, since a risk-averse norm that says "it's safer not to trust strangers" might have been easily transferred to the lab environment (however, it must be noted that trust is not very well repaid on average). Data from the trust game thus seem to identify the existence of a sort of "minimum" trust level among members of a society (since it arises in absence of any typical trust-enhancing condition such as reciprocal knowledge, repeated interaction, and so forth), and the fact that such level is above zero is not at all a trivial result.

Behaviour of the responders in the ultimatum game is often interpreted in terms of subjects coming to the lab with a set of pre-codified social norms that guide them in their daily, repeated interactions outside the lab, and `mistakenly' applying these norms to the one-shot context of the experiment, in absence perhaps of other clues due to the relative unfamiliarity of the laboratory environment (see also Camerer and Fehr, 2003). However, individuals seem able to capture the difference between repeated and one-shot games. Fehr and Fischbacher (2001) conducted experiments on repeated UG under two differing treatments. In one case, subjects played each time with a new opponent and the proposers were not informed about the past behaviour of the responders they were facing; in another treatment, proposers were allowed to know the history of the responder's choices up to that time. The second treatment clearly creates an opportunity for reputation formation; in fact, the responders have an incentive to reject low offers in early rounds, knowing that this information will be made available to the proposers, to increase the probability to receive high offers in later rounds. The data show that indeed, most responders take advantage of the `reputation formation' opportunity by raising their acceptance threshold (i.e., by rejecting higher offers, on average, than in the control treatment). These results show that individuals are perfectly able to distinguish between one-shot and repeated interactions, and modify their behaviour accordingly. Hence, it seems implausible to think of responders' behaviour in the standard UG as triggered by the automatic application of a social norm to a mistakenly perceived situation.

Rather, the data suggest that social norms, although being nurtured in centuries of evolution through tight, face-to-face social contact and repeated interaction, possess nonetheless a property of `abstractness' that cause them to emerge even in the social vacuum of the laboratory environment among pools of perfect strangers; hence, a responder in the ultimatum game might `rationally' decide to punish an unfair offer either out of a purely visceral reaction of anger, or also with the purpose to teach the other person to adhere to a norm of fairness in the future, even though most likely he won't be the direct beneficiary of such a change in attitude (though he may indirectly benefit from it, as long as fairness norms are increasingly adopted by members of a population). The fact that in multi-person public goods or dictator games, indirect punishment is observed (i.e., punishment by a third party who is not at all personally affected by the unfair act of the person being punished) is a further indication of the fact that – probably – the one-shot interaction of the lab is taken just as a temporal episode of a much larger game among members of a society.

Among the more recent studies on the ultimatum game, the work by Henrich et al. (2001a, 2001b; see also Bowles et al. 2003) stands out as particularly rich and innovative. A pool of economists and anthropologists has conducted experiments on the trust, ultimatum and dictator games with subjects belonging to 15 different small-scale societies around the world. The societies were carefully selected so as to present very heterogeneous characteristics in terms of economic organization and social and cultural traits. The basic results of this cross-cultural study can be summarized as follows: first, the canonical prediction of self-interested behaviour is never observed in any of the societies investigated. Second, there is a considerably higher variability across groups than in any previous cross-cultural study of this type. The behaviour of the proposers in the UG vary widely in the sample populations, with mean offers ranging from 26 to 58 per cent, and modal offers from 15 to 50 percent. Similarly, behaviours of the respondents in the sample include rejections of offers larger than 50% of the pie, and acceptance of offers below 20%. Third, a large part of the cross-group variability can be explained by society-level differences in the degree of economic organization, as measured by two key variables: 1) the magnitude of payoffs to cooperation, and 2) the degree of market integration. The first variable measures the extent to which economic production within a society is based on group activities that require the cooperation of members beyond the family domain (such as for example, whale-hunting). If such cooperative activities are largely responsible for the surplus produced, these societies must necessarily develop norms that govern surplus sharing. The second variable measures the extent to which members of the society rely on some form of exchange. The larger the diffusion of exchange habits, the higher should be the development of abstract sharing norms that hold with kins and strangers alike.

With regard to the behaviour of proposers in the UG, the authors find that offers are positively and significantly correlated with the two variables, which can account for roughly 60 per cent of the observed between-group variability in behaviour, while individual-level variables such as sex, age, income do not seem to bear any explanatory power. Hence, willingness to cooperate and share seems to be positively correlated with the intensity of market exchange and cooperative economic production existing within a society. These findings are remarkable for several reasons; first, they suggest that individuals responded to the new situation of the experiment by comparing it to corresponding familiar situations in their daily lives, and acted accordingly. In other words, people seem to have brought their own life experience into the lab, which should be good news for all experimental economics, since it suggests that the behaviour we generally observe in our experimental laboratories is after all not very distant from what happens in the `real' world. In other words, these results increase our confidence in the `ecological' validity of the experiments.

Secondly, the results support the hypothesis that preferences are endogenously generated through the influence of economic and cultural institutions and are not an exogenous datum, as mainstream economics usually postulates. Third, the development of market institutions, maybe counter-intuitively, seems to foster rather than inhibit fairness and altruism-based norms of behaviour.

4. Social rationality: experimental game theory

While the first two sections have briefly reviewed contemporary research issues on preferences, here we will telegraphically highlight some of the key issues that are currently being investigated in the domain of social interaction as exemplified and modelled by game theory.

