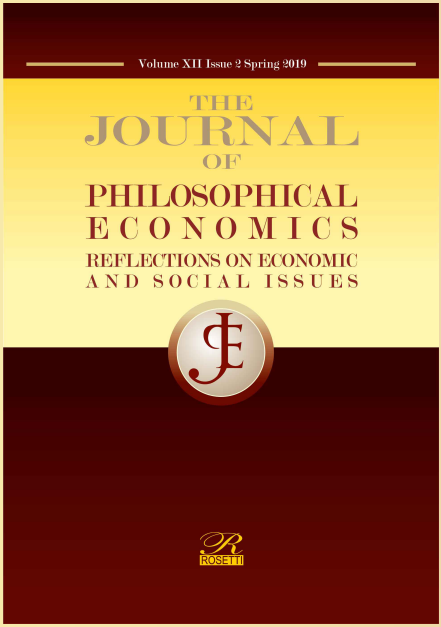
THE JOURNAL OF PHILOSOPHICAL ECONOMICS: REFLECTIONS ON ECONOMIC AND SOCIAL ISSUES

Volume XII Issue 2 Spring 2019

ISSN 1843-2298



Economic experiments versus physical science experiments: an ontology-based approach

María Caamaño-Alegre
José Caamaño-Alegre

ROSETTI

Economic experiments versus physical science experiments: an ontology-based approach

María Caamaño-Alegre

José Caamaño-Alegre

Abstract: This article applies an ontology-based approach to economic experiments, emphasizing their differences with respect to physical science experiments. To contextualize our discussion, a conciliatory Weberian view of the similarities and differences between natural and social sciences is provided. After that, some ontological features of the social sciences' domain are highlighted, together with their problematic effect on experimental economics. Specifically, we focus on human beings' representational capacities and intentionality, their cultural and conventionally mediated forms of social interaction, and the holistic openness, instability and uncertainty of the social world. Finally, we emphasize the severe under-determination of theory by evidence affecting social science, as well as the related problems of empirical ambiguity, confirmatory biases and propensity to pseudoscientific practices in experimental economics.

Keywords: social ontology, experimental validity, natural science, Weber, under-determination

Introduction

Over the last decades, one of the most remarkable trends in empirical economics is the growing use of experimental and quasi-experimental methods — this tendency has been bibliometrically assessed by List (2009) and Hamermesh (2013). Experimental economics has evolved from a marginal position among economists to gaining acceptance into mainstream economics, therefore deserving an in-depth and lively methodological debate.[1] This evolution basically affects lab and field experiments, as 'thought experiments' are associated to theoretical rather than experimental economics (Thoma 2016; Reiss 2016), and 'natural experiments' are

Received: 20 July 2018

rare and do not share some of the main traits and challenges of most common economic experimentation (Meyer 1995; DiNardo 2008).[2] We, therefore, use the term 'economic experiments' in a restricted sense as referring to lab and field experiments in economics. Economic experiments are especially relevant to any attempt at contrasting natural science and social science experimentation, since economists often attribute to themselves a higher scientific status among social researchers (Fourcade et al. 2015).

In this article, we are going to explore this contrast from an ontology-based perspective, focusing on some ontological traits of the social domain as a continuous source of issues challenging economic experimentation.[3] Although our approach fits well with the current interest in social ontology, we intend here neither to endorse nor to discard the *ontological foundationalism* recently criticized by Sugden (2016) and Lohse (2017). Our purpose is not to sustain that 'ontological investigations play a central role for the social sciences as they are the foundation for the explanation of social phenomena, social regularities, and the effects that social phenomena have on individual behaviour', taking Lohse's characterization of ontological foundationalism. Without going that far, we only assume that the adequacy of a given social research methodology — economic experimentation, in our case— to the ontological features of the social domain is a legitimate question that is worth inquiring about. Both the social right not to be fooled by overstated self-attribution of scientificity, and the efficient steering of the collective effort in social science, require an open and critical examination of such issue. By analysing the ontological sources of methodological problems, our discussion raises some fundamental questions that need to be addressed before it is possible to develop a methodological stance on important issues like, for example, in what cases economics should refrain from emulating physics and what innovations could best handle the ontological challenges.

Although the ontological attributes highlighted here are separately and fragmentarily considered in the literature, the global picture we provide calls into question the current relationship between mainstream and experimental economics, where the latter is meant to reinforce the scientific credentials of the former. Our ontology-based approach warns against the use of economic experimentation as an excuse to attribute to mainstream economics an even more 'perfect' resemblance to physics (see Drakopoulos 2016). Conversely, it encourages the recognition of economics as a genuine social science and, consequently, also a broader and more committed methodological pluralism in the economic field.[4] Moreover, we

are aware that many of the abovementioned ontological features hold for social science experimentation more generally, yet, since mainstream economists have traditionally been more reluctant than other social scientists to acknowledge their implications, economic experiments are particularly challenged by such features.

The remaining of the article is organized as follows. First, we provide a brief and balanced account of the similarities and differences between natural and social sciences, based on Weber's conciliatory attempt to overcome the so called 'battle of the methods' (*Methodenstreit*). The purpose of the section is to frame our discussion of the differences between both fields of sciences in a more comprehensive view where the commonalities, nuances and degrees of closeness between the two fields are also acknowledged. Then we highlight some ontological features of the social domain and point out their problematic incidence on experimental economics. Specifically, our discussion revolves around human beings' representational capacities and intentionality, their cultural and conventionally mediated forms of social interaction, and the holistic openness, instability and uncertainty of the social world.[5] After that we draw attention to the strong under-determination of theory by evidence in social science, which comes hand in hand with empirical ambiguity, confirmatory biases and proclivity to pseudoscientific practices in experimental economics. Finally, we end with some concluding remarks on the methodological relevance of ontological issues.

Similarities and differences between natural and social sciences: the conciliatory Weberian view

Max Weber's abovementioned attempt to overcome the battle of the methods resulted in a conciliatory view that has become widely accepted. Indeed, an important lesson learnt from Weber is that some important characteristics for a long time considered exclusive to natural sciences are nevertheless also shared by social sciences, and conversely (Huff 1982a, pp. 82-84, 89-93). With respect to the first set of characteristics, Weber pinpoints the following aspects connecting natural and social sciences:

i) the recognizability and computability of observed regularities;
ii) the need to operate with non-observable variables or theoretical constructs that account for relations between observable variables; and
iii) the possibility of developing causal-nomological explanations.

As for those features usually ascribed to social science but also present in natural science, Weber mentions the following:
i) the uncertainty and incalculability affecting both social and natural events; [6]
ii) the radical incompleteness of all scientific knowledge, which manifests itself in the statistical or probabilistic character of all scientific laws;
iii) the under-determination of theory by observation, which is evident in the possibility of devising an infinite number of theories proving equally empirically adequate; and
iv) the interpretability of both social and natural events, all of which would be equally subject to meaning attribution.

The above lists have been considerably extended in contemporary discussions of the matter. The unavoidable resort to idealizations, the employment of simulations, the creation of new phenomena in the laboratory or the use of data models as empirical basis constitute only a few current examples of other shared features unifying social and natural sciences.[7] Interestingly, the technical notion of 'ideal type', which Weber introduced in order to characterize what he regarded as a methodological device peculiar to social science, has been increasingly vindicated as valuable also to natural science (Ramsey 1994, Godek 2016). More specifically, the deformational modelling involved in concepts like 'free market' or 'rational agent' fits into Leszek Nowak's general typology of deformational procedures as an example of quantitative deformation (positive potentialization).

