Experiments, Theory – and Reality?

Bruno S. Frey and Iris Bohnet*

1. Introduction

Not long ago, it was taken to be a matter of course that experiments are impossible in economics. This view has regularly been proclaimed in textbooks and was reiterated in the respective heading of the Encyclopedia Britannica as late as 1991 (p. 395): "...there is no laboratory in which economists can test their hypotheses". Over the last few years, however, the situation has completely changed: experiments are often used and their potential is highly praised. And rightly so. Game theory, market equilibrium models and theories of choice under uncertainty have greatly benefited from the insights gained in experiments. Theorists should better understand in how far their logically correct assumptions and implications are met by the behavioural evidence provided by experiments. In view of this success story, it may now be time to take a look at experiments from a different perspective. In this paper, we wish to critically evaluate whether the present-day experimental movement also copes with 'real world problems', i.e. with issues which might have not been addressed by theory yet.

Theory-driven experiments should be complemented by problem-driven research. This does not imply giving up theory but changes the focus of attention. The following real-life as well as theoretical puzzle might help in understanding our point of view. In most European countries, the use of referenda is rejected because in a highly complex world, important decisions cannot be left to little motivated, little educated and little informed citizens. Even in Switzerland, with its tradition of direct democracy, it has repeatedly been demanded that the scope for popular referenda and initiatives be curtailed. The political issue thus lies on the table. The challenge to social scientists and, in particular, to political economists, is to investigate whether these views are warranted. How can we – theoretically and empirically – rely on consumer sovereignty in the market and at the same time refuse citizens' sovereignty in politics?

Interestingly, the supposedly incapable citizens have taken decisions which – to say the least – have not prevented the most direct democracy in the world, Switzerland, from becoming one of the richest and best organized countries. Voters who are expected to be ignorant from a theoretical point of view (Downs 1957), participate in political decision-making, demand information and engage in political debates. These institutionalized communication opportunities induced by referenda and initiatives are a crucial factor favouring socially beneficial collective decisions (Bohnet and Frey 1994). While it is possible to understand the role of political discussion by consulting historical and political accounts, it is extremely difficult to isolate the specific parameters determining a discussion's effects. Here experiments can close the gap.

We ran a series of experiments on cooperation (Prisoner's Dilemma) and on fairness (Ultimatum and Dictator Games) which capture particular features of the political process, and introduced verbal and non-verbal communication. As has been well established in the experimental literature (e.g. Ledyard 1995), verbal communication significantly increases cooperation in the one-shot Prisoner's Dilemma. More interestingly, we found that a component of verbal communication, mere identification, suffices to induce individuals to cooperate: Cooperation rises from 12% under anonymous conditions, to 23% with non-verbal communication, and to 78% when discussion is allowed. In the Dictator Game, non-verbal communication is even more important: Allocators share the money received equally as soon as they are identified while in anonymity, they only pass on 26% of their endowment. In contrast, we observed no such increase in fairness in the Ultimatum Game because an equal split is already induced by the recipient's possi-

* We are grateful for helpful suggestions to Reiner Eichenberger and Felix Oberholzer. The financial support of COST A7 is gratefully acknowledged.

1 See e.g. Möckli (1994) for Switzerland, and more generally, Butler and Ranney (1994).
goods games or with behavioural anomalies; those investigating decision theory, on the other hand, are not interested in market experiments.4

In so far as the research topics of experimental economics are theory-driven, 'relevance' is no prevalent criterion nor is it required that the experimental setting is related to real-world problems. If applied to reality at all, researchers tend to refer to the importance of distributional and fairness issues in economic policy in quite general terms.5 Such 'hand waving' stands in stark contrast to the great emphasis put on precision when experiments are designed and conducted. An indicator of this disinterest for real-life applications is the lacking integration of econometric studies into the respective fields. To our knowledge, e.g. none of the studies on Ultimatum or Dictator Games seeking to understand how fair individuals are, confronts the results with econometric estimates of fairness or altruism, e.g. in the form of charitable giving.6

A second consequence of experimental economics' take-off into a field of its own is the major emphasis put on technique. The mastery of technique tends to become a goal of its own.7 The 'value' of an experiment is increasingly judged according to how well the technique is applied, and most praise is received by those who use the most advanced and sophisticated research methods. There is a good reason for this, which has to do with the organization of international research. To evaluate the extent to which a scholar masters the techniques is possible on the basis of commonly accepted standards. Judging 'relevance', on the other hand, is subjective, and there are no handy criteria available by which to proceed. Technique thus serves as a signalling device on the academic labour market (see Frey and Eichenberger 1992).