As it is well known, standard game theory (by standard we mean game theory as it is explained in advanced textbooks) is a highly formal mathematical language used to describe the behaviour of super-rational individuals in strategic contexts. Although being born as a branch of applied mathematics and having been originally developed with the intention of making it the science of military conflict, its diffusion within economics has been rapid and overwhelming, and related fields of the social sciences like law and political science have recently began to make an extensive use of it. While nowadays theoretical and applied I.O., contract theory, microeconomics, and even individual decision theory as applied to inter-temporal choice are completely pervaded by subtle game-theoretic nuances and increasingly sophisticated and obscure equilibrium concepts, it is quite interesting to note that, already in the 40s, several military officers at RAND corporation, where game theory had had its initial moments of glory, soon started to be absolutely dissatisfied with its developments, and began to consider its whole analytical apparatus simply irrelevant as an aid to solve complex international policy problems. Philip Mirowski (2002), in his historical reconstruction of the intellectual and political fervour of those years at RAND, reports a couple of

quite illuminating statements. Alex Mood, responsible for the game theory section at RAND, wrote that the theory of games "has developed a considerable body of clarifying ideas and a technique which can analyze quite simple economic and tactical problems…these techniques are not even remotely capable, however, of dealing with complex military problems" (qu. in Mirowski, 2002: 323). In a similar spirit, Edward Quade stated:
"Game theory turned out to be one of these things that helped you think about the problem and gives you maybe a way of talking about the problem, but you can't take one of these systems analysis problems and set it up in a game and solve the game. Only in the very simplest cases would you actually be able to get a game theory solution. It gives you a background of how to think about a problem. I find discussions like the discussion of the Prisoner's Dilemma—its very helpful in illuminating your thought but you don't solve the problem that way, you know." (interview with Edward Quade, qu. in Mirowski, ibid.: 329).

What was already obvious to military officers fifty years ago seems yet not so obvious to mainstream economists fifty years later: namely, that game theory can be a useful tool to analyze strategic interaction problems in a novel and rigorous way, but its solution concepts are largely irrelevant except for the simplest games (and even here there might be disagreement, as Quade's opinion about the way to "solve" the Prisoner's Dilemma suggests).

It was Thomas Schelling's book on *The strategy of Conflict*, published in 1960, which rejuvenated the interest in the theory of games as something that could be actually applied to real strategic problems (Mirowski, 2002). Interestingly, Schelling's book does not contain a single formula; instead, it is full of semantically poignant concepts (`focal point' being the prominent example) and hardly uses any game-theoretic tools beyond the payoff-matrix.

Experimental economics as applied to strategic interaction has contributed significantly to putting life back into game theory, thus increasing its descriptive relevance, its predictive capabilities and its prescriptive value. Basic game theory, in fact, can be quite usefully applied as an abstract language that permits to classify the infinite variation of socio-economic interactions between two or more individuals on the basis of a few basic features (see also Camerer (2003) on this point). In Schelling's words: "the greatest contribution of game theory to social science in the end is simply the payoff matrix itself that, like a truth table of a conservation principle or an equation, provides a way of organizing and structuring problems so that they can be analyzed fruitfully and communicated unambiguously (qu. in Mirowski, ibid.: 330). Experimental game theory has taken the basic tools of game theory and has substituted most of its arcane equilibrium concepts with `reasonable' predictions inferred from the experimental evidence.

To summarize in a nutshell some of the developments in this area: it is by now largely acknowledged even by the most orthodox defenders of equilibrium concepts that, assuming players interacting in a game do play consistently with Nash equilibrium, they do so only after a process of learning and not right from the start. Literature on learning in repeated games has literally burst in the last decade, and several competing models have been tested against each other in comparison with experimental data (see, e.g., Fudenberg and Levine, 1998; Camerer and Ho, 1999). The two main families of learning models are reinforcement learning, which assumes players base their choices of strategy in every round only on the basis of their past earnings' history and without forming beliefs about the behaviour of the opponent, and belief-based learning models, which on the contrary assume players keep track of the opponent's behaviour to formulate a best reply. Evidence is mixed, although reinforcement models seem to not capture learning patterns very well in context in which information about the behaviour of others is available, since human subjects seem to respond to a certain extent to the moves of other players when given the opportunity to observe them. Some hybrid models (Camerer and Ho, 1999; Camerer, Ho and Chong, 2002) which combine features from both families seem to fit data well in some classes of games. Literature on learning has contributed significantly to obtain important insights on out-of-equilibrium behaviour in interactive context, although it must be said that these studies rely on a somewhat restrictive notion of learning itself, implicitly or explicitly identifying it with ``adaptive behaviour'', i.e., the process by which people interacting with each other or with ``Nature'' modify their actions in the effort of - supposedly - improving their performance. Few would object to the fact that learning as the word is normally used in common language hints at a far more comprehensive set of cognitive and behavioural processes, of which the realm of adaptive behaviour just defined may constitute nothing but a small part. A university student who learns the principles of quantum physics, or a musician who learns to play the piano probably share very little with experimental subjects who "learn'" (or do not learn) to pick the best available action (amongst a fixed set) to respond to their opponents in a stationary world. In other words, learning processes in real life involve far more than just actions, but include changes in procedural routines as well as in (more or less conscious and elaborate) mental representations. A few experimental studies have recently started to address the issue. Merlo and Schotter (1999), for example, have shown that in repeated experiments, subjects tend to learn different aspects of the task depending on whether they receive an immediate payoff according to performance in each single trial, or, on the contrary, are let to play `for free' for a while in view of a single, final shot in which they get paid. In the latter case, subjects' performances substantially improve. The authors argue that the first payoff environment renders subjects more myopic and adaptive, while the second induces more farsighted and intelligent explorations in the space of solutions. Weber (2003) has studied learning in a dominance-solvable game in which subjects receive no feedback about other players' choices in previous rounds. Studying learning in no-feedback conditions is important in order to test whether subjects are able to change their behaviour over time on the basis of some abstract reasoning about the game's features, rather than simply by reacting to others' behaviour. Moreover, in many real world situations, feedback is missing or ambiguous, hence the importance of studying no-feedback learning in the lab.