However, as Weber himself noted, these common features should not mask the key underlying difference between both sciences, namely, the intentional character of social events as opposed to the non-intentional character of natural events.[8] More recently, Alex Rosenberg has also adhered to this view maintaining that, in his own words: 'it is intentionality that makes the difference between the methodologies of the social and the natural sciences' (Rosenberg 2015, p. xi). The intentional character of human actions amounts to a crucial ontological difference that in turn originates a methodological one, since only in social science will causal explanation require taking behavior out of the domain of the psychological and putting it into the domain of the culturally determined forms of responding to the world (Weber 1949, pp. 66-76; Huff 1982b, pp. 208-209). Even if scientific work on social phenomena aims at the discovery of laws, it should also include intentional notions like purposes, reasons and motives in non-nomological kinds of explanations.

Ontological features of the social domain: their problematic incidence on experimental economics

As suggested in the previous section, the very object of study in the social sciences has several distinctive features, with no analogue in the domain of physics. Our purpose here is not to systematically overview them, but only to emphasize how some of the most commonly acknowledged ones have a problematic incidence on experimental economics.

Representational capacities and intentionality

Planets, electrons, magnetic fields or wave-lengths do not have representations, neither of the surroundings, nor of themselves. Their behaviour is not determined by any kind of representation and, therefore, no recognition of experimenters studying them may affect their conduct. Conversely, scientists do not have to create hypotheses about which representations should be ascribed to them in order to explain their behaviour. The situation is the opposite in the case of human subjects, who have mental states, such as perceptions, thoughts, beliefs, or desires oriented or referring both to things in their environment and to themselves. It is important to note that intentions are dependent on both representations and values or desires. An agent's intention to do something stems from his or her representation of that something as well as from the value that the agent attaches to it. Let us recall that, according to Davidson (1963 [1989]), in order to be able to rationally interpret human action, we must be able to identify motives. These consist of reasons and intentions, the former entailing an epistemic attitude regarding representations' fitting with reality — for example, the representation of me being able to win a race, together with my desire to join the race, may give me a good reason to decide to take part in the race. Hence, motives and their constituents may, to a greater or lesser extent, arise from beliefs, that is, from interwoven dispositions to think in certain ways. The acknowledgment of these essential elements of what we usually call 'human behaviour' underlies Rosenberg's (2015, p. 35) remainder that social science aims to explain human *action* and not mere behaviour. As opposed to mere behaviours like the beating of the heart or the reflexive withdrawal from painful stimuli, actions are something that we do to our bodies, not something that merely happens to them. [9]

The awareness of experimental subjects, together with their representational capacities and intentionality are a key issue in economic experiments. As Bardsley

(2005) remarks, these experiments generally consist in manipulating information, whose effects are operative only in so far as subjects are aware of it. This contrasts not only with what happens in physical science experiments, but even with what is the case in medical trials, where causal factors operate independently of the subjects' awareness of it. In other words, Hawthorne effects and demand effects in economic experiments do not have an analogue in physical science experiments, and their parallelism with placebo effect is limited by the absence of an economic 'therapeutic effect' not mediated by subjects' awareness. [10] Going further, Bardsley argues that this leads to a break in the strict identity of laboratory and target variables, an identity underlying the demonstrative force of natural science experiments. As stated by him, social science labs cannot equalize between treatments any potential impact of extraneous factors, therefore lacking the same degree of control as the ones in natural science and constituting an 'irreducible difference' between natural and social science experiments.

Subjects' awareness and experimenters' inability to fully equalize covariates across groups are not only an issue in lab experiments, but also in field experiments. Awareness allows subjects to perceive, in experimental manipulations, uncontrolled cues that may trigger norm-driven behaviour and lead to invalid inferences, and these consequences may actually be even worse in field experiments due to the higher risk of drawing unwarranted policy conclusions (Jimenez-Buedo and Guala 2016, p. 19). The famous RAND Health Insurance Experiment, carried over from 1974 to 1981 on more than 5,800 individuals in six different US locations, can illustrate the challenge of equalizing covariates in field experiments. To assure that a set of subjects' characteristics are balanced across groups, the original RAND investigators apply a stratified random assignment to groups and investigate attrition as a potential source of biases. However, three decades later, Aron-Dine et al. (2013, p. 207) find that differential non response or attrition across groups generates unbalanced pre-randomization covariates that are potentially biasing. We cannot dismiss the possibility that these imbalances in available observable factors coexist with and result from other non-available and/or unobservable factors, linked to the subjects' representational capacities and intentionality.

More specifically, subjects' awareness of the experimental nature of the situation they confront may give rise to experimenter effects, low involvement, and different artefacts, [11] which in turn exacerbates the inscrutability of subjects' motives. For instance, in an economic field experiment on the effectiveness in raising money of voluntary contributions *vis à vis* charity lotteries, Landry et al. (2006) find an experimenter-effect due to the physical attractiveness of the female solicitor.

The risk of subjects' 'unreal' involvement, as well as their feeling of artificiality in the tasks they face, are some of the concerns emerging from Hogarth's (2005, p. 259) comparison between experimental demands regarding judgments of willingness-to pay and willingness-to-accept on trivial issues, like small gambles, and those on consequential issues, like compensation awards in civil trials. [12] Experimenters in economics have to deal, not only with the holistic features of human decision making, but also with the inter-individual divergence in the behavioural manifestation of shared motives, and, conversely, with the inter-individual convergence in the behavioural manifestation of different motives — see, for instance, Farina et al.'s (2009) attempts to disentangle strategic from other-regarding motivations for trust and reciprocity. Therefore, experimenters have to constantly reevaluate everything, including themselves, their data interpretations, and the way they communicate and solicit word descriptions within experiments (Smith 2010, p. 4). Similar difficulties have been stressed by Jimenez-Buedo and Guala (2016), who discuss how economic experimenters struggle to understand a threefold process involving the way subjects construe the experimental task, react to the expectations of the rest of participants, and rely on cues triggered by the experimental manipulation.

It is needless to say that language plays an essential role both in the development of objective, precise and abstract representations, and in the communication of human thoughts, beliefs, desires, intentions, etc. But linguistic access to subjects by social researchers provides only indirect, partial, and very often ambiguous access to their representations and intentional states. As an example of how experimental economics faces this kind of difficulties, Smith (2010, p. 4) discusses the case where subjects involved in an ultimatum or dictator game describe their experiences in terms of 'unfairness', posing the question as to whether they mean something about the rules, the outcomes, the outcomes under the rules, or other circumstances surrounding the experiment. While the experimenters use 'fair' in a specific outcome sense entrenched in the utilitarian tradition, the subjects might use it either in a procedural sense or in a colloquial sense, having in mind justice. Since experimenters' direct access to what the subjects mean is unfeasible, they must carefully consider how subjects see the circumstances surrounding decision-making in order to interpret their speech.

A phenomenon closely connected to all above features is what we could call 'biographical holism', which makes any human action dependent on the entire past of the person at hand. Indeed, the cognitive as well as the intentional development of a human being remains open to be influenced, not only by experiences and

information gathered through the interaction with the surrounding, but also by the inner recollection and treatment of those experiences and information. In contrast to social beings, physical entities usually exhibit a greater stability, as, in normal circumstances, only a limited range of phenomena would prompt a structural change in them — and, as long as they keep their structural properties, no knowledge of their past has predictive relevance. To put it shortly, human individuals are more historically-sensitive than physical entities. As a consequence, knowledge requirements for accurately predicting human behaviour seem particularly harder to be met, while, on the other hand, knowledge requirements for predicting the behaviour of a physical entity can often be accurately met once its internal structure has been determined. In this vein and regarding experimental economics, Herbert Gintis (2000, p. 319) highlights that biographical holism renders it impossible to eliminate all uncontrolled influences:

> In physics and chemistry (...) this goal can be achieved because elementary particles, and even chemical compounds, are completely interchangeable, given a few easily measurable characteristics (atomic number, energy, spin, chemical composition, and the like). Experiments in human social interaction, however, *cannot* achieve this goal, even in principle, because experimental subjects bring their personal history with them into the laboratory. Their behavior is therefore *ineluctably* an interaction between the subject's personal history and the experimenter's controlled laboratory conditions.

