UNIVERSITA' DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI SCIENZE ECONOMICHE
"Hyman P. Minsky"
Via dei Caniana 2, I-24127 Bergamo, Italy
Tel. +39-035-2052501; Fax: +39-035-2052549

Quaderni di ricerca del Dipartimento di Scienze Economiche "Hyman P. Minsky"

Anno 2002 n.8

On the Robustness of Economic Models

Francesco Guala e Andrea Salanti

Comitato di Redazione

Riccardo Bellofiore
Annalisa Cristini
Giancarlo Graziola
Andrea Salanti
Luigi Bonatti
Piero Ferri
Riccardo Leoni

- La Redazione ottempera agli obblighi previsti dall'art.1 del D.L.L. 31.8.1945, n.660 e successive modificazioni.
- Le pubblicazioni del Dipartimento di Scienze Economiche dell'Università di Bergamo, consistenti nelle collane dei Quaderni e delle Monografie e Rapporti di Ricerca, costituiscono un servizio atto a fornire la tempestiva divulgazione di ricerche scientifiche originali, siano esse in forma definitiva o provvisoria.
- L'accesso alle collane è approvato dal Comitato di Redazione, sentito il parere di un referee.

On the Robustness of Economic Models*

Francesco Guala
University of Exeter (UK)

Andrea Salanti
University of Bergamo (Italy)

Abstract. We investigate the different ways in which the results of theoretical models can be 'robust'. We identify three kinds of 'robustness': (1) robustness to changes in the model's idealisations; (2) robustness to changes in the 'background' conditions; (3) robustness to changes in the implied causal mechanism. Each of these is discussed and illustrated by means of examples from economic practice.

Address for correspondence:
Dr Francesco Guala
Centre for Philosophy of the Social Sciences
University of Exeter
Amory Building, Rennes Drive
Exeter EX4 4RJ
UK

Email:f.guala@ex.ac.uk

* Grants from, respectively, the Cognitive Science Laboratory of the University of Trento and the Department of Economics of the University of Bergamo are gratefully acknowledged. We would like to thank Alessandro Turrini, two anonymous referees, and the audience of the 2001 HES conference for some useful comments on an earlier version. The usual caveats apply.

1. On the Robustness of Economic Models

Robustness is supposedly a desirable attribute, and scientific progress is often related (among other things) to the discovery of increasingly robust theoretical and/or empirical relations. Like other normative terms, however, robustness works as a comprehensive 'umbrella' under which various different ideas are subsumed. Since different kinds of progress are associated with different kinds of robustness, it is important to distinguish between them properly.

Gibbard and Varian (1978) provide one of the few philosophical accounts of economic modelling that makes explicit use of the idea of robustness. They characterise an important class of models as "caricatures"[1] and ask when and why such caricatures may be "helpful in understanding a situation":

One way is by yielding conclusions that are *robust*, in the sense that they do not depend on the details of the assumptions. When a theorist applies a model that caricatures a situation, one hypothesis he may entertain is this: the conclusions of the applied model roughly depict some features of the situation, and that is because (1) the assumptions of the model caricature features of the situation, and (2) the conclusions are robust under changes in caricature. A principal way of testing this hypothesis may be to try out models with disparate caricatures of the same complex aspect of reality (pp. 674-5, italics in the original).

These and other remarks in the same essay suggest that the kind of robustness Gibbard and Varian have in mind has to do with changes in the model's 'idealisations', in the terminology that we shall use throughout this paper. Surely this captures an important

1 To be more precise, they first distinguish between "ideal" and "descriptive" models, just to say that their attention will primarily focus on the latter. Then descriptive models are said to be "approximations" or "caricatures". The difference between the two is treated as one of degree rather than of substance. In their own words: "A caricature differs from an approximation … not only in its simplicity and inaccuracy, but in its deliberate distortion of reality. When a model is applied as an approximation, the goal is to distort as little as is compatible with a given degree of simplicity and tractability. A caricature involves deliberate distortions of reality for other reasons — to isolate the effects of one of the factors involved in the situation, or to test for robustness under changes of caricatures" (p.676).

virtue of models, but it is by no means the only sense of robustness that can be detected in the economic literature (or hinted at during presentations at seminars and workshops). First of all, consider that models caricature real entities and systems in roughly two ways, by distortion and (simultaneously) by omission of some aspects of reality.2 Thus, a distinction between robustness to changes in the model's idealisations (sections 1-2) and robustness to changes in the 'background' conditions (section 4) follows quite naturally. Moreover, if we focus on the causal mechanism(s) depicted by means of the model (and supposedly at work in reality), a third notion of robustness, with respect to changes in the causal mechanisms themselves, may be envisaged (section 5). In the remaining sections we focus on the practical consequences of robustness criteria. In particular, we look at the difficulty of drawing inferences from robustness in one domain to robustness in another one (section 6), and to the testing procedures that should be put at work in order to evaluate the robustness of economic models (section 7). Finally, we must stress that robustness has been discussed (albeit under different headings) in the econometric literature and, more recently, in philosophical studies of causation. Our paper follows in the path of these two literatures, trying to extend the analysis to the area of economic modelling. Section 3 discusses the link between causation and modelling, and highlights the links between our paper and previous work in these areas.