Other important findings concern behaviour in coordination games. Coordination games are characterized by the existence of multiple equilibria, which makes predictions often indeterminate. Furthermore, some of the proposed equilibrium refinements seem to have a dubious validity. Collecting a large body of experimental evidence in this class of games thus seems particularly compelling in order to build empirically-sound theories of equilibrium selection. Coordination games can be roughly divided in two classes: those with equilibria that are payoff-equivalent for all players involved (like matching games or experimental market entry games) and those whose equilibria can be Pareto-ranked. The first experiments on the latter class of games were conducted in the early 90's by Van Huyck and his collaborators at Texas A&M University, and their first results generated a considerable stream of further experiments. Their findings are remarkable in that they demonstrate how easily coordination failure may arise in a simple setting in which players have full information about the game structure and payoffs. In particular, in their experiment on the median action game (Van Huyck et al. 1991), the end of game equilibrium was *always* determined by the historical accident of first period play, and it was *always* an inefficient equilibrium. Economists who deny the existence of strong forms of history-dependence in economic life (e.g., Margolis and Liebowitz, 1995) should pay attention to these results. In the experiments on the minimum effort game (Van Huyck et al. 1990) the limit outcome was always the most inefficient one.

Several other studies have investigated the conditions that might affect coordination in these games. Van Huyck, Battalio and Beil (1993) have imposed participation costs to players such as to rule out inefficient equilibria by forward induction. Berninghaus and Ehrhart (1998) have proved that an extended time horizon improves coordination in the minimum effort game. Keser, Ehrhart and Berninghaus (1998) have shown that the nature of interaction (local versus global) also matters in determining which equilibrium will be selected. Goeree and Holt (1999) have shown that changes in the payoff function such as reducing the costs of deviation from equilibrium play, substantially improve coordination in the minimum effort game. Allowing pre-play costless signaling was shown to increase efficiency in the median action game (Blume and Ortmann, 2000). In a similar game (Devetag, 2003), players were able to `hill-climb' the strategy space in a coordinated manner due to costly signalling by some of the players who chose to `teach' others the path to the efficient equilibrium.

Overall, these studies highlight that convergence to efficient equilibria in coordination games is far from trivial. At the same time, there are several conditions that improve the efficiency of the achieved coordination, either by increasing the individual incentive to engage in risky explorations, or by allowing players to influence each other's beliefs and actions through their choices.

While the evidence on coordination games with Pareto-ranked equilibria points at the ubiquity of coordination failure, experimental findings on games with Pareto-equivalent equilibria show on the contrary how coordination success can emerge in the aggregate almost `by magic' (Kahneman, 1988) out of heterogeneous, boundedly rational strategies that do not seem to appeal to any equilibrium principle. This invocation of magic and invisible wizards able to concoct the ingredients of exuberant irrational behavior in order to produce a rational potion is indeed a recursive feature of some literature in experimental economics, and Kahneman should be aware of illustrious precedence: 'Why is that human subjects in the laboratory violate the canons of rational choice when tested as isolated individuals, but in the social context of exchange institutions serve up decisions that are consistent (as though by magic) with predictive models based on individual rationality?' (Smith, 1991: 894). Vernon Smith, who conceived the program of experimental economics as designed to repair the averages of the rationality postulate, is right to call for the spell of magic to save the concept of equilibrium.

To be fair, it must be ascertained that both Kahneman and Smith frequently crossed borders of the that traditional concept of rationality. Kahneman, a behaviouralist, made a career emphasising the differences between *homo economicus* and the more pedestrian *homo sapiens*. And Vernon Smith argued, at his Nobel lecture, for a difference to be established between constructivist, i.e. deductivist and theoretical rationality according to the neoclassical pattern, and "ecological rationality" based on the understanding of emergent rules from social interaction and coordination in a smithian way (Smith, 2003).

Game theory explored several instances of these dilemmas. Market-entry games are one instance of such games in which n players must decide whether or not to enter a market with a certain fixed capacity c, where c is usually $< n$. If the number of entrants is above the market capacity, all entrants suffer losses, while if entrants are less than c, opportunities for gain are left unexploited. In some cases players who decide to stay out earn a fixed positive payoff. Several experiments have shown that the number of entrants in each period is remarkably close to the market capacity c, although outcomes vary widely depending on features of the experimental design such as, e.g., the parameters of the payoff function, group size, matching protocol (e.g., Zwick and Rapoport, 2000). In some cases, groups seem to self-organize so as to achieve a higher degree of coordination than that implied by the mixed strategy equilibrium solution (Bottazzi and Devetag, 2003). Recent experiments investigate coordination patterns in more complex environments, such as players interacting over a network with both exogenous and endogenous link formation (Corbae and Duffy, 2000). Other very innovative experiments have started to explore the emergence of linguistic coordination in complex tasks (Weber and Camerer, 2003). These last studies show the creative potential of experimental research, especially when it is implied as a means to discover behavioural patterns without being too constrained by the obsession of theory-testing.