Individuals do not only passively react to whatever stimula they receive; rather, they are cognitively and intentionally active in representing or interpreting social situations they find themselves in. On the other hand, not all individuals see things the same way when exposed to the same circumstances or the same options. Agents' representations thus constitute an essential variable whose specific characteristics, in each experiment, cannot be taken for granted, otherwise it becomes a serious source of confounds. Rightly identifying representation, however, is not an easy task since, in building representations as well as related intentions, a whole array of complex and holistic factors come into play, varying depending on each subject's biography, cultural and social background, etc.

Cultural and conventionally mediated forms of social interaction

Conventions ruling social behaviour are established on the basis of shared representations and beliefs of higher orders (i.e., beliefs about others' beliefs, and about others' beliefs about others' beliefs...). Now, social conventions are embedded

in a culture or, to put it differently, they are socially constructed rather than given by nature. They result from decisions concerning how to coordinate actions given certain interests. Hume's classical account of conventions, further developed by D. Lewis in contemporary philosophy, emphasizes the role played by high order beliefs. Following a convention requires sharing beliefs about other people's beliefs concerning other people's beliefs that all the others follow the same convention for certain reasons (i.e. the 'I know that you know that I know' kind of scenario). A convention, however, need not be established as an explicit agreement, raising the problem of how to identify conventions in those cases where no explicit formulation is available. Property, for instance, is presented by Hume as a paradigmatic example of a social convention, in this case one governing the enjoyment of goods. Cultural elements, like language, politics, art, economy can all be regarded as (implying) social systems based on conventions.

Within the philosophy of economics, the conventional and cultural features intrinsic to the social domain have been recently stressed by Uskali Mäki, who accordingly introduces a new technical term to refer to the kinds of entities populating the social domain, as opposed to the kinds of entities inhabiting the natural domain. According to him, the economic world is furnished with *commonsensibles*, i.e. things like money, prices, wages, preferences, expectations, beliefs, contracts, exports, and so on (Mäki 2005, pp. 243-245; 2009, pp. 87-88). Mäki notes that, unlike physical entities, many of the above things are not regarded as having a mind-independent existence but just an inquiry-independent one. Commonsensibles moreover, although including many unobservables like beliefs and preferences, belong to the world of ordinary, common sense experience, rather than to that of counter-intuitive theoretical entities as the ones postulated in physics. Generalizing Mäki's point, we can conclude that, with respect to ontological questions, the observable/unobservable divide plays a minor role in social science when compared to its role in natural science. Commonsense unobservables are a major part of the ontological basis of social phenomena and, consequently, a wide range of folk psychology presuppositions are involved in understanding social science's domain of inquiry.

A crucial challenge here for experimental economics is that of implementing in the lab, without deception [13], those mediating conventions shaping social interaction. In this respect, Bardsley (2005) brings up, among other examples, Ball et al.'s (2001) attempt to explore the effects of status in markets. These experimenters define status as 'a ranking in a hierarchy that is socially recognized'. However, despite their

claim that social status was implemented, Bardsley contends that nothing in the experiment is recognized as a ranking in a hierarchy outside the experiment. Status could hardly be imparted in the experiment at all, so the pseudo-status created in the experiment would have very few conceivable real-market consequences. Access to subjects' interpretations is required in order to improve the experimenters' interpretation of the experiment.

As a proponent of the Cambridge social ontology, Siakantaris (2000) notes that distortions originated in the laboratory create 'an institution implying its own rules', a special kind of society, and sees the 'internal-relationality' of socio-economic mechanisms as an obstacle to their lab isolation. Similarly, Tony Lawson, the leading figure from the same school of thought, emphasizes that statuses or social positions and social forms in general are mutually constitutive and, therefore, impossible to isolate through controlled experiments (Lawson 2009b, p. 765; 2015a, p. 44). The same way that employers, for instance, do not exist unless in relation to employees, managers do not exist unless in relation to subordinates. Analogously, markets cannot exist insulated from firms, property rights or monetary systems. Lawson (2015b, p. 318) concludes that, due to the interconnectedness and mutual constitution of many different features of social reality, it is impossible to experimentally isolate its individual components.

To further illustrate the drawbacks and difficulties in implementing conventionally mediated social entities and commonsensibles in economic labs, let us consider Smith's (1962) famous experiments on double oral auctions. After claiming that these experiments are 'designed to study some of the hypotheses of neoclassical competitive market theory,' he acknowledges, however, 'that they are intended as simulations of certain key features of the organized markets and of competitive markets generally, rather than as direct, exhaustive simulations of any particular organized exchange.' (Smith 1962, p. 111) Indeed, neither money nor commodities supposedly exchanged in his experiments are real, and supposed preferences are not subjects' genuine, homegrown ones. Each experiment encompasses a sequence of 'trading periods' of barely a few minutes, and in most of them each 'buyer' and 'seller' is allowed to effectively exchange only a single unit of the fictitious commodity. Both buyers and sellers are instructed to think of themselves as making a pure profit, but no monetary payoff is gotten by them. Quantity-adjusting decisions by the marketers and speculative purchases are not accommodated. We are fully aware that all of these design choices are deliberate, and quite a few of them are relaxed or modified by Smith himself in later experiments. Nevertheless, the point here is

what epistemic status can be attributed to such all too artificial experiments, and how they can be interpreted in terms of their relevance to real-world markets. This would appear as highly debatable if it was not for new series of experiments that provide evidence of robustness. Moreover, not all of Smith's (1962) choices have easily available alternatives, as implementing certain commonsensibles is not a simple task. Inducing preferences or valuations, for instance, is an expeditious way to dodge the very difficult issue of uncovering them from the traders, given their strong incentives to misrepresent their preferences and the serious weaknesses of contingent valuation surveys used to this effect (Al-Ubaydli and List 2017).

The diverse, contingent and evolving nature of the regularities stemming from conventions points directly to the high localism characteristic of social regularities, as opposed to natural regularities. Different cultures and even different groups within a culture may create different conventions, which would also be subject to historical change. Natural events are not ruled by conventions and, therefore, are not affected by this sort of localism. The nomological scope and stability exhibited by natural regularities is not comparable to that so far exhibited by social regularities. Yet, in so far as regularities, however limited, are identifiable, predictive explanations are still possible and ideographic explanations by retroduction are not the only option (Mitchell 2009; Byrne and Uprichard 2012). To this regard, Jimenez-Buedo and Guala (2016, p. 19) have pointed out that the validity of inferences drawn from economic lab experiments on social norms are in most cases limited to a particular culture. In economic field experiments as emblematic as the ones developed by Duflo and Banerjee at the Abdul Latif Jameel Poverty Action Lab, the results are considered, too, as highly dependent on context (Favereau 2016, pp. 216, 219).

Holistic openness, instability and uncertainty

The biographical and social holisms tackled in the last two subsections give rise to an openness, instability and uncertainty in social phenomena more pervasive than those found in physical phenomena. While, outside labs, the physical world consists of open systems, these systems' components or causal factors tend to be essentially atomistic, intrinsically closed and stable, isolatable and susceptible to be triggered under experimentally controlled conditions (Pratten 2015, pp. 74-75). Human beings and socio-economic entities, by contrast, tend to be intrinsically open, intertwined and evolving. Now, holism comes in different degrees and modalities. The scope of holism differs drastically depending on the degree of inter-dependency between the

parts or aspects conforming certain unitary entity or phenomenon. Also, holism may be stable or dynamic depending on whether the holistic system is closed (like a syntactic system) or open (like an economic system). In contrast to syntactic systems, where we know both the elements involved and their contribution in relation to the rest, social systems like real markets are highly complex and dynamic, where both the elements involved, and their contributions are uncertain.