2. Robustness to changes in the model's idealisations

One idealises by describing a real-world feature in a very specific, sometimes even extreme or idiosyncratic, fashion. An aspect of the model is characterised so as to instantiate just one of the various forms that that feature can take in the real world. A typical idealisation is involved for example when a result 3 (e.g. the existence of an

---

2 We shall refer to these two distinct aspects of model-building activity with the terms '*idealisation*' and '*abstraction*' respectively. We should warn the readers that this terminology is by no means standard in the philosophy of science and the methodology of economics literature. The lack of a shared convention can be appreciated for example by looking at the essays in Hamminga and De Marchi (eds. 1994). In this paper we follow, e.g., Cartwright (1989, Ch. 5).

3 In this paper we follow scientific jargon and use the term 'result' somehow loosely, to refer to theorems, predictions, existence or non-existence proofs, but also e.g. to the demonstration of the existence of a (causal) relation between modeled entities.

equilibrium) is proved to hold for an economy with a given number of agents, although we in fact know that the number of agents in the real world will almost certainly differ from the one in the model. Or when we analyse an economy populated by risk-neutral individuals – whereas different, more varied attitudes towards risk might conceivably be the case in reality.

A most famous – and, arguably, most drastic – idealisation dates back to the birth of marginalist economic theory. According to equilibrium models, the price of a good is determined in a competitive market by the supply of that good, on the one hand, and the tastes of the consumers and their budget constraints, as embodied in demand schedules, on the other. Orthodox theory says relatively little about the institutional arrangements that allow for the formation of a 'clearing price'. The story of a Walrasian auctioneer matching supply and demand by *tâtonnement* is clearly an idealised assumption which 'stands for' a whole range of different institutions at work in real-world markets. Until a few years ago – that is, before the development of auction theory – it was commonly assumed that the results of equilibrium theory would apply across a wide range of different market institutions. Some of the best work in auction theory has been devoted to investigate the robustness of equilibrium theory to changes in the institutional arrangement. Perhaps not so surprisingly, it has been theoretically demonstrated that institutions indeed matter for the formation of equilibrium prices in competitive markets. Some mechanisms are quite efficient at that, others less, still others are not efficient at all.[4] Equilibrium models are only partially invariant to changes in the Walrasian auctioneer idealisation.

The progressiveness associated with increasing robustness to changes in idealisations should be apparent. The idea is that a given result has been rigorously demonstrated with respect to a very specific model-economy only for reasons of simplicity, clarity, or mathematical tractability. But the validity of the result itself should not *depend on* any of the specific idealisations used in the proof. Often, this independence will be self-evident to the informed reader. At other times, the author promises to be able to provide a more general proof, or points to a body of literature (for example in pure mathematics) providing the basic tools to construct a general demonstration of the required sort.

---

4 Cf. McAfee and MacMillan (1987) and Milgrom (1989) for a survey of results.

3. Robustness in the 'core' and in the 'boundaries'

An important function of theoretical models of the above kind is to provide concrete examples of the way in which causal relations can hold between factors or variables. This is done by constructing artificial worlds in which such relations hold, and by demonstrating that they do therein. Thus, for example, Akerlof's (1970) famous 'market for lemons' is a model economy (featuring only two types of traders, with highly idealised utility functions, four kinds of car, etc.) in which the variable symmetric/asymmetric information about the quality of potentially tradable goods respectively promotes/prevents effective trade.

Akerlof demonstrates that this is the case by running simple algebraic calculations for each case. Nevertheless, he is asking us to believe that his result (the causal relation between the asymmetrical distribution of information and the level of trade) will not change when the core causal variable (the distribution of information) takes different values from those considered in the model.5 Note that Akerlof is interested in establishing a fundamental discontinuity between the case of perfectly uniform information and the presence of even the slightest form of information asymmetry. He is concerned, among other things, with demonstrating that a relation among variables that was customarily assumed to hold well (price differentiation for goods of different but recognisable quality) is instead fragile to changes in the variables themselves (that is, when buyers are uncertain about quality).
Thus, we have to distinguish between the invariance of a causal relation across changes in other circumstances and the invariance of that same relation to changes in the values taken by its variables (given the *same* circumstances).6 The latter is strictly speaking a sub-case of idealisation robustness, since we are in both cases concerned about variations in the explicit features of a model. But not all features have the same status. The distinction

5 More generally, as Sugden (2000) points out, Akerlof implicitly asks us to believe that changes in the idealised and abstract features of the model will not matter for the result itself – which is supposed to continue to hold also in neighbouring systems that are 'close' to Akerlof's model-world (i.e. that are identical up to changes in the idealised and/or omitted features). Or, in other words, that his model is both 'idealisation robust' (in the sense specified above) and 'abstraction robust' (in the sense specified below).
6 We have borrowed this distinction from Jim Woodward (1997; 2000), who specifically applies it to causal relations *in the real world*. On causation in models vs. in the real world, see below.