5. What to make of the evidence?

The neoclassical economics' image of the economic agent is that of an individual who is perfectly able to state his preferences in a coherent manner for whatever pairs of goods possibly conceivable whose consumption can occur at any moment in time between now and the remote future; and who coldly accepts an extremely unfair treatment from another agent as long as there is some (however small) material reward to achieve. This is not meant to say that the traditional conception of perfect rationality has not evolved at all within the mainstream economics tradition; however, the core of the discipline by and large persists with the "business as usual" approach of relying on optimization techniques to model economic behaviour.

The evidence here discussed shows a radically different picture: real humans sometimes experience difficulties in making complicated tradeoffs between consumption of different goods at different moments in time, sometimes reveal inconsistent preferences and sometimes simply do not know how much to value things they are not familiar with and hence rely on all sorts of arbitrary clues. At the same time, real humans do not grow up in a social vacuum, but, on the contrary, are continuously engaged in all types of social interactions with other individuals and have developed relatively stable patterns of behaviour based on simple norms such as reciprocity and cooperation. Hence, they happen to punish unfair treatments, even though this may cost them something. Furthermore, they are not hyper-rational but they are capable, to some extent, to reason about the behaviour of other agents and devise, if not optimal, sensible behavioural heuristics that at times happen to produce remarkably efficient aggregate results (…but not always!).

All in all, these findings seem to vindicate years of contemptuous dismissal on the part of mainstream economists of the criticism coming from economic sociologists, namely the formers' obsession with methodological individualism and the belief that the economy could be studied in isolation from the other aspects of society. As it is well known, economics has proceeded in its own way by and large ignoring the critiques. However, the findings on the experiments on social preferences recall ideas and concepts that were already familiar to economic sociologists fifty years ago. For example, Polanyi extensively analyzed primitive economies and he defined `reciprocity' as one of the basic templates of economic action to be found in all societies (Polanyi, Arensberg and Pearson, 1957). Likewise, both Polany (Polanyi et al. 1957) and Granovetter (1985) long ago introduced the notion of `embeddedness' to refer to the need to consider economic transactions as occurring within a larger network of social relations. In some sense, therefore, it might be said that contemporary experimental economics is bringing economics back to the past of its sister social sciences. However, the innovation consists in the methodology implied, which combines the advantages of the experimental method with the rigorous analytical classification of social interactions provided by basic game theory. Furthermore, if economists have a lot to learn from psychologists and sociologists, the reverse is true as well, because experiments also demonstrate that individuals respond to *incentives* in choosing their behaviour. Far from mechanically applying pre-codified norms in a semi-automatic mode once a situation gets categorized in a certain way, individuals appear sensitive to the incentive structure of the situation they face. Hence, the key issue is not that of deciding between the two opposite extremes of *homo economicus* and *homo sociologicus*, or between methodological individualism and structural determinism, but precisely the unpacking of the individual evaluation procedure, in which incentive-driven behaviour combines with reason and emotions to produce decision outcomes.

The recent attempts at incorporating the empirical evidence on preferences by means of preserving the standard framework of utility maximization (e.g., models of intertemporal choice that utilize hyperbolic discounting, utility functions that incorporate reciprocity (Rabin, 1998) and `inequality aversion (Fehr and Schmidt, 1999) do not look convincing for manifold reasons: firstly, they appear to assume the problem away by simply reverting to the canonical maximization paradigm with a mere modification of the utility function's argument. This is an old problem in discussions about rationality and self-interest in economics; advocates of the rationality paradigm often argue that postulating that agents maximize simply amounts to saying that they do the best they can to choose their preferred alternative among their `perceived' option set. And assuming self-interest merely implies that if an agent behaves altruistically, then it is in his self-interest to do so. However, this view of rationality and self-interest deprives the two notions of any content, and renders them simply unfalsifiable (see Vriend, 1996, for a discussion). Secondly, these models can generally only account for part of the evidence, and some can be as easily contradicted by experiments as the more traditional ones. Hence, one might still rely on the "as if" assumption in modelling economic behaviour in terms of utility maximization and self-interest as long as the predictive power that can be gained this way is such to justify the short-cut. When predictions are poor, it might be worthwhile to proceed all the way down and inductively search for empirical regularities.

6. Finding patterns of structure and evolution: economics back to the Cambridge tradition

Economics evolved and changed, how much strange this may appear to those unadvertised. As the first generation of the neoclassical revolution (1871-4) was paradoxically reasserted by the Cowles general equilibrium program and the rise of econometrics for structural estimation, the canon was slowly displaced to the radical affirmation of rational expectations. And, when the unchallenged beauty of the *homo economicus* was parading, including in game theory, experimental and computational economics emerged as new icons for a fin-de-siècle economics turned more empirical and realistic. This recent movement was partially surveyed in this paper.

It was not, of course, a smooth and uniform movement. But it was intentional. When Merrill Flood, the Robot Sociologist – as Mirowski calls him – and inventor of the Prisoner's Dilemma, invited Herbert Simon to Rand, in 1952, he provided the framework and the intellectual motivation for a research on the Turing machines as main tools for simulation. At the same time, von Neumann was proposing his theory of automata to represent social institutions, and in particular to conceive of markets as evolving computational entities. Both alternatives were contradictory with neoclassical economics and its very mode of theorizing, and the authors understood the effect of their deeds. Simon explained, quite candidly, that "[when Allan Newell and I began discussing the prospect of simulating thinking in the summer of 1954, it was due to] my longstanding discontent with the economists' model of global rationality, which seemed to me to falsify the facts of human behavior and to provide a wholly wrong foundation for a theory of human decision making" (Simon, 1979: ix). And von Neumann did not even try to hide how much he despised neoclassical economics.