Drawing on this holistic reality, Leonardo Ivarola (2017) argues that, since people's actions are the necessary bridge in any causal relationship between economic variables, each of the latter can make different causal contributions depending on how factors, both endogenous and exogenous to the subjects, influence their actions once the variable has been activated. To use his example, an increase in the quantity of money (causal variable) may either increase the national income of a country, or lead to a rise in the general level of prices, or even have no impact in the observed macroeconomic variables (possible effects), depending on the people's actions connecting cause and effect. These actions, on the other hand, may be unstable because of both the interpretations that agents make of signals of the world and contextual conditions.

Furthermore, Ivarola (2017, pp. 218, 223) notices that the unstable complexity of social phenomena may mark a limit to the method of isolation, a method that presupposes a singular connection between each causal variable and a certain kind of effect. Isolating a variable from other variables makes sense when we assume that the latter are disturbing factors altering the causal power of the former. If, rather than as disturbing factors, those variables are acknowledged as originating different paths in turn associated with multiple causal contributions of the same variable, then, the set of variables originating different paths should be explored. To be clear, the goal would no longer be isolating a singular contribution between cause and effect, but the multiple contributions of the same variable — he appeals to the notion of 'possibility-tree' or 'open ended result' as that suitable to capture the basic structure of socioeconomic processes. This requires an extensive explorative work leading to robust *non-Galilean idealizations* or *structural assumptions* (Cartwright 2009) [14], i.e., assumptions whose fulfilment is necessary for a theorized causal relationship to hold. Although some sort of isolation and causal stability is still needed if we are seeking explanations with some degree of generality, it should be compatible with the fact that social processes occur in open systems where variables have plural rather than single causal contributions.

Regarding economic experimentation, Guala (2012, p. 610) calls attention to the fact that experiments have manipulative, representative and isolative analogies with models and notes that, at least at the beginning of a research program, relatively simple experiments that instantiate the isolative assumptions of their simple modelling counterparts tend to be used. According to him, however, there remains at least a fundamental difference between them: '*models simulate whereas experimental systems do not*', the former 'are conceptual entities, whereas experiments are made of the same "stuff" as the target entity they are exploring and aiming at understanding.' (Guala 2012, p. 611) But to what extent can we assume that economic experiments and their target entities are made of the same stuff if experiments are designed to close and isolate what is constitutively open and holistic? As Siakantaris points out (2000, p. 270), the conditions of dominance and privacy applied in experimental economics are intended to ensure that microeconomic phenomena occurring at the lab are isolated from all 'extraneous' factors. Dominance implies that monetary incentives designed by the experimenters make any subjective cost or benefits associated with participating in the experiment comparatively negligible; privacy, on the other hand, requires that each experimental subject be informed only about his/her reward schedule, not about the rewards of any other participants. The condition of privacy would help to remove possible influences between the subjects in the laboratory. Subjects in an experimental situation should act as atomistically as possible. The influence of all kinds of factors internal to each of the subjects, like mood, age, gender, occupation, cultural aspects..., must be minimized as they may interfere with the experimental result. Adopting Santos' (2007) taxonomy, it could be said that all these issues may undermine the crucial 'material' of behavioural experiments (the human participants), while the material of technological experiments (the market institution) is also under dispute [15]. Therefore, the epistemic value of economic experiments relative to economic models might, to some extent, be put into question. The tighter the control is, the lesser the openness, instability and uncertainty that experimental results may have, but also, as Santos (2009) shows, the more these results may be the outcome of economists' actions. Hence, the epistemic value of experiments is partly a function of the degree to which the experimental control softens or over-constraints human agency.

Compared to Smith's double auction experiments discussed in the previous section, the famous ultimatum bargaining experiment conducted by Güth et al. (1982) leaves much more room for human agency and motivational variety. Nevertheless, subjects are isolated from any history of each other's interactions, since bargaining

pairs are determined stochastically and, when the experiment is repeated a week later, each subject usually has to expect a different opponent. To some extent, they are also isolated from mutual interaction, as simultaneous moves are not accommodated, and subjects are seated far enough from each other so as to exclude verbal communication. In a sense, every player finds herself in a 1-person game, because the only thing she has to do is to make a choice that is good for herself — if performing as a 'proposer', she only needs to anticipate the future decision of her 'responder'. Although Güth et al. claim that their experiments are useful to analyse certain aspects of bargaining behaviour, they deliberately avoid the discussion about 'whether ultimatum bargaining games can adequately represent real bargaining situations' (Güth et al. 1982, p. 368). As in Smith's (1962) experiments, the epistemic status of this first ultimatum game experiment and its relevance to real-world bargaining behaviour would, therefore, appear as highly disputable, if it were not for the fact that new series of experiments provided evidence of robustness.

Severe under-determination of theory by evidence: ambiguity, confirmatory bias and pseudoscientific practices

In a previous section, we mentioned the under-determination of theory by observation among the features mentioned by Weber as usually ascribed to social science but also present in natural science. As Psillos (2005) explains, evidence may underdetermine theory in the sense that the former cannot prove the truth of the latter or render it probable, with the epistemic implication that 'belief in theory is never warranted by the evidence.' This sort of under-determination is an issue even in natural science, as confirmation holism raises some serious difficulties in determining whether disconfirming evidence refutes a theoretical postulate, some auxiliary assumption or the description of initial conditions. Yet, there are cases where contrary evidence turns out discriminating enough, that is, cases in which some recalcitrant, salient disconfirming evidence points to a flaw in the theoretical postulate, as no empirically sound variation of the other elements could make it admissible to keep the theoretical postulate. However, the ontological peculiarities of the social domain — already examined in the previous sections — make the under-determination more severe in social science than in natural science, making it more difficult to decide when a theoretical assumption, rather than some auxiliary assumption or the description of the initial conditions, should be discarded given some disconfirming evidence.

A wide variety of authors have acknowledged the above problem. According to Vernon L. Smith, the under-determination or Duhem-Quine problem is at the heart of Hertwig and Ortmann's (2001) criticism of experimental economics. Some Smithian examples of *ex post facto* reinterpretations in game experiments (Table 1) bring to mind some of the abovementioned peculiarities of the social domain. From a Bayesian framework, Søberg's (2002) working paper concludes that experimental reproductions do not solve the problem of non-falsifiability per se, and that the Duhem-Quine challenge is not overcome by assigning different subjective degrees of beliefs to various conjunctions of main and auxiliary hypotheses. However, Søberg's (2005) final version, which has been published as an article in the *Journal of Economic Methodology*, ends on a more optimistic note and concludes that the Bayesian strategy does provide a rationale for the recalcitrant observed regularities. Francesco Guala (2005) agrees that the Duhem-Quine problem is only insurmountable for an ultra-deductivist approach to hypothesis testing. In any case, the importance of the ontological peculiarities of the social domain emerges again in connection with the under-determination problem. Establishing auxiliary hypotheses about a stable domain, where only a limited number of variables are relevant, is very different from formulating those hypotheses for a highly unstable, open domain that most often entails complex and unstable background conditions for phenomena occurring within it. To put it bluntly, auxiliary hypotheses in social science can be as difficult to establish and as conjectural as theoretical hypothesis.