between ‘core relations’ and ‘boundary’ conditions is an old one in economic thinking.7 Most models are built in order to demonstrate that a rigorous relation holds between certain selected features (or variables) of the model itself. Ideally we would like a result to be completely independent from boundary assumptions, but to a certain extent dependent on the core assumption of the model. If a result holds entirely independently of a core assumption, then that assumption seems pretty useless.8 Core variables must be robust, then, *but only to changes in their values*. Indeed, the robustness of the ‘core’ relations to changes in the value of their variables comes first in the scale of appraisal criteria.9 Only when such robustness has been established in a particular set-up, robustness to changes in the remaining parts of the set-up itself (idealisation-robustness) or in its implicit parts (abstraction-robustness) is taken into consideration as a further element of interest.10

Notice, however, that what may represent a core variable or relation in some context, can also be legitimately regarded as a boundary variable or relation in a slightly different context. For example, risk attitudes among traders are clearly a boundary variable within Akerlof's model of ‘the market for lemons’. Simple linear utility functions are used for reasons of simplicity and in order to focus on the main issue, i.e. the consequences of asymmetric information among traders on the very existence of market equilibrium. But risk attitudes, to remain in the realm of the economics of information, come to the forefront in principal-agent models, where agent's risk aversion is crucial in determining the loss of efficiency, due to the necessity of having the agent to bear all the risk connected with the probabilistic distribution of outcomes for a given level of his/her effort11. In these models, assumptions about risk attitudes constitute (some of) the core variables, whose importance

---

7 With variations in terminology, it goes back to at least Machlup (1955). See Mäki (2000, p. 328) and the references therein.

8 We must thank Roger Backhouse for pointing this out.

9 This does not mean that the core causal relations should be *absolutely* robust to changes in the values of their own variables. In the so-called special sciences (for ontological as well as epistemic reasons) scientists are often content with *partially* invariant relationship (Woodward, 2000).

10 A lot of philosophy of science misleadingly focuses on the scope and range of scientific generalisations, whereas in fact disciplines like economics or medicine are primarily aimed at identifying genuine causal relations. Generality is less important than causal efficacy, and this is reflected in the preference given to robustness in the core relations of a model (a typical mark of causation) rather than in the boundary conditions (which determine the scope of a model's applicability).

11 Cf., for instance, Kreps (1990, ch. 16).

for the derivation of certain results across various settings is apparent. The distinction between core and boundary variables, therefore, is pragmatic to the extent that it depends on the interests of the modeller (the sort of result that he/she intends to demonstrate).

4. On the notion of causation

Some clarification is now due about the concept of causation used here: first, we think it is acceptable to talk of causation as occurring both in models and in the real world, and accordingly we shall often shift from one level to the other in our discussion. The reason for this is simple: by 'models' we (like many philosophers and scientists) intend primarily 'model-worlds' – sets of objects endowed with properties and relations, some of which will typically be causal in character.12 Some such model-worlds will be concrete, others will be purely abstract (mathematical set-theoretic structures, for example). Philosophers sometimes talk of models as purely linguistic entities, but this is misleading, as syntactical 'models' cannot always be interpreted causally, or at least not unequivocally so, whereas models must provide us with information about causal mechanisms in the first place. Take some typical symmetric relationship such as the ideal gas law ($PV=nRT$) or the quantity theory of money ($MV=PT$): both pick up by implicit definition a number of model-objects. For some of them, it will be true that changes in temperature cause changes in volume (or changes in money supply cause changes in prices) but this will not apply to *all* the systems described by the two equations. Causation is a property of systems, not of equations or models-as-linguistic-entities.13

Secondly, and roughly, we subscribe to the view that X is a cause of Y if (given the right set of background conditions) X can be used to change the value of Y, and these changes are brought about *via* X (and the intermediate factors connecting X to Y) only (see e.g.

12 If we take models to be *structures* – sets of abstract or concrete entities and relations – it becomes natural to think of causation as applying to both models and real-world systems. On the ambiguous boundaries between model-based demonstrations, simulations, and experiments see Guala (2002), Morgan (2002), and Boumans and Morgan (2001).

13 We have postponed a definition of models until this section simply because in many instances the ambiguity between linguistic entities and objects does not matter.

Cartwright 1989; Humphreys 1989; Woodward 2000).14 Thus, for example, in a sufficiently large economy a change in an individual agent's income causes a change in her consumption. This might sound odd in the case of general equilibrium models, where more primitive factors like tastes, technology and endowments seem to do all the important explanatory work. In this paper we follow Hausman (1990) in holding that the factors determining a general equilibrium are separable and can be causally ordered. Where many different variables play a causal role, they can be either highlighted as primary or relegated among the 'other things' in the background, depending on the modeller's interests. Moreover, there is in principle no conceptual difficulty in granting causal role to intermediate entities like supply and demand schedules, provided they are connected to their effect in the right way. Equilibrium explanations can thus be construed as causal ones in the sense of causation specified above.