---

4 The practitioners of these subfields rarely, if ever, cite each other which reflects the missing interaction well.
5 Exceptions are rare considering the boom in experimental economics. First steps to make experimental results politically relevant include the allocation of airport landing slots by the US-Civil Aeronautics Board and the auction of T-bills by the US-Department of Treasury (see Portney 1994: 15).
6 An interesting real life analogy to Ultimatum Games in the form of boycotts is reported by Camerer (1990: 315) who, however, does not analyze any field evidence in depth.
7 Technique is, of course, important. We do not argue for sloppy or soft experiments but for judging technical skills in terms of their adequacy for the problem at hand.

2.2 Internal and External Validity

The two aspects of 'self-defined problems' and 'self-serving technique' are related to psychologists' distinction between internal and external validity of experiments. When research is theory-driven, the experiments are difficult or impossible to apply to real-world problems. Thus, the external validity is low. Sophisticated techniques lead to high internal validity. While it would be preferable to reach maximum validity in both dimensions, such a situation is unattainable taken that the researchers involved act under the constraints discussed above. Limited resources, time and effort force the experimenters to invest most effort into a technical sophistication unrelated to real-life issues. It follows that there is a trade-off between the two dimensions of validity (Figure 1). We have argued that the differentiation into a field of its own and the internally driven research process going with it, leads to a position in the vicinity of B rather than of A.

![Figure 1: The relationship between internal and external validity](image)

As we have made clear, the trade-off between external and internal validity does not constitute a 'menu of choice'. The position reached is the result of the internal dynamics of academic scholarship which depend on the institutional conditions of scientific research.
The goal of this section has been to show that there are many experiments beyond the laboratory. Much is lost if economic experimentation is restricted to the latter arguing that only then influences can be tightly controlled. It should at least be acknowledged that this comes at a cost in terms of lower external validity.

4. Experiments and Endogenous Institutions

The advantage of experiments that 'the economic environment is very fully under the control of the experimenter' (Roth 1988: 974) has one important draw-back: The experimenter defines which parameters may be used for decision-making. We then know how individuals behave confronted with a well-specified problem and restricted by a clearly defined institutional setting. This is important when testing theory – we need to explore the limits of our theoretical approaches by varying institutional parameters. However, it does not help us in learning about reality, about how things really are. Omitted alternatives might be the variables decisive for problem-solving – or, thinking of Galileo Galilei, a pebble lying on a table does not fall. Nor are people always trapped in a prisoner's dilemma. As artificial as it sounds to put a light body on a much heavier body to investigate the movement of bodies, as unrealistic is it to force individuals into one-shot anonymous prisoner's dilemma games without exit possibilities to learn about how people behave in dilemma situations.

Human beings are creative problem-solvers. They may transform one-time interactions into repeated games or at least into decisions with sequential moves. Also, free-riders do not disappear in a crowd of anonymous others but they are identified or even talked to. Most of all, people avoid difficult, unproductive situations by institutionalising decision-making mechanisms allowing for coordination and cooperation. Many of these enlargements of individuals' possibility space have of course been experimentally tested.10 Almost no attention, however, has been given to the demand for alternative institutional mechanisms.

While the pebble could not tell Galileo that it cannot fall if the table is not taken away, a human being can state his/her preferred process to solve a problem. Survey results indicate that individuals have quite clear ideas about which decision procedures they like. They typically do not want pure chance (e.g. lotteries) to decide, they are sceptical about the price mechanism's capability to deal with surplus demand situations in a desired way, while they go for more democratic, more discursive mechanisms.11 A process-oriented view, therefore, does not so much stress an efficient outcome but rather efficient procedures.

Process-oriented experiments also confront individuals with well-specified problems but leave room for the evolution of institutions. The demand for communication varies according to the decision problem at hand and the involved cost (1.). If individuals are free to transform the structure of a game, they try to avoid prisoner's dilemmas by either exiting or by searching for a coordinating mechanism (2.).

4.1 The Demand for Communication

Verbal communication seems to be especially relevant in public good type settings while in bargaining games, it either does not affect the result at all (Ultimatum Game) or is not needed as non-verbal communication suffices to induce fairness in the Dictator Game, i.e. to divide the sum received equally. The incentive to free ride in the Prisoner's Dilemma is only overcome if the individuals involved communicate with each other and thereby are able to converge on a common, mutually advantageous cooperative strategy. Under anonymity as well as under identification such a strategy transcends the individuals' possibilities as a collective decision is needed in order to pursue it. Do people anticipate the potential of communication to enlarge their possibility space?