Table 1 A few examples of *ex post facto* reinterpretations when tests of game theory fail

Treatment protocol and/or context
- Face-to face bargaining
- Anonymous dictator, ultimatum, and other extensive form games
- Same; any experiment
- Same, but use advanced graduate students
- Any experiment
Theory rescuing interpretation
- Social context loses control over preferences
- Reparameterize with other-regarding utility to fit the data
- Unsophisticated undergraduates
- All subjects need experience: repeat with new strangers after each play
- Payoffs too low

Source: Smith (2001, p. 428).

The severe under-determination of economic theory by the evidence gathered from experimental economics and econometric research goes hand in hand with a persistent ambiguity in empirical results. This ambiguity is in turn related to the lack of correlation between patterns of behaviour and psychological properties, something aggravated by the multiplicity of psychological variables (both representational and intentional) oriented in turn towards a multiplicity of dynamic cultural and social variables. If the strict identity of laboratory and target variables is under question, biographical and social forms of holism acknowledged, and structural assumptions distrusted, experimental outcomes will be dramatically open to interpretation. Whether observations from economic experiments agree with the predicted actions or not, 'the inference of truth from observation is inherently ambiguous, and [experimenters] face the daunting challenge of using [their] experimental skills and imagination to reduce this ambiguity.' (Smith 2010, p. 3) Experimenters' inferences from their subjects' lab behaviour may fail due to the inter-individual divergence in the behavioural manifestation of shared motives, and, conversely, to the inter-individual convergence in the behavioural manifestation of different motives. To put it differently, both the correlation and the lack of correlation between individuals' manifestation of motives may be spurious, since there is guarantee neither that people sharing the same motives will behave the same way, nor that people having different motives will behave differently. In limiting cases, the concern may arise that the experimental basis for a hypothesis is no less hypothetical or better confirmed than the hypothesis for which that basis should provide support. Contrary to measurement instruments employed in natural science (scales, thermometers, interferometer, oscilloscope), no material but mental intervention is attempted through economic experiments and other social research instruments whose causal mechanism remains obscure. Not surprisingly, in his analysis of hypothesis validity, Wiggins (1968, p. 423) recalled Cicourel's claim that sociological variables are so obscure that it is not possible to specify how they can be manipulated in an experiment, rising the objection that 'it is possible to examine social behaviour experimentally only by using the principles of social behaviour.' In the social domain, a complex and unstable surrounding affects complex, unstable subjects in complex, unstable ways that include a continuous and peculiar feedback between both surroundings and subjects, i.e., a feedback where the causal power of (self-)representational capacities gives rise to loops and a variety of reactions to each other's representations. Unstable domains in natural science, like the one studied in meteorology or epidemiology, do not present loops — that is, phenomena do not become what scientists think about them by virtue of their very thinking about

them— , and there are a limited well defined range of possible variables operating in ways that are also limited and often well-known.

All of the above affect not only lab experiments but also field experiments, despite the latter being often viewed as ruled by relatively theory-free research designs establishing what works and what does not work. The ontology-related ambiguity of empirical results impinges upon the external validity of economic field experiments, albeit those researchers engaged in large-scale Randomized Controlled Trials (RCTs) are prone to think that their large datasets and careful experimental design ensure validity and scientificity. However, as Hausman notes, things are not that simple, for 'without knowledge of the mechanisms [underlying an estimated impact], it is all too easy for an intervention that works splendidly at a specific time and place to fail abysmally when tried elsewhere' (Hausman 2018). This issue is analysed in detail by Ruzzene (2015) in relation to those cases where RCTs were employed for purposes of evaluating development policies and improving policy-making in developing countries. She emphasizes that a given policy or program outcome might depend on the specific features of the experimental setting and the particular implementation of the program, as well as on an array of context-specific conditions. Furthermore, she criticizes three available responses to the problem of external validity in field experiments: replication, cross-country regressions and structural modelling of the causal structure of interest. Replication may indeed provide misleading evidence on external validity because both a failed result and a successful one may have more to do with (often unknown) conditioning factors than with the efficacy of the intervention under the circumstances selected in the primary experiment. This resembles the above idea that both the correlation and the lack of correlation between events may be spurious, since neither the former necessarily means convergence in causal mechanisms, nor the latter necessarily suggests divergence.

The severe under-determination and high ambiguity of results within economic experiments appear to leave more room for confirmatory biases and pseudoscientific practices than is left in physical-science experimentation. Fifty years ago, Wiggins (1968, pp. 399-400) already warned against the risk that an experimenter's expectation of subjects' behaviour biases results towards confirming such expectation. According to him, this could occur in the two following ways: (1) the behaviour of the experimenter causes on the subjects a behaviour consistent with the former's expectation or hypothesis; (2) the experimenter's hypothesis causes her/him to perceive the behaviour of her/his subjects as consistent with

it. Regarding experimental economics, it has even been argued that some of its prescriptive methodologies, like script enactment, repeated measures, performance-based payments, and absence of deception, 'force participants to conform to a normative expectation that they must behave rationally and in accordance with their self-interest' (Van Vugt 2001, pp. 429-430). Certainly, this charge could now be partially dropped as, in practice, experimental subjects' have shown a wide range of 'anomalies' or departures from such *homo œconomicus* model. [16] However, economic experimenters' tendency to assume that a testing procedure is 'right' or 'wrong' depending on whether or not it leads to confirmatory predictions has also been authoritatively criticized (Smith 2010, p. 6). Confirmatory bias is also objectionable on the grounds that, in social science, there is a much higher risk of modifying the subject under study in a way that fits the researcher's expectations.

Despite all the above, we do not intend to globally qualify experimental economics as pseudoscientific because, first, there is no unanimous clear-cut demarcation criterion for science, and, second, not-yet-scientific knowledge should be distinguished from pseudoscientific, merely apparent knowledge, where there is a deceptive or distorted element involved. [17] Throughout the last decades, the general tendency in philosophy has been to combine the understanding of 'science' as a family resemblance concept (in the Wittgensteinian sense) with the recognition of multi-criterial indicators revealing particular pseudoscientific issues. As with any other vague concept, the lack of a clear-cut definition is compatible with finding criteria to identify clear cases falling out of the scope of the concept. Mario Bunge, Philip Kitcher, Paul Thagard, Gerhard Vollmer, Martin Mahner and Sven Ove Hansson are some of the authors who have been working in this direction. In his 1996 paper, Hansson mentions the unwillingness to test and the systematic disregard of refuting information as some characteristic features of pseudoscientific practices that, when occurring in scientific disciplines, are usually masked under some rhetorical or technical moves. The cluster approach to demarcation vindicated by these authors is the one gaining wider agreement today. In this view, the scientific status of a research depends on a variety of heterogeneous factors ranging from the independent testability of auxiliary hypothesis to the search of both confirmation and disconfirmation (Mahner 2013).

In experimental economics, we find a problem that is closely connected to the ones pointed out by Hansson, namely, the dogmatic endorsement of false theoretical postulates on human behavior in the face of contrary evidence. Instead of beginning by developing a realistic theory of such behavior based on experimental or

observational evidence, the dominant approach is to retain the traditional *homo œconomicus* assumption as the starting point, and then to modify the analysis by incorporating observed deviations. Going along this line, Joffe (2017, p. 12) contends that axioms do not represent an actually occurring process in terms of causal mechanisms. Furthermore, Starmer (2005) notes that, while economic experiments have prompted the emergence and current proliferation of non-expected utility models, an implicit pre-commitment to normatively appealing principles is still kept. This pre-commitment would be manifest in the preservation of normatively attractive principles such and monotonicity and transitivity (in spite of contrary evidence) in the proposed revisions of the expected utility theory. Furthermore, there has been a dramatic lack of attention to those experimentally observed deviations that seem to escape from any rational vindication. In sum, the neglect of empirically established features of the social domain leads to the preservation of false theoretical assumptions despite recurrent disconfirming evidence.