Indeed, both Simon and von Neumann provided tentative escapes from the game-theoretical universe of the infinite reproduction of PD as the core of the investigation on the genetic code of society. In his 1984 book, "The Evolution of Cooperation", Axelrod noticed the contamination of research, since "The iterated PD has become the *E. Coli* of social psychology" (Axelrod, 1984: 28), and of course this applies to economics as well. Simon and von Neumann, each in a different sense, departed from game theory as mimicking experimentation and reinterpreted the social dimension of economic behaviour through their simulacra and automata.

The inductive approach we are arguing for values these precursors and these traditions. Indeed, the longstanding norm of selection of legitimate theories as those that could provide a formal and complete representation of the universe through a self-contained model, proved infectious in economics. Binmore denounced one of the implications of this sort of theorizing: "Not only are abstractions introduced that do not necessarily admit an operational referent; at the same time, operationally relevant factors are abstracted away together. This, in itself, is not necessarily invidious. However, the Bourbakist ethos makes it inevitable that factors that are not taken into account of formally are not taken into account at all" (1987: 183).

Abandoning this representation of the action of an agent as a vector in commodity space, several economists suggested instead to model the agent as a Turing machine, following Simon (Velupillai, 2002), or the institutions as algorithms (Scarf, 1989), in any case as information-processors. Along the same path, Nelson and Winter had long ago suggested an economic model mimicking some biological processes of selection and creation of variation, insisting that processing information and taking decision implies a non-mechanical approach (Nelson, Winter, 1982: 68). Although many different alternatives are envisageable,[2] it is in the convergence of experimental and computational economics that new fields of research, both empirically oriented and theoretically supported, are being suggested.

It is true that, for some, this movement is just a reincantation of the past. Kyndland and Prescott, for instance, reconstruct the NBER-Cowles debate in order to argue that neoclassical economists should not have criticized Mitchell and his followers for lack of theory, since it is convenient to measure without theory (Kyndland, Prescott 1990:3) – just as, one might say, calibration is measurement without reality. It is beyond the scope of this paper to comment on this topic,[3] but it nevertheless should be noted that an inductive approach aims at obtaining information about the structure and evolution of the economies in real life, and should preferably not be deviated to statistical demonstrations of pure artefacts.[4]

While considering these developments, Velupillai registers the failure of identification of conservation principles in economics and in social sciences, in contradiction to what happens in natural sciences (Velupillai, 2000: 65). Yet, the study of forms of communication and processing of information, of the architecture of institutions and in particular of the rules and regulations forming markets, suggests that coordination, the maximization of structural stability, is both the relevant subject for economics and empirically identifiable. In this framework, coordination refers not to a state of equilibrium, since it is an evolving and adaptive process, but to a dynamics of routines, of conservation as well as of transformation of societies. In that sense, experiments are relevant, since they may provide comparisons on the cultural and social value of information, on group and individual learning under different conditions, i.e. the processing of decision making – the Turing machines and their human complex counterparts.[5] Consequently, experimental economics promises an inductive approach to economics in the rigorous sense of an empirically supported research that is not bounded by restrictive notions of social interaction.

But there is also a second possible meaning for an inductive approach. Computational simulation of complex behaviour of difficult models is now feasible, what undoubtedly enlarges the role of models since permitting more interaction among variables, or nonlinearities. Considering that the models are generally underdetermined in relation to the social features they describe, this is a major progress. Furthermore, computational simulation permits also different statistical approaches to identified processes, using induction in a novel way, for instance challenging the non-critical

[2] Mirowski argues that the main contenders are the visions propounded by Kenneth Judd, Alain Lewis, Herbert Simon, Daniel Denett, and of course his hero John von Neumann (Mirowski, 2002: 523 f.).

[3] One of us developed a criticism of the RBC literature (Louçã, 2003).

[4] Another implication of an inductive approach is emphasized by Velupillai, who argues that a modern theory of induction may provide an alternative to the stochastic approach introduced with the econometric revolution pioneered by Haavelmo in economics (Velupillai, 2000: 71 fn).

[5] Bowles and a number of other researchers developed several experiments testing these cultural processes, giving experimental economics an anthropological dimension it badly needs (Bowles et al., 2002, 2003).

generalization of probabilistic statistics and using alternative statistics based dynamic geometrical and topological measures built on the structure of correlations. Computational economics explores new ground, for instance a novel approach to dynamics, described as the evolution of geometric relations of time series data (e.g., Vilela-Mendes, Araújo, Louçã, 2003), highlighting the changing patterns in the structure of the markets and how does it react to extreme events, with are part of the story and not undesirable outliers, as it may be the case for conventional smoothing statistics. Powerful descriptive statistics may be built on these assumptions, and even probabilistic generalizations are possible under disciplined conditions.