5. Robustness to changes in the 'background' conditions

The main feature distinguishing idealisation robustness from the one discussed in this section is that idealisations are explicitly defined in the model itself. Most often, they are also clearly highlighted as such. But a model-caricature can betray reality also by deliberately *omitting* some of its features. In this case it is appropriate to talk of an abstraction, in the etymological sense of 'taking away', stripping an entity or system of some of its properties. The crucial point is that the abstracted features do not figure in the model-economy, although they are sometimes mentioned in the informal commentary that goes with it. The omitted aspects are kept 'in the background' as conditions, factors, etc. that are supposedly irrelevant for the purpose at hand.15

---

14 Of course this is not a reductive definition of causation (something that has defied the efforts of generations of philosophers), but merely a partial articulation. For a thorough discussion of various theories of causation, see Hausman (1998a).

15 Economists often express this idea by saying that a theoretical result is valid *ceteris absentibus* (other things being absent) or, less appropriately, *ceteris paribus* (other things being equal). Morgan (2001) discusses at length the relation between models and the informal commentaries (or 'stories') that accompany them. Against the received wisdom, she argues that background stories are essential for the interpretation and usage of models.

Some of the background factors can usually be accounted for by purely theoretical means, by making a model more complicated (that is, by 'concretising' the model). This happens when, e.g., the model of a national economy without the foreign sector is built for heuristic purposes. A number of well-known modelling techniques allow economists to add the omitted feature back and thus test the robustness of the result to changes in the level of abstraction. If carried on, however, eventually the process of concretisation will inevitably lead from the theoretical to the empirical realm.16

To illustrate this point we can use the same examples as in the previous sections, but focusing this time on the actual material set-up – the physical arrangements, boundary conditions, etc. – in which a market mechanism is implemented. Experimental economists have shown that certain institutions are quite fragile and others considerably robust to changes of this kind. The frequency of bidding rounds, the way in which buyers are seated in the room, or whether they interact via a computer rather than orally or in writing, for example, may make a substantial difference in terms of market outcome. The way in which the system is 'shielded' from disturbances, how smoothly the information flows, the 'transparency' of the rules, etc. are not included as explicit variables in auction theory but constitute the main concern of experimenters and mechanism designers when it comes to actual application. 17 Models of auction systems omit these factors because theoretical language abstracts from them, despite the fact that eventually they affect the theory's scope of application.

These problems are by no means limited to economics. Students of experimental science have provided several similar examples from physics, a 'classic' being the construction of lasers. LASER is an acronym standing for Light Amplified by Stimulated Emission of Radiation. A laser can be described at several levels of abstraction. In general, a laser is a machine able to produce a highly coherent radiation, a piece of apparatus with the capacity of 'lasing'. Its structure can be described at a high level of abstraction in the following way (taken from a textbook):

---

16 This does not mean that it *is* normally carried on to such consequences, nor necessarily that it *should* (although of course there are often good reasons to do that).

17 More on this in Guala and Salanti (2001) and Guala (2001).

> What they [lasers] have in common is an active material (e.g. the ruby) to convert some of the energy into laser light; a pumping source (e.g. the flashtube) to provide energy; and mirrors, one of which is semi-transparent, to make the beam traverse the active material many times and so become greatly amplified.18

As Nancy Cartwright (1989, Ch. 5) points out, the kind 'laser' is identified by such a structure, rather than by the particular materials (some of which are mentioned between brackets in the above quote, by way of examples) used in specific instantiations. Thus, lasing can be reproduced across a range of circumstances, using different materials, different kinds of shielding, different sources of power, different triggering mechanisms, and so on. Students of experiments like Harry Collins (1985) have illustrated convincingly how the job of laser-construction is by no means entirely dictated by theoretical accounts of the abstract causal structure. The choice of the 'right' materials, and the way in which the various components are put together, are equally important. But usually there is no *one* way in particular in which the job has to be done.

Scientists use abstract theoretical models instead of (or alongside) painstaking descriptions of a specific experimental set-up precisely because they hope (often correctly so) that a certain causal mechanism be at least *relatively* robust to changes in the background circumstances, the choice of materials, the boundary conditions etc. that were used in the original, benchmark experiment or prototype.19 Clearly, we are here discussing abstraction-robustness in empirical contexts. Such a shift is natural when discussing abstraction robustness, because the omitted properties can rarely if ever be *all* added back to a model by theoretical reasoning. The process of concretisation will most often start in the purely theoretical realm, and almost inevitably be carried on by experimenting or simulating with material systems. Idealisation and abstraction robustness, in other words, are often supposed to hold across model- and real-worlds.

---

18 L. Allen and D.G.C. Jones (1967) *Principles of Gas Lasers*, New York: Plenum Press; quoted in Cartwright (1989, p. 216).

19 That 'relatively' must be stressed here. Surely causal mechanisms robust across a wide range of background circumstances and material instantiations are particularly valuable, but are also quite rare in

6. Robustness to changes in the causal mechanism itself

There is however yet another kind of model-robustness that is *prima facie* irreducible to the other kinds of robustness-talk. This usage of the term is quite common in economics, where progress is often identified with the proliferation of models featuring *different* causal mechanisms. As an example, let us take the case of the theory of imperfect competition, a field of research where caricatures abound.20 The number of different representations of the situation facing the imperfectly competitive firm is portentous, and in different models we may find exactly the opposite specification of the game that the agents are supposed to play.