There are at least three ontological factors ultimately involved in the kind of social scientists' pseudoscientific practice consisting in the dogmatic acceptance of false theoretical postulates, without providing, for such postulates, any form of concretization *à la* Nowak — i.e., any modification of postulates corresponding to the specification of certain empirical parameters (Borbone & Brzechczyn 2016, pp. 3-4). First, there is the fact that rational patterns of behaviour (or inferences) constitute a main target in social inquiry. Hence any attempt at explaining social phenomena requires the identification of rational patterns according to some normative framework, as opposed to what happens in natural science, where no such norm-relative identification is needed, but, rather, salient empirical regularities and basic causal relationships suffice as the target for some theorizing. Second, given that the highly ambiguous and variable nature of social phenomena and its incidence in the under-determination of theory by observation is unparalleled in natural science, pseudoscientific assumptions may be more easily kept in social science, even if disconfirming evidence clearly suggests their falsity. Third, the dynamical and holistic features of the cultural and conventional side of the social domain result in highly complex and intertwined forms of social interaction. The scientific reliance on idealizing assumptions in order to theoretically decompose complex interactions and isolate causal variables could ultimately lead social scientists to keep plainly false assumptions. They could argue that the extreme complexity of social phenomena calls for extreme idealizing assumptions in the form of clearly false assumptions that, notwithstanding, would successfully fulfil certain strategic epistemic roles. Yet, once that the intended epistemic roles of

idealizing assumptions have been carefully specified, it becomes obvious that certain false assumptions are kept even though they do not contribute to the isolation of any real causal variable.

Concluding remarks

Experimental economics appears nowadays as a fashionable field, being one of the most prominent expressions of the current empirical turn in economics. It comfortably fits with the traditional commitment to naturalistic research methodology, endorsed, at least rhetorically, by mainstream economists. Experimental economics goes hand in hand with the greater attention paid to cognitive social psychology and the bases of human behaviour, an attention which has tended to make economics 'recognizably more like the other social sciences it once feigned to distain' (Rosenberg 2015, p. x). These are seemingly good news to those who, as in our case, advocate a methodological pluralism in the discipline, a pluralism much more far-reaching than current predominant math-modelling and econometrics — although formulating a pluralistic methodology for empirical economics goes well beyond the purposes of this article, we suggest the use of historical and comparative methods, case studies and survey research, in addition to the most common econometric analyses and to nowadays expanding economic experimentation.

Economic experiments have usefully contributed to undermine economists' recalcitrant adherence to that *a priori* and narrow assumption known as *homo œconomicus*. However, this healthy effect does not guarantee a high experimental validity of such experiments, but rather an a more fully recognition of the value of experimental economics by mainstream economists. This recognition paves the ground for major contributions from experimental economics, but its potential to revamp economic science is limited by the ontological peculiarities of its object of study, and it may fizzle out because of the experimenters' kowtow to false theoretical postulates and its subjugation to econometricism (Morgan and Patomäki 2017, p. 1410). If something emerges from our discussion is that *social ontology matters*, which implies that economic experimentation leaves more room for interpretation and have less demonstrative force than physical-science experimentation. Logically enough, this should moderate the 'scientificity' claims made on the basis of economists in white coats and urges them to develop an appropriate toolbox to deal with the array of validity challenges arisen from ontological traits extraneous

to the physical world. Paying attention to these traits is crucial for both the critical assessment of scientific practices and the development of methodological innovations sensitive to the ontological peculiarities.

Endnotes

[1] See, for instance, the monographs by Guala (2005) and Santos (2010) or the more recent handbook edited by Fréchette and Schotter (2015), apart from innumerable articles in a wide variety of academic journals.

[2] As El Skaf and Imbert (2013) note, computer simulations and thought experiments are usually seen as having different roles and epistemologies than those of experiments in their proper sense. Regarding Mäki's characterization of models as experiments, see infra, last section.

[3] We use 'ontological' as referring to those properties' characteristic of the way in which certain kind of entities and relations exit.

[4] According to Lawson (2009a), each of the so called 'new research programs' in contemporary economics comes to constitute — at least in most of the cases— a mere 'heterodoxy within the mainstream' or a 'mathematical economics heterodoxy'. They do not satisfy the need of 'a more genuinely pluralistic economics, one that does not support an unreasoned insistence on mathematics only and everywhere.' See also Dow (2004) for a view focused on the ontological foundations for methodological pluralism.

[5] Attributing a holistic nature to something means that we do not think of it as constituted by the addition of single, independent parts, but by the mutual relation between them. Organisms such as the human body form holistic entities, while collections and aggregates such as apples in a basket or sugar cubes are not holistic, just consisting in an addition of single, independent elements.

[6] This connection between physical and mental or social phenomena has been emphasized by Patrick Suppes (1985, 2008).

[7] Some influential authors that have discussed the above topics are, respectively, Ramsey (1994) and Godfrey-Smith (2009), Hartmann (1996) and Morgan and Morrison (1999), Hacking (1983, ch. 13), and Suppes (1962).

[8] Such intentionality has been understood within the philosophical tradition, at least since Brentano, as the oriented dimension of the human mind. Such dimension would make it possible for the mind to develop mental representations (like impressions, ideas, memories, etc.) standing (most often) for other things (perceptual properties, kinds of entities, events, etc.). In order to avoid ambiguity, we will use the term 'intentionality' only in its ordinary sense, as referring to intentions to attain certain goals, and not in the philosophical one just mentioned, which can be equated to representational capacities. The reason for this is that the ordinary sense of 'intentionality' seems the prevalent one in social science.

[9] In what follows, when we use the term behaviour regarding social science's domain, it will embrace only the kind of behaviour that consist of human actions.

[10] For a systematic comparison between clinical trials and economic field experiments, see Favereau (2016). She draws attention to the fact that the latter are not framed in any ex-ante knowledge like that resulting from the preclinical phase (studies on animals), phase I (previous tests on healthy individuals), and phase II (optimal dose) of a therapeutic trial. Because of this gap that economic field experiments may fail to provide univocal explanations and clear recommendations.

[11] According to Hilpinen (2011), the expression 'artefact' in this context is 'used to refer to experimental results which are not manifestations of the natural phenomena under investigation, but are due to the particular experimental arrangement, and hence indirectly to human agency.'

[12] A well-known example of artefact in experimental economics is the one resulting from the dictator game, recently analysed by Jimenez-Buedo (2015).

[13] The prohibition of deception is a main rule governing experimental economics, one adopted in order to avoid, in present and future experiments, the distortions induced by the subjects' distrust. However, this rule is not without cost, as it allows neither to avoid experimental subjects' strategic responses to the true purpose of the experiment, nor to produce interesting situations hardly obtainable in an honest way, being mainly these two methodological motivations what justifies the use of deception by psychological experimenters (Hertwig and Ortmann 2001, p. 397).

[14] Galilean idealizations are those introduced only to neutralize the impact of perturbing factors and to solve tractability problems.

[15] See, for instance, Hogarth (2005, pp. 258-259), where he questions the match on all relevant dimensions between market experiments and the characteristics of 'real-world' markets.

[16] This model of human behaviour has been the focus of one of the main objections against the traditional, neo-classical paradigm in economics — a paradigm that embeds such model as an *a priori* and narrow assumption. Under the *homo œconomicus* approach, the goals of economic agents tend to be taken as given and the strategy to pursue those goals is supposed to be based on the principle of subjectively expected utility maximization. Outside experimental economics, Herbert Simon partially challenged such approach by introducing the empirically supported principle of the bounded rationality of economic agents, who would be restricted by cognitive limitations of different sorts. However, experimental economists have paved the way to question the mainstream approach more deeply by adding new empirically supported theses about goals to the ones that Simon had established about cognition. For instance, Reinhard Selten (2003) showed the important role that some previously unrecognized goals play in the economic behavior, being his experiments on solidarity especially revealing to this respect. Moreover, intransitivities in choice behavior and preference reversal represent two of the experimentally tested phenomena disconfirming the classical principle of utility maximization (González 2003, p. 76).