In our experiments, we asked subjects whether they preferred to play the three games with or without communication. When subjects are told that the experimental design allows for communication before they have to decide but that they may also play anonymously if they

---

10 For repeated games, see e.g. Isaac, McCue and Plott (1985) or Andreoni (1988); for sequential move games, Macy (1991); for the relaxation of anonymity by verbal and non-verbal communication, Frey and Bohaei (1995); for exit, Vanberg and Congleton (1992).

11 For an evaluation of the different decision-making mechanisms by survey, see Kahneman, Knetsch and Thaler (1986) and Frey and Pommerene (1993), and for an application to a real life siting problem, Oberholzer, Frey, Hart and Pommerene (1995).
In the Dictator- and the Ultimatum Games, it is again the economics majors who decide less often for communication. Perceived social pressure induces allocators to talk to their recipients indicating a need for verbal justification. This result is counter-intuitive as allocators expose themselves to potential social pressure exactly by talking to their recipients.16

4.2 Transforming the Decision Structure

Common pool resources are not always depleted and individuals often contribute to public goods. While Ostrom, Gardner and Walker (1994) forcefully demonstrate that informal institutions such as shared norms, exchanged promises and a group identity help in finding productive solutions, we focus on formal institutions. Public Choice scholars have long recognized the potential of social contracts to increase cooperation, ex ante as well as ex post (Mueller 1989, Frey and Kirchgässner 1994). Individuals are more likely to, ex ante, agree on rules of cooperation behind the veil of ignorance because it is only by such an agreement that a socially efficient provision of public goods may be guaranteed ex post. Covering all potential public goods by one contract increases the cost of defection. Not to cooperate in one specific instance would not only keep the public good in question from being efficiently supplied but would also destroy or at least endanger the social contract as such.

A nuclear waste disposal facility is such a public good (Oberholzer, Bohnet and Frey 1995). It is socially beneficial but locally very costly. Its completion is more likely if the society manages to transform an independent decision in a one-shot prisoner's dilemma into a dependent decision in a 'serial interaction prisoner's dilemma' (Granovetter 1978: 142). An individual community's calculus is thereby changed: The benefits of defecting, i.e. not having to bear the local cost, must then be weighted against the cost of defecting, i.e. the destruction of the socially beneficial contract. Provided a social contract covering locally unwanted but socially desired goods exists, a specific community destined to host the facility acts like a link in a chain. It learns from what other chain members did before when confronted with such a project. 'Serial interaction' or 'step-wise decision making' (Macy 1991) leads to the same conclusion as do repeated prisoner's dilemma games: individuals play 'tit for tat' (Axelrod 1984).

Direct democracy is able to produce such a structure: referenda include all the citizens to decide about the production of nuclear energy in the first place. It is not a dictator or some representatives in government and parliament to whom the buck may be passed but to each citizen individually. In 1990, the Swiss indeed included nuclear technology into their social contract by rejecting a proposition suggesting to abandon nuclear energy. Econometric evidence further shows that it was the people who approved of nuclear energy in the national referendum in 1990 who also supported the siting of the nuclear waste repository in their home town later on, in 1993.17 Thus, it seems possible to induce cooperation in a specific prisoner's dilemma if the game is not played in isolation but is part of a social contract including all the citizens.

One of the few series of experiments allowing for different processes to deal with public goods supports the importance of democratic decision-making. Wilke, Rutte and Wit (1986, in the Netherlands) and Messick and Samuelson (1986, in California) had subjects being part of groups of six harvest points from a replenishable common pool resource worth 300 points. Individuals were told that they could collect between 0 and 10 points during each trial until the resource was depleted or the experimenters stopped the experiment. The overuse condition was simulated by a replenishment rate of 3 % leading to the depletion of the resource in ten trials; the underuse condition with a replenishment rate of 17 % kept the resource at a level of 300 points in ten trials and induced individuals to suboptimally use the resource. In both conditions, the experiment was interrupted after ten trials and individuals were given the chance to transform the process of decision-making. Subjects could choose between continuing to collect points individually or delegating decision power to some elected group member who would then collect for all and allocate the points to the other

16 This may be taken as indicating that students did not believe in the total anonymity of the decisions and felt that verbal justification is better than no justification in case their anonymity were relieved - but this is pure speculation. For a critical discussion of the problem, see Hoffman, McCabe, Shachat and Smith (1992).