Let us focus in particular on two models that employ two-stage oligopolistic games. Kreps and Scheinkman (1983) model the case of perfectly substitutable commodities as a game with quantity precommitment in the first stage and price competition in the second. All this is done in order to show that, under appropriate assumptions, there exists a unique equilibrium of Cournot-like type. Benassy (1986), who is instead concerned with product differentiation, investigates the existence of Bertrand-Edgeworth equilibria with differentiated commodities. After having approvingly referred just to the previous article as "a recent and interesting contribution on timing in the pure substitutes case" (n. 6, p. 58), in the last section of his paper he presents a two-stage game where "the idea is that agents choose prices in a first stage, and then production in a second stage" (p. 72).

The proliferation of models of this kind may be interpreted as evidence of our inability to detect the best representation of the situation we want to investigate and/or the true mechanism at work. Often empirical (casual, field, or experimental) evidence does not help in choosing the best model. Usually the problem is not a complete lack of some form of empirical evidence, but the lack of *decisive* evidence. So how can the robustness of a result

---

the special sciences (that is, in all science with the exception of fundamental physics). See also footnote 9 above and Woodward (2000).

20 Since the early pioneering contributions to this field it clearly emerged that the problem of modeling agents' conjectures may receive a plurality of answers, each one with its own pros and cons. Some other topics in modern microeconomics, however, could have served as well for our scope. Let us mention, for instance, the microeconomic theory of information, signaling, contracts, incentives and mechanism design.

to radical changes in the theoretical causal mechanism that produced it be considered progressive? For each real-world situation, at most *one* of these models can tell the true causal story; the fact that we have many alternatives converging on the same result does not, *per se*, imply that we are getting closer to the truth at all. 21

Perhaps theoretical economists rely on an argument of the same sort as that used for 'robustness to changes in idealisations': if conclusions do not change very much under different assumptions, this means that that particular detail is not so important after all. In some cases, this may be true: if the only aim is to predict, assuming that nothing much will change in the institutional set-up, an instrumentalist attitude may be justified. But in many other circumstances a 'detail' such as the kind of game that agents are supposed to play *is* important. This is surely the case when it comes to intervention, for example when the legislator intends to implement some sort of regulation. Very often, indeed, only knowledge of the structure of the game will allow us to predict the ways in which the agents will react to a change in rules and/or incentives.

Proof that a result can be achieved by means of various alternative mechanisms, however, may convey progress in at least two other (non-instrumentalist) ways. First, as some social scientists recognise, different mechanisms may be at work in the real world. It is just natural to expect, then, that "when the competitive paradigm is abandoned, and replaced by more realistic assumptions", the new paradigm will look like a "highly coloured patchwork of often unrelated investigations". This will be necessary "[to] mirror the incredible complexity of real market phenomena" (Gabsiewicz and Thisse 1999, p. lii). Secondly, suppose that modelling is conceived in a constructivist way, as ultimately aimed at mechanism-design and application. 'Functional' or 'goal-oriented' reasoning22 plays a big role in some areas of economics. Such reasoning is typical of technological enterprises. Technology aims at the construction of devices able to produce reliably a certain effect or perform a given function. Technological reasoning therefore often proceeds 'backwards'

21 For an essentially identical argument in an empirical context, see Cartwright (1991). In the terminology of Backhouse (1997, pp.100–01), we may say that robustness to the causal mechanism increases the *generality* of a result, but diminishes its *scop*e. (Where the 'scope' is the class of *identifiable* situations of applicability.)

22 Not to be confused with the 'functional explanations' that are so popular and controversial in some branches of the social sciences.

from the desired outcome to the actual set-up that can bring it about. Scientific theories of course play an important role in this heuristics, by suggesting causal schemes that might help the engineer in completing his task.23

Auction theory can provide us once more with an example. According to an important seminal paper by Vickrey (1961), auctions with *prima facie* different rules can lead to exactly the same result. An ascending English auction in fact is strategically identical, from the viewpoint of game theory, to a sealed-bid Second-Price auction – and both formats have the (theoretical) capacity of generating Pareto-efficient allocations. At the level of application, the fact that two different mechanisms can lead to the same result increases the range of options of the mechanism designer. Combining theoretical reasoning with considerations of robustness on the other dimensions, applied economists can identify the institutional arrangement that is better suited to the task at hand.24

Or, to move to an even more general level of analysis, let us take the concept of efficient price that is so central in modern economic theory. In principle, there are many ways in which a Pareto-efficient price system can be determined: a central planner informed about all the traders' preferences, for example, might simply set the price at the appropriate level, identified after some (very demanding) calculations.25 Orthodox economists, however, are quite sceptical about allocation procedures of this kind, first of all because most real central institutions, no matter how powerful, rarely have the capacity (and/or the incentive) to gather enough relevant information and to perform the required calculations. Instead, these economists suggest that decentralised systems of allocation like competitive markets can under the 'right' circumstances generate efficient prices and allocations. A desirable socio-economic *function*, in other words, can in theory be performed by different mechanisms. A result can thus be obtained in many different ways, and be 'robust' across a range of different causal mechanisms – although it is to be expected that in practice some of these mechanisms will be difficult or even impossible to implement. As mentioned

23 For an interpretation of economics along these 'constructivist' lines, see Simon (1969).

24 A detailed case study of mechanism design and implementation is provided in Guala (2001).

25 Alternatively, the same planner could try to play the role of a Walrasian auctioneer proceeding by trials and errors.

above, for a causal mechanism to be implemented robustness must be demonstrated not only across model-worlds, but also from models to the real world.