[17] The very notion of pseudoscience has raised a long philosophical controversy in so far as it was for some decades associated with strict demarcation criteria for science. Both the confirmationist criterion advocated by logical positivists and the falsifiability one proposed by Karl R. Popper found a strong opposition from the historicist approach developed by Thomas S. Kuhn, Paul K. Feyerabend, Irme Lakatos and Larry Laudan, among others. Most philosophers of science from the sixties to nowadays admit that there is no clear-cut demarcation criterion for science, since non-scientific theories can be highly confirmed, and no theory is completely free from anomalies. Moreover, given that confirmation is a matter of degree, any demarcation criterion based on this feature would be rendered vague.

Acknowledgements

We are thankful to Daniel Kuchler for helpful comments on an earlier version of this paper. This research was financially supported by the research projects 'Models and Theories in Physical, Biological, and Social Sciences' (PICT-2014-1741, ANPCyT, Argentina), 'Political and Economic Consequences of the Descentralization' (CSO2013-47023-C2-2-R, Spanish Ministry of Economy and Competitiveness).

References

Al-Ubaydli, Omar and J.A. List (2017), 'Field experiments in markets', in Abhijit Vinayak Banerjee and Esther Duflo (eds.), *Handbook of Field Experiments*, Vol. 1, Amsterdam: North-Holland, pp. 271-307.

Aron-Dine, Aviva, L. Einav, and A. Finkelstein (2013), 'The RAND health insurance experiment, three decades later', *Journal of Economic Perspectives*, 27(1), 197-222.

Ball, Sheryl, C. Eckel, P. J. Grossman, and W. Zamem (2001), 'Status in markets', *Quarterly Journal of Economics*, 116, 161-88.

Bardsley, Nicholas (2005), 'Experimental economics and the artificiality of alteration', *Journal of Economic Methodology*, 12(2), 239-251.

Borbone, Giacomo and Krzysztof Brzechczyn (2016), 'The role of models in science: An introduction', in Giacomo Borbone and Krzysztof Brzechczyn (eds.), *Idealization XIV: Models in Science*, Poznań Studies in the Philosophy of the Sciences and the Humanities, Volume 108, Boston: Brill/Rodopi, pp. 1-10.

Byrne, David and Emma Uprichard (2012), 'Useful complex causality', in Harold Kincaid (ed.) *The Oxford Handbook of Philosophy of Social Science*, Oxford: Oxford University Press, pp. 109-129.

Cartwright, Nancy (2009), 'If no capacities then no credible worlds. But can models reveal capacities?', *Erkenntnis*, 70, 45-58.

Davidson, Donald ([1963] 1989), 'Actions, reasons, and causes', in: *Essays on Actions and Events*, Oxford: Clarendon Press, pp. 3-19.

DiNardo, John (2008), 'Natural experiments and quasi-natural experiments', in Steven N. Durlauf and Lawrence E. Blume (eds.), *The New Palgrave Dictionary of Economics*, 2nd ed., London: Palgrave Macmillan.

Dow, Sheila C. (2004), 'Structured pluralism', *Journal of Economic Methodology*, 11(3), 275-290.

Drakopoulos, Stavros (2016), 'Economic crisis, economic methodology and the scientific ideal of physics', *The Journal of Philosophical Economics: Reflections on Economic and Social Issues*, X(1), 28-57.

El Skaf, Rawad and Cyrille Imbert (2013), 'Unfolding in the empirical sciences: experiments, thought experiments and computer simulations', *Synthese*, 190, 3451-3474.

Farina, Francesco, N. O'Higgins, and P. Sbriglia (2009), '*Suit the action to the word, the word to the action*: Eliciting motives for trust and reciprocity by attitudinal and behavioural measures', *Research in Economics*, 63, 253-265.

Favereau, Judith (2016), 'On the analogy between field experiments in economics and clinical trials in medicine', *Journal of Economic Methodology*, 23(2), 203-222.

Fourcade, Marion, E. Ollion, and Y. Algan (2015), 'The superiority of economists', *Journal of Economic Perspectives*, 29(1), 89-114.

Fréchette, Guillaume R. and Andy Schotter (eds.) (2015), *Handbook of Experimental Economic Methodology*, Oxford & New York: Oxford University Press.

Gintis, Herbert (2000), 'Beyond *Homo economicus*: evidence from experimental economics', *Ecological Economics*, 35(3), 311-322.

Godek, Lidia (2016), 'On deformational modeling. Max Weber's concept of idealization', in Giacomo Borbone and Krzysztof Brzechczyn (eds.), *Idealization XIV: Models in Science*, Poznań Studies in the Philosophy of the Sciences and the Humanities, Volume 108, Boston: Brill/Rodopi, pp. 64-80.

Godfrey-Smith, Peter (2009), 'Models and fictions in science', *Philosophical Studies*, 143, 101-116.

González, Wenceslao J. (2003), 'Rationality in experimental economics: An analysis of Reinhard Selten's approach', in Maria C. Galavotti (ed.), *Observation and Experiment in the Natural and the Social Sciences*, Dordrecht: Kluwer, pp. 71-83.

Guala, Francesco (2005), *The Methodology of Experimental Economics*, New York: Cambridge University Press.

Guala, Francesco (2012), 'Experimentation in economics', in Uskali Mäki (ed.), *Handbook of the Philosophy of Science, Volume 13: Philosophy of Economics*, Amsterdam: Elsevier, pp. 597-640.

Güth, W., R. Schmittberger, and B. Schwarze (1982), 'An experimental analysis of ultimatum bargaining', *Journal of Economic Behavior and Organization*, 3(4), 367-388.

Hacking, Ian (1983), *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge University Press.

Hamermesh, Daniel S. (2013), 'Six decades of top economics publishing: Who and how?', *Journal of Economic Literature*, 51(1), 162-172.

Hansson, Sven Ove (1996), 'Defining pseudoscience', *Philosophia Naturalis*, 33, 169–176.

Hartmann, Stephan (1996), 'The world as a process: simulation in the natural and social sciences', in Rainer Hegselmann, Ulrich Müller, and Klaus Troitzsch (eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view*, Dordrecht: Kluwer, pp. 77-100.

Hausman, Daniel M. (2018), 'Philosophy of economics', in Edward N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy*, Fall 2018 Edition, https://plato.stanford.edu/archives/fall2018/entries/economics/.

Hertwig, Ralph and A. Ortmann (2001), 'Experimental practices in economics: A methodological challenge for psychologists?', *Behavioral and Brain Sciences*, 24, 383-451.

Hilpinen, Risto (2011), 'Artifact', in Edward N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/archives/win2011/entries/artifact/.

Hogarth, Robin M. (2005), 'The challenge of representative design in psychology and economics', *Journal of Economic Methodology*, 12(2), 253-263.

Huff, Toby E. (1982a), 'On the methodology of the social sciences: A review essay - part II', *Philosophy of the Social Sciences*, 12(1), 81-94.

Huff, Toby E. (1982b), 'On the methodology of the social sciences: A review essay - part III', *Philosophy of the Social Sciences*, 12(2), 205-219.

Ivarola, Leonardo (2017), 'Socioeconomic processes as open-ended results. Beyond invariance knowledge for interventionist purposes', *Theoria. An International Journal for Theory, History and Foundations of Science*, 32(2), 211-229.