17 For the econometric analysis, see Oberholzer, Frey, Hart and Pommerehne (1995).
matters). The relationship to empirical reality becomes increasingly loose, and self-contained academic motivations take over.

(2) Experimental economics is defined in an increasingly narrow way and is confined to laboratory trials. There is no connection to social experiments which have some advantages in terms of external validity: as the latter are undertaken in real-life, and are addressed to policy issues, they make it easier to bridge academic research and reality.

(3) In order to strictly control the relevant parameters, and to therewith maximize internal validity, experiments normally allow individuals to react only in specific ways prescribed by the experimenter. Yet, human beings are not passive automatons but creative problem-solvers. Confronted with a particular experiment they use their resources to come to grips with these sometimes awkward situations. Most importantly, they resort to verbal and non-verbal communication to improve their position, use institutions in order to overcome or to transform the situation to their benefit, or rationally choose the exit option. Only ‘open’ experiments allow to study and evaluate such important reactions.

(4) With respect to the explanation of human behaviour, experimental economics departs from modern economics. The economic approach to human behaviour attributes changes in behaviour to changes in constraints. Institutions (in the sense used by Modern Institutional Economics) play a major role for explaining how individuals act because they systematically affect the constraints. Experimental economics, in contrast, tends to explain behaviour by referring to ‘types of persons’, and therefore does not empirically enrich the generalized relative price effect to which much of the recent success of economics as a social science can be attributed. This shortcoming can be remedied by putting more emphasis on ‘institutional experiments’.

References


Autorenverzeichnis

Prof. Dr. Otwin BECKER, Alfred-Weber-Institut für Sozial- und Staatswissenschaften der Universität Heidelberg, Grabengasse 14, D-69117 Heidelberg I.

Dr. Iris BOHNET, Institut für empirische Wirtschaftsforschung der Universität Zürich, Blümlisalpstr. 10, CH-8006 Zürich.

Prof. Dr. Friedel BOLLE, Europa Universität Viadrina, Wirtschaftswissenschaftliche Fakultät, Postfach 776, D-15227 Frankfurt (Oder).

Prof. Dr. Ronald BOSMAN, CREED/Department of Economics, Universität von Amsterdam, Roeterstraat 11, NL-1018 WB Amsterdam.

Dipl.-Volksw. Gregor BRÜGGELAMBERT, Universität GH Essen, Fachbereich Wirtschaftswissenschaften, Universitätsstraße 12, D-45141 Essen.

Prof. Dr. Ernst FEHR, Institut für empirische Wirtschaftsforschung der Universität Zürich, Blümlisalpstr. 10, CH-8006 Zürich.

Dr. Hans FEHR, Universität Tübingen, Fachbereich Wirtschaftswissenschaft, Mohlstraße 36, D-72074 Tübingen.

Prof. Dr. Bruno S. FREY, Institut für empirische Wirtschaftsforschung der Universität Zürich, Blümlisalpstr. 10, CH-8006 Zürich.

Dr. Simon GÄCHTER, Institut für empirische Wirtschaftsforschung der Universität Zürich, Blümlisalpstr. 10, CH-8006 Zürich.

Prof. Dr. James J. HECKMAN, University of Chicago, Department of Economics, 1126 E 59th St., Chicago, Ill. 60637, USA.

Dr. Maik HEINEMANN, Universität Hannover, Fachbereich Wirtschaftswissenschaft, Königsworther Platz 1, D-30167 Hannover.

Prof. Dr. Jeffrey A. SMITH, University of Western Ontario, Department of Economics, Social Science Center, London, ON N6A5C2, Canada.

Prof. Dr. Reinhard TIETZ, Steinhausenstr. 23, D-60599 Frankfurt (Main).

Prof. Dr. Wolfgang WIEGARD, Universität Tübingen, Fachbereich Wirtschaftswissenschaft, Mohlstraße 36, D-72074 Tübingen.

Prof. Dr. Frans van WINDEN, CREED/Department of Economics, Universität von Amsterdam, Roeterstraat 11, NL-1018 WB Amsterdam.

Prof. Dr. Geert WOLTJER, Universität Maastricht, Wirtschaftswissenschaftliche Fakultät, Postfach 616, NL-6200 MD Maastricht.