There are several reasons for exploring these new inductive methods. One is the evidence of widespread nonlinearities in economic series, a trivial hypothesis considering the nature of social interaction and structures, which can be empirically and statistically confirmed (Louçã, 1997). Indeed, complex nonlinear relations may emerge even in very simple models whether there are increasing returns or other externalities, from the simple fact of some agents being price settlers and not price takers, from the dynamics of learning, from complex dynamics in preferences and technologies and still from the lag structure itself (Brock, Hsieh, LeBaron, 1991: 32-4) – and models are timid replications of reality, in which all these factors may combine. Furthermore, this is not new: there is a long and respectable tradition in economics, although somewhat misregarded, dealing with nonlinearities, from Keynes' psychological laws and organic unity to Schumpeter's deviant behaviour of entrepreneurs. Nonlinearities are everywhere, simply because they are the model representation of social complexity: "No doubt, if our means of investigation should become more penetrating, we should discover the simple under the complex, then the complex under the simple, then again the simple under the complex, and so on, without our being able to foresee what will be the last term (Poincaré, 1903: 132).

But this is no trivial task. For ages, equilibrium became a theory of theories and a paradigm of paradigms: it defined the legitimate models of proof and theorising, it imposed the choice of techniques and the dominance of abstract reasoning and deduction of the general laws. This approach ended in general failure, namely the analysis it provided for crises and change, irregularities and interdependency. Consequently, the lack of a deterministic, complete and definitive knowledge has been felt by many to be a loss, in the context of the shipwreck of positivism. This paper has argued precisely the reverse: such an evolution of economic science, now in more turbulent waters, simply brings it back to the investigation of reality. Nothing is lost and everything is transformed, since that is the condition of the world: stable and mutant, unstable and structural, an indeterministic world of which the scientist is a part. Ignorance is certainty about certainty: the wisdom which knows that it only knows something is wiser than that which claims to be able to know everything.

A return to the Cambridge tradition is certainly part of this process of redefinition of economics. And, since finding patterns of structure and evolution is the very purpose of science, economics may be at the brink of a new era, with the combined contributions of experimental and computational economics, permitting the simulation of new generations of models aimed at the description of real markets and social choice. Yet, for long, the understanding of human agency escaped this intention. Changing the map of social sciences, experimental and computational economics may now allow for a new inductive, empirically oriented and theoretically renewed perspective. But only the future will answer this query: is it possible to rebuild economics as a sensitively realistic social science? Our answer is a commitment.

References

Abel, Andrew (1990), 'Asset prices under habit formation and catching up with the Jones', *American Economic Review*, 80, 38-42

Allais, Maurice (1953), 'Le comportement de l'homme rationnel devant le risqué, critique des postulats et axioms de l'ecole Americaine', *Econometrica*, 21, 503-546

Ariely, Dan, Loewenstein, George, Prelec, Drazen (2003), '"Coherent Arbitrariness": Stable Demand Curves without Stable Preferences', *Quarterly Journal of Economics*, February, 73-105

Axelrod, R. (1984), *The Evolution of Cooperation*, New York: Basic Books

Benzion, Uri, Rapoport, Amnon, Yagil, Joseph (1989), 'Discount rates inferred from decisions: An experimental study', *Management Science*, 35, 270-284

Berg, Joyce, Dickhaut, John, McCabe, Kevin (1995), 'Trust, reciprocity, and social history', *Games and Economic Behavior*, 10, 122-142

Berninghaus, S.K., and Ehrhart, K.M. (1998), 'Time Horizon and Equilibrium Selection in Tacit Coordination Games: Experimental Results', *Journal of Economic Behavior and Organization*, 37, 231-248

Binmore, Ken (1987), 'Modelling Rational Players', *Economics and Philosophy*, 3: 179-214

________ (1999), 'Why Experiment in Economics?', *Economic Journal*, 109: F16-24

Blume, A., Ortmann, A. (2000), 'The Effects of Costless Pre-play Communication: Experimental Evidence from a Game with Pareto-ranked Equilibria', Charles University Center for Economic Research and Graduate Education working paper

Boldrin, Michele, Christiano, Lawrence, Fisher, Jonas (2001), 'Habit persistence, asset returns, and the business cycle', *American Economic Review*, 91, 149-166

Bottazzi, Giulio, Devetag, Giovanna (2003), 'A Laboratory Experiment on the Minority Game, *Physica A*, 324, 1-2, 124-132

Bowles, Samuel, Choi, Jung-Kyoo, Hopfensitz (2002), *The Coevolution of Individual Behaviors and Social Institutions*, ISF working paper

Bowles, Samuel, Gintis, Herbert (2002), 'The Inheritance of Inequality', *Journal of Economic Perspectives*, 16: 3, 3-30

Bowles, Samuel, Henrich, Joe; Boyd, Robert; Camerer, Colin; Fehr, Ernest; Gintis, Herbert (Eds.) (2003), *Cooperation, Punishment and Self-Interest*, Oxford: Oxford University Press

Brock, William (2002), 'Some Santa Fe Scenery', in David Colander (Ed.), *The Complexity Vision and the Teaching of Economics*, New York: John Wiley & Sons, 29-49

Brock, W., Hsieh, D., LeBaron, B. (1991), *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*, Cambridge, US: MIT

Camerer, Colin (1995), 'Individual Decision Making', in A. Roth, J. Kagel (Eds.), *The Handbook of Experimental Economics*, Princeton: Princeton University Press

_______ 2003), *Behavioral Game Theory*, Princeton: Princeton University Press

Camerer, Colin, Fehr, Ernst (2002), 'Measuring Social Norms and Preferences using Experimental Games: A Guide for Social Scientists', working paper University of Zurich (forthcoming in Henrich, Boyd, Bowles, Camerer, Fehr, Gintis and McElreath (Eds.), *Foundations of Human Sociality: Experimental and Ethnographic Evidence from 15 Small-Scale Societies*)

Camerer, Colin, Ho, Tech (1999), 'Experience-weighted attraction learning in normal form games', *Econometrica*, 67, 827-874