Jimenez-Buedo, Maria (2015), 'The last dictator game? Dominance, reactivity, and the methodological artefact in experimental economics', *International Studies in the Philosophy of Science*, 29(3), 295-310.

Jimenez-Buedo, Maria and F. Guala (2016), 'Artificiality, reactivity, and demand effects in experimental economics', *Philosophy of the Social Sciences*, 46(1), 3-23.

Joffe, Michael (2017), 'Causal theories, models and evidence in economics: some reflections from the natural sciences', *Cogent Economics & Finance*, 5, 1280983.

Landry, Craig E., A. Lange, J. A. List, M. K. Price, and N. G. Rupp, (2006), 'Toward an understanding of the economics of charity: Evidence from a field experiment', *Quarterly Journal of Economics*, 121, 747-782.

Lawson, Tony (2009a), 'Heterodox economics and pluralism: reply to Davis', in Edward Fullbrook (ed.), *Ontology and Economics: Tony Lawson and His Critics*, London and New York: Routledge, pp. 93-129.

Lawson, Tony (2009b), 'The current economic crisis: its nature and the course of academic economics', *Cambridge Journal of Economics*, 33, 759-777.

Lawson, Tony (2015a), 'A conception of social ontology', in Stephen Pratten (ed.), *Social Ontology and Modern Economics*, London: Routledge, pp. 19-52.

Lawson, Tony (2015b), 'Methods of abstraction and isolation in modern economics', in Stephen

Pratten (ed.), *Social Ontology and Modern Economics*, London: Routledge, pp. 315-337.

List, John A. (2009), 'An introduction to field experiments in economics', *Journal of Economic Behavior & Organization*, 70, 439-442.

Lohse, Simon (2017), 'Pragmatism, ontology, and philosophy of the social sciences in practice', *Philosophy of the Social Sciences*, 47(1), 3-27.

Mahner, Martin (2013), 'Science and pseudoscience how to demarcate after the (alleged) demise', in Massimo Pigliucci and Maarten Boudry (eds.), *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, Chicago: University of Chicago Press, pp. 29-43.

Mäki, Uskali (2005), 'Reglobalizing realism by going local, or (how) should our formulations of scientific realism be informed about the sciences?', *Erkenntnis*, 63, 231-251.

Mäki, Uskali (2009), 'Realistic realism about unrealistic models', in Harold Kincaid and Don Ross (eds), *The Oxford Handbook of Philosophy of Economics*, Oxford: Oxford University Press, pp. 68-98.

Meyer, Bruce D. (1995), 'Natural and quasi-natural experiments in economics', *Journal of Business and Economic Statistics*, 13(2), 151-161.

Mitchell, Sandra (2009), 'Complexity and explanation in the social sciences', in Chrysostomos Mantzavinos (ed.), *Philosophy of the social sciences: philosophical theory and scientific practice*, Cambridge, New York: Cambridge University Press, pp. 130-145.

Morgan, Jamie and H. Patomäki, (2017), 'Contrast explanation in economics: its context, meaning, and potential', *Cambridge Journal of Economics*, 41, 1391-1418.

Morrison, Margaret and M. S. Morgan, (1999), 'Models as mediating instruments', in Mary S. Morgan and Margaret Morrison (eds.), *Models as mediators: Perspectives on natural and social science*, Cambridge: Cambridge University Press, pp. 10-37.

Pratten, Stephen (2015), 'The scope of ontological theorising', in Stephen Pratten (ed.), *Social Ontology and Modern Economics*, London: Routledge, pp. 69-94.

Psillos, Stathis (2005), 'Underdetermination thesis, Duhem-Quine thesis', in Donald Borchert (ed.), *Encyclopedia of Philosophy,* Vol. 9, 2nd edition, Detroit: Macmillan, pp. 575-578.

Ramsey, Jeffry L. (1994), 'Ideal reaction types and the reactions of real alloys', *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1, 149-159.

Reiss, Julian (2016), 'Thought experiments in economics and the role of coherent explanations', *Studia Metodologiczne*, 36, 113-130.

Rosenberg, Alexander (2015), *Philosophy of Social Science*, 5th edition, Boulder, CO: Westview Press.

Ruzzene, Attilia (2015), 'Policy-making in developing countries: from prediction to planning', *Journal of Economic Methodology*, 22(3), 264-279.

Santos, Ana C. (2007), 'The 'materials' of experimental economics: technological versus behavioral experiments', *Journal of Economic Methodology*, 14, 311-337.

Santos, Ana C. (2009), 'Behavioral experiments: how and what can we learn about human behavior', *Journal of Economic Methodology*, 16(1), 71-88.

Santos, Ana C. (2010), *The Social Epistemology of Experimental Economics*, Routledge: London & New York.

Selten, Reinhard (2003), 'Emergence and future of experimental economics', in Maria C. Galavotti (ed.), *Observation and Experiment in the Natural and the Social Sciences*, Dordrecht: Kluwer, pp. 63-70.

Siakantaris, Nikos (2000), 'Experimental economics under the microscope', *Cambridge Journal of Economics*, 24, 267-281.

Smith, Vernon L. (1962), 'An experimental study of competitive market behavior', *The Journal of Political Economy*, 70(2), 111-137.

Smith, Vernon L. (2001), 'From old issues to new directions in experimental psychology and economics', *Behavioral and Brain Sciences*, 24(3), 428-429.

Smith, Vernon L. (2010), 'Theory and experiment: what are the questions?', *Journal of Economic Behavior & Organization*, 73, 3-15.

Søberg, Morten (2002), 'The Duhem-Quine thesis and experimental economics: A reinterpretation', *Statistics Norway,* Research Department Discussion Papers, 329.

Søberg, Morten (2005), 'The Duhem-Quine thesis and experimental economics: A reinterpretation', *Journal of Economic Methodology*, 12(4), 581-597.

Starmer, Chris (2005), 'Normative notions in descriptive dialogues', *Journal of Economic Methodology*, 12(2), 277-289.

Sugden, Robert (2016), 'Ontology, methodological individualism, and the foundations of the social sciences', *Journal of Economic Literature*, 54(4), 1377-1389.

Suppes, Patrick (1962), 'Models of data', in Ernest Nagel, Patrick Suppes, and Alfred Tarski (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford: Stanford University Press, pp. 252-261.

Suppes, Patrick (1985), 'Explaining the unpredictable', *Erkenntnis*, 22, 187-195.

Suppes, Patrick (2008), 'A revised agenda for philosophy of mind (and brain)', in Wilhelm K. Essler, Dagfinn Follesdal, and Michael Frauchiger (eds.), *Lauener Library of Analytical Philosophy*, Frankfurt: Ontos Verlag, pp. 19-51.

Thoma, Johanna (2016), 'On the hidden thought experiments of economic theory', *Philosophy of the Social Sciences*, 46(2), 129-146.

Van Vugt, Mark (2001), 'Self-interest as self-fulfilling prophecy', *Behavioral and Brain Sciences*, 24(3), 429-430.

Weber, M. (1949), *The Methodology of the Social Sciences*, trans. by E. A. Shils and H. A. Finch. Glencoe, IL: The Free Press.

Wiggins, James A. (1968), 'Hypothesis validity and experimental laboratory methods', in Hubert M. Blalock, Jr. and Ann B. Blalock (eds.), *Methodology in Social Research*, New York: McGraw-Hill, pp. 390-427.

María Caamaño-Alegre is Associate Professor of Philosophy of Science at the University of Valladolid (Spain) (mariac@fyl.uva.es).

José Caamaño-Alegre is Associate Professor of Political Economy and Public Finance at the University of Santiago de Compostela (Spain) and member of the Governance and Economics research Network (GEN) (jose.caamano@usc.es).