The famous socialist calculation debate of the thirties is a particularly interesting case in point. During that debate, following Barone's (1908) seminal contribution, the analytical tools of marginalist economics were applied to different institutional frameworks, in order to demonstrate the universal validity of the 'economic calculus' of resource allocation. One result, namely the universal validity of the 'economic calculus' for an optimal solution to the problem of resource allocation, was taken at the time to be robust to changes both in the model's idealisations (the usual assumptions of general equilibrium analysis) and in the implied causal mechanisms (a competitive market process as opposed to a centralised allocation system). The problem was that both 'mechanisms' were handled with models that, as we can now appreciate, were not particularly robust to changes in some important 'background' conditions, such as information, incentives, number of competing firms, etc.).26

7. Induction from robustness

Drawing conceptual distinctions may be an enjoyable philosophical game in itself, but is ultimately of little interest unless the distinctions bear some methodological weight. In the rest of this paper we shall try to put the above taxonomy at work in order to explore its practical implications. Let us start with a seemingly trivial remark: applied economists or econometricians place a very different emphasis on the testability (or falsifiability)27 of models in comparison with pure theorists. Philosophical accounts often follow

---

26 One of the few scholars who seemed to have realized the practical irrelevance of such an analytical result was Schumpeter (1992, Ch. 16), who pointed out, quite straightforwardly, that "a probable maximum ... as such establishes the economic rationality of that type of socialism exactly as the competitive maximum establishes the rationality of competitive economy. *And in neither case does this mean very much*" (p. 184, n. 9, italics added).

27 We shall leave aside here the two main questions about falsificationism in economics, that is: (1) whether economists have taken falsificationist precepts seriously or have simply paid lip service to them; and (2) whether falsificationism could be ever coherently and successfully practiced in economics (and in other sciences, for that matter). The debate on these issues has fostered a huge amount of literature that would be impossible even to simply mention here. For a brief review see Salanti (1998).

practitioners in distinguishing sharply between purely theoretical and applied models, and in proposing different criteria of appraisal for each kind. Whereas purely theoretical models should be heuristically fertile, simple, elegant, or capture the 'essence' of economic reality, applied models are to be appraised on the basis of their falsifiability, degree of confirmation, or various statistical criteria as specified in econometrics textbooks. 28

This methodological dualism appears misguided in the light of the above discussion. In this paper we have argued that appraising a model involves primarily an evaluation of its robustness to changes of various sorts, and that model-robustness is a sub-category of a more general family of robustness notions that stretch across empirical and theoretical contexts29. Robustness (in all its manifestations) is a virtue of theoretical *and* applied models alike, and by focusing on it we can achieve a unified, encompassing account of economic criteria of appraisal.

The kernel of truth in methodological dualism lies in the fact robustness proofs at the theoretical level do not justify inductive inferences to robustness in the empirical realm. Of course in some instances robustness across empirical contexts is in effect supported a priori by putting forward theoretical considerations, but robustness will stretch only as far as the relevant theory does. If people donate their blood for free, surely they will donate more if they are paid to do so (i.e. if monetary incentives are added to altruistic motivation)? In reality, they do not (Titmuss, 1970). Or, to take another well-known example, people display inconsistent patterns of choice in preference reversals experiments. Preference

---

28 See, for instance, Boland (1989, pp. 2-8) or Hausman (1992, pp. 78-82). The distinction often takes the form of a recognition of the peculiar methodological problems of macroeconomics and/or macroeconometrics: cf. Hicks (1979, pp. viii-ix), Vercelli and Dimitri (1992, part I), Backhouse (1995 and 1997, Chs. 11-13), Granger (1999), Backhouse and Salanti (2000, Introduction), Hoover (2001b).

29 Sugden (2000) implicitly proposes a dualist account of appraisal criteria in theoretical vs. empirical contexts. His main message is that putting theoretical models at work involves making inductive inferences. Robustness proofs are an obvious way of justifying such inductions from a specific model to other neighbour models in which some of the assumptions, idealisations, abstractions, etc. are relaxed. According to Sugden, however, robustness arguments are able to support model-to-model inductive inferences only. Model-to-world inferences are/should be built on different grounds, namely their *credibility*. But even granted that we can provide some epistemic ground to the notion of credibility, this account is incomplete: robustness should hold not only across models, but across real-world circumstances as well. We want relations that can be relied upon not only in more than one modelling context, but in several empirical contexts too – and how can credibility grant it for us?

reversals tend to diminish when the same task is repeated and the monetary incentives raised. Therefore, one is led to think that preference reversals surely must diminish also when the monetary incentives are raised, even though there is no repetition. But the frequency of reversals actually *increases* under the latter circumstances (Grether and Plott, 1979).