Camerer, Colin, Ho, Tech, Chong, Kuan (2002), 'Sophisticated experience-weighted attraction learning and strategic teaching in repeated games, *Journal of Economic Theory*, 104, 137-188

Caplin, Andrew, Leahy, John (2001), 'Psychological expected utility theory and anticipatory feelings', *Quarterly Journal of Economics*, 166, 55-79

Chamberlin, Edward (1948), 'An Experimental Imperfect Market', *Journal of Political Economy*, 56:95-108

Corbae, Dean, Duffy John (2002), 'Experiments with Network Economies', University of Pittsburgh working paper

Devetag, Giovanna (1999), 'From Utilities to Mental Models: A Critical Survey on Decision Rules and Cognition in Consumer Choice', *Industrial and Corporate Change*, 8(2), 289-351

_______ 2003), 'Coordination and Information in Critical Mass Games: An Experimental Study', *Experimental Economics,* 6: 53-73

Ellsberg, Daniel (1961), 'Risk, Ambiguity and the Savage Axioms', *Quarterly Journal of Economics*, 75: 643-69

Elster, Jon (1985), 'Weakness of will and the free-rider problem', *Economics and Philosophy*, 1, 231-265

Fehr, Ernst, Fischbacher, Urs (2001), 'Retaliation and Reputation', mimeo, University of Zurich

Fehr, Ernst, Schmidt, Klaus (1999), 'A Theory of Fairness, Competition and Co-operation', *Quarterly Journal of Economics*, 114, 817-868

Frederick, Shane, Loewenstein, George, O'Donoghue, Ted (2002), 'Time Discounting and Time Preference: A Critical Review, *Journal of Economic Literature*.

Freeman, Christopher; Louçã, Francisco (2002), *As Time Goes By*, Oxford: Oxford University Press

Fudenberg, Drew, Levine, David (1998), *The Theory of Learning in Games*, Cambridge, Mass.: MIT Press

Goeree, J.K., Holt, C.H. (1999), 'An Experimental Study of Costly Coordination', University of Virginia, Department of Economics working paper

Granovetter, Mark 91985), 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology*, 91: 481-510

Grether, D., Plott, C. (1979), 'Economic theory of choice and the preference reversal phenomenon, *American Economic Review*, 69, 623-638

Hahn, Frank (1994), "An Intellectual Retrospect", *Banca Nazionale del Lavoro Quarterly Review*, V48: 245-58

Henrich, Joseph, Boyd, Robert, Bowles, Samuel, Camerer, Colin, Fehr, Ernst, Gintis, Herbert, McElreath (2001a), 'In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies', Amer. Econ. Rev. Papers and Proceedings, May, 73, 78

Henrich, Joseph, Boyd, Robert, Bowles, Samuel, Camerer, Colin, Fehr, Ernst, Gintis, Herbert, McElreath (2001b), 'Cooperation, Reciprocity and Punishment in Fifteen Small-Scale Societies', working paper, Santa Fe Institute

Hirschman, Albert (1977), *The Passions and the Interests*, Princeton: Princeton University Press

Hsee, Christopher, Abelson, Robert, Salovey, Peter (1991), 'The relative weighting of position and velocity in satisfaction', *Psychological Science*, 2:4, 263-266

Jermann, Urban (1998), 'Asset pricing in production economies', *Journal of Monetary Econo*mics, 41, 257-275

Kagel, John, Roth, Alvin, (1995), *The Handbook of Experimental Economics*, Princeton: Princeton University Press

Kahneman, Daniel (1988), 'Experimental economics: A psychological perspective', In R. Tietz, W. Albers, R. Selten (Eds.), *Bounded Rational Behavior in Experimental Games and Markets*, New York: Springer-Verlag, 11-18

Kahneman, D., Knetsch, J., Thaler, R. (1991), 'Anomalies: The Endowment effect, loss aversion, and status quo bias', *Journal of Economic Perspectives*, 5, 193-206

Kahneman, Daniel, Tversky, Amos (1979), Prospect Theory: 'An Analysis of Decision Under Risk', *Econometrica*, 47:2, 263-291

Kahneman, D., Wakker, P.P., Sarin, R. (1997), 'Back to Bentham? Explorations of experienced utility', *Quarterly Journal of Economics*, 112, 375-405

Keser, C., Ehrhart, K.M., and Berninghaus, S.K. (1998), 'Coordination and Local Interaction: Experimental Evidence', *Economics Letters*, 58, 269-275

Kyndland, F., Prescott, E. (1990), "Business Cycles: Real Facts and a Monetary Myth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 3-18

Laibson, David (1994), 'Essays in hyperbolic discounting', Ph.D. dissertation, MIT

Laibson, David (1997), 'Golden eggs and hyperbolic discounting', *Quarterly Journal of Economics*, 112, 443-477

_______ 1998), 'Life-cycle consumption and hyperbolic discount functions', *European Economic Review, 42, 861-871*

Lichtenstein, Sara, Slovic, Paul (1971), 'Reversals of preference between bids and choices in gambling decisions', *Journal of Experimental Psychology*, 89(1), 46-55

Loewenstein, George (1987), 'Anticipation and the Valuation of Delayed Consumption', *Economic Journal*, 97, 666-684

_______ (1999), "Experimental Economics from the Vantage-point of Behavioural Economics", *Economic Journal*, 109, F25-F34

Loewenstein, George, Prelec, Drazen (1992), 'Anomalies in intertemporal choice: Evidence and an interpretation', *Quarterly Journal of Economics*, May, 573-597

Loewenstein, George, Weber, Roberto, Flory, Janine, Manuck, Stephen, Muldoon, Matthew (2001), 'Dimensions of time discounting', paper presented at conference on Survey Research on Household Expectations and Preferences, Ann Arbor, MI, Nov. 2-3, 2001

Louçã, Francisco (1997), *Turbulence in Economics*, Cheltenham: Edward Elgar

_______ (2003), "Swinging all the Way – The Education of Doctor Lucas and Foes", *History of Political Economy*, forthcoming

Margolis, Stephen, Liebowitz, S.J. (1995), `Path Dependence, Lock-in and History', *Journal of Law, Economics and Organization*, Summer, 205-26

Mellers, B.A., Schwartz, A., Cooke, A.D.J. (1998), 'Judgment and Decision Making', *Annual Review of Psychology*, 49, 447-477

Merlo, Antonio, Schotter, Andrew (1999), 'A Surprise-Quiz View of Learning in Economic Experiments', *Games and Economic Behavior*, 28, 25-54

Mirowski, Philip (2002), *Machine Dreams – Economics Becomes a Cyborg Science*, Cambridge: Cambridge University Press

Morgan, Mary (2001), "The Curious Case of the Prisoner's Dilemma: Model Situation? Exemplary Narrative?", Research Memoranda 7-2001, Faculty of Economics and Econometrics, University of Amsterdam

Nelson, Richard; Winter, Sidney (1982), *An Evolutionary Theory of Economic Change*, Cambridge: Harvard University Press

O'Donoghue, Ted, Rabin, Matthew (1999), 'Incentives for procrastinators', *Quarterly Journal of Economics*, 114:3, 769-816

Poincaré, Henri (1903), *Science and Hypothesis*, New York; Walter Scott Pub.

Polanyi, Karl, Arensberg, Conrad M., Pearson, Harry W. (1957), *Trade and Market in the Early Empires*, New York: The Free Press

Rabin, Matthew (1998), 'Incorporating fairness into game theory and economics', *American Economic Review*, 83, 1281-1302

Roth, Alvin (1995), "Introduction", in A. Roth, J. Kagel (Eds.), *The Handbook of Experimental Economics*, Princeton: Princeton University Press

Roth, Alvin, Kagel, John, Ed. (1995), *The Handbook of Experimental Economics*, Princeton: Princeton University Press

Rubinstein, Ariel (2000), 'Is it "economics and psychology"? The case of hyperbolic discounting', Tel Aviv University and Princeton University discussion paper

Samuelson, Paul (1937), 'A note on measurement of utility', *Review of Economic Studies*, 4, 155-161

Scarf, H. (1989), "Mathematical Programming and Economic Theory", Cowles Foudation Discussion Paper, 930

Schelling, Thomas (1960), *The Strategy of Conflict*, Cambridge: Harvard University Press

_______ (1978), *Micromotives and Macrobehaviors*, New York: E.W. Norton

Simon, Herbert (1979), *Models of Thought*, New Haven: Yale University Press

Smith, Vernon (1962), "An Experimental Study of Competitive Market Behavior", *Journal of Political Economy*, 70: 111-137

_______ (1991), "Rational Choice: The Contrast Between Economics and Psychology", *Journal of Political Economy*, 99 (4), pp. 877-897

_______ (2002), 'Method in Experiment: Rhetoric and Reality', *Experimental Economics*, 5(2), 91-110

_______ (2003), "Constructivist and Ecological Rationality in Economics", *American Economic Review*, 93: 3, pp. 465-508

Thaler, Richard (1981), 'Some empirical evidence on dynamic inconsistency', *Economic Letters*, 8, 201-207

Tversky, Amos, Kahneman, Daniel (1974), 'Judgment under Uncertainty: Heuristics and Biases', *Science*, CLXXXV, 1124-1131

Van Huyck, J., Battalio, R., and Beil, R. (1990), 'Tacit Cooperation Games, Strategic Uncertainty, and Coordination Failure', *American Economic Review*, 80, 234-248

Van Huyck, J., Battalio, R., and Beil, R. (1991), 'Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games', *Quarterly Journal of Economics*, 106, 885-910

Van Huyck, J., Battalio, R., and Beil, R. (1993), 'Asset Markets as an Equilibrium Selection Mechanism: Coordination Failure, Game Form Auctions, and Forward Induction', *Games and Economic Behavior*, 5(3), 485-504

Velupillai, Kumaraswamy (2000), *Computable Economics*, Oxford: Oxford University Press

Vilela-Mendes, Rui; Araújo, Tânya; Louçã, Francisco (2003), "Reconstructing an Economic Space from a Market Metric", *Physica A*, 323: 1-2, pp. 635-50

Von Neumann, John, Morgenstern, Oskar (1964), *Theory of Games And Economic Behavior*, 1944. Reprint, New York:Wiley

Vriend, Nicolaas (1996), 'Rational behavior and economic theory', *Journal of Economic Behavior and Organization*, 29, 263-285

Weber, Roberto (2003) `Learning' with no feedback in a competitive guessing game', *Games and Economic Behavior*, in press

Weber, Roberto, Camerer Colin (2003), 'Cultural conflict and merger failure: An experimental approach', *Management Science* 49(4): 400-415.

Zwick, Rami, Rapoport, Amnon (2002), 'Tacit Coordination in a Decentralized Market Entry Game with Fixed Capacity', *Experimental Economics*, 5(3), 253-272