A priori arguments for robustness are often unreliable, when we move from theoretical to empirical contexts. When we build a machine (be it in the social or in the natural sciences), the steps in the concretisation procedure (the process of turning a theoretical into a material structure) must inevitably be carried out in practice, because we lack the knowledge of what would happen if certain idealisation were relaxed or certain background factors brought into the picture. Similarly, we do not know a priori whether certain structural relations will bring their robustness or invariance properties with them once they are transported from the realm of ideas into that of the material world. Or when we move from one real-world situation to another real-world situation that differs in some respects.

8. Robustness and the hypothetico-deductive model

Theoretical and applied models, to sum up, are appraised in the light of very similar criteria. In his advanced microeconomics textbook, David Kreps (1990) argues that

[t]he standard acid test is that the theory should be (a) testable and (b) tested empirically, either in the real world or in the lab. But many of the models and theories … have not been subjected to a rigorous empirical test, and some of them may never be. Yet, I maintain, models untested rigorously may still lead to better understanding, through a process that combines casual empiricism and intuition (p. 7).

We have tried to give a more precise articulation of the activity of model-appraisal than in Kreps's vague remarks. Model appraisal is a systematic activity, although as we have seen some room is left to mathematical and practical intuition. A central theme of this paper is that robustness is a key evaluative concept in both the theoretical and the empirical realm. Model-to-model robustness is an important dimension of scientific progress. However, we agree with Kreps that it will not automatically translate into model-to-world, or even world-to-world robustness. Eventually, only testing will do – but what kind of test?

The hypothetico-deductive model of testing is perhaps the most resilient legacy of the empiricist philosophy that dominated the middle part of the XX Century.30 It can be found, for example, in textbooks of research methods in the social sciences, despite the fact that it is rarely genuinely practised by scientists. If we are right, and economists (more or less consciously) appraise models in light of their robustness properties, then it follows that the HD model should be abandoned too. Consider the cornerstone of falsificationism, i.e. the idea that a single counterexample is enough to falsify a hypothesis (provided, of course, that the test has been properly implemented). This may be a good reason to reject a candidate for a universal law of nature, but not to reject a model that aims at capturing some (partially) robust relation. Appraisal in terms of robustness are more complicated than that.

First of all, one must trade off between different kinds of robustness – progress (more robustness) in one dimension does not necessarily guarantee progress in another dimension. Secondly, robustness appraisal is highly sensitive to pragmatic considerations and to the interests of the scientist. Surely economic models are not robust to changes in the fundamental constants of physics, but nobody seems to care particularly about that. In some cases, it is important that models be robust to changes in certain institutional arrangements, but in other cases it does not matter either. (For instance, when we are analysing and/or predicting the functioning of a single given market, and we are confident that the basic institutional arrangements are not changing in the course of our study. But if we aim at predicting the future of China's economy, we better take changes in institutional structure into account.)

Empirical testing of robustness will often take the form of investigations of the *limits* of economic models, i.e. of the domain in which their relations seem to hold to a certain degree of approximation. Outside such domains, it will be dangerous or unwarranted to apply a model of a certain kind. Notice that, contrary to Popper (1963) and others, this

---

30 Cf. Popper (1934) and Hempel (1966) and, for an application to economics, Friedman (1953).

attitude should not be conflated with instrumentalism or conventionalism. It is indeed perfectly legitimate to endorse a realist view of economic science, and at the same time to recognise that its models have a limited domain of applicability. Consider the most fundamental theoretical building block, the rational economic agent that is ubiquitous in neoclassical economic theorising. Its basic features (perfect rationality and coherent preferences, for instance) are just idealisations or refinements of the belief-desire model of action of folk psychology.31 It is reasonable to assume that models with rational agents will describe accurately the behaviour of economic systems where the 'background conditions' are right for the implementation of purely rational behaviour on the individuals' part. Indeed, a good deal of research is devoted today to the study of what makes rational behaviour possible in reality.32

One may argue that such an attitude towards the limitations of a scientific discipline is unscientific and unproductive. The right attitude, so the argument goes, is to try to supersede domain-specific theories and to encompass them by means of more general theories.33 But there is little hope that this strategy will pay off in disciplines like economics. First of all, to doggedly seek entirely general theories would most likely lead us to transgress the boundaries of the discipline and seek a reduction to a more fundamental science (be that psychology or, more likely, neurobiology and physics). But this may take too long to be done (if it is possible at all) and we still need economic science in the meantime. Moreover, causally interesting relations do exist at the typical level of theorising of economics (that is, relations among firms, consumers, prices and interest rates, as opposed to neurons, genes, or whatever), and we should learn to use them and refine them, not dispense with them in order to look for an unattainable Holy Grail. The precepts of generality and universality make a lot of sense for a science aimed at capturing the fundamental laws of nature – but this is not the aim of the social sciences, whose main goal is rather the discovery of pragmatically useful relations upon which intervention, control and social reform can be based.

---

31 Cf. Maki (1996) and Hausman (1998b).

32 See e.g. Plott (1995).

33 Popper (1957).

9. Final remarks

Keynes famously claimed that "economics is a science of thinking in terms of models joined with the art of choosing which models are relevant to the contemporary world".34 Apart from the somewhat outdated distinction between 'science' and 'art', this passage is misleading if taken to assert that the choice of the relevant models is a matter of pure intuition. Economists, to be sure, have developed some strategies directed at making such a kind of choice grounded on some shared (and hopefully sound) criteria.

Some of the ideas defended above are not new. Robustness to changes in idealisations, implying that for a particular result a number of different assumptions are – so to speak – interchangeable, vindicates a basic intuition behind Milton Friedman (1953) famous 'as if' interpretation.35 But this is to be expected, because interpreting robustness as a criterion of progress seems to be more in line with economists' practice. As we have seen, robustness in a model is considered as a first step towards finding out a structural relation in the real world. Robustness criteria of appraisal, in other words, bear not only on the choice of our theoretical models, but also on the tests that we use to evaluate their empirical adequacy.

We hope that, at this stage of the argument, the sense in which a more robust relationship is preferable to a less robust one should be apparent. Both senses of robustness examined so far mirror the way in which econometricians talk of (and search for) 'structural' (or 'autonomous', or 'invariant', or 'superexogenous') relationships in empirical contexts: relations that can be used for intervention across a wide range of situations.36 To establish robustness in a model is therefore supposed to be a preliminary step to (but sometimes

---

34 From a letter to Roy Harrod of 4 July 1938, here quoted from Hausman (1984, pp. 300-1).

35 To such an extent, perhaps, that the latter becomes pretty redundant, at least with reference to "boundary conditions": if a particular assumption can be substituted with another without changing the relevant conclusions, there is no need to place upon it any particular importance from the point of view of its explanatory role. We are grateful to Roger Backhouse for such a remark.

36 The distinction between robustness to changes in the value of the variables appearing in the core causal mechanism (given a fixed set of boundary conditions), and robustness across changes in other conditions parallels the distinction between 'exogenous' and 'super-exogenous' causes drawn by some econometricians (cf. Engle, Hendry and Richard 1983; Hoover 2001a), as well as the philosophical distinction between simple 'causes' and 'capacities' or 'tendencies' (cf. Cartwright 1989, Hausman 1992, Humphreys 1989).

even a substitute for) establish a structural relation in the real world (that is, a relation between real-world variables which is robust to changes in their real-world values, in boundary conditions and perhaps also across different material arrangements).

Finally, we have argued that hypothetico-deductivism is for universal laws, not for robust models. Again, this is not a particularly original remark: the shift in focus from covering-law to causal explanation in the recent philosophical literature has gone hand in hand with the search for new, non-hypothetico-deductive, empirical tests. Most of these tests, as a matter of fact, stress the importance of invariance to intervention, changes in variables, parameters, and background conditions as hallmark of causation.[37] In this paper, by extending these notions to the appraisal of theoretical models we hope to have identified a set of methodological norms that economists, in a sense, have known all along.

37 Cf. Cartwright (1989), Humphreys (1989), Pearl (2000), Woodward (2000), Hoover (2001a).

References

Akerlof, G. (1970) "The market for 'lemons': Quality uncertainty and the market mechanism", *Quarterly Journal of Economics* 84: 488-500.

Allen, L. and Jones, D.G.C. (1967) Principles of Gas Lasers, New York: Plenum Press.

Backhouse, Roger E. (1995) Interpreting Macroeconomics. Explorations in the History of Macroeconomic Thought, London and New York: Routledge.

Backhouse, Roger E. (1997) Truth and Progress in Economic Knowledge, Cheltenham: E. Elgar.

Backhouse, Roger E. and Salanti, Andrea (Eds) (1999) Macroeconomics and the Real World. Vol. I: Econometric Techniques and Macroeconomics, Oxford: Oxford University Press.

Barone, E. (1908/1935): "Il ministro della produzione nello stato collectivista", Gionale degli economisti e rivista di statistica, Vol.37, 1908, pp.267-293. (English translation: "The Ministry of Production in the Collectivist State", in F.A. von Hayek, ed., Collectivist Economic Planning, London: G. Routlegde & Sons, 1935, pp.245-290).

Benassy, Jan-Pascal (1986) "On the existence of Bertrand-Edgeworth equlibria with differentiated commodities", in Werner Hildenbrand and Andreu Mas-Colell (eds) Contributions to Mathematical Economics. Essays in Honour of Gérard Debreu, Amsterdam: North-Holland, 57-57 [Reprinted in Gabszewicz and Thisse (1999), 117-138]

Boland, Lawrence (1989) The Methodology of Economic Model Building: Methodology after Samuelson, London: Routledge.

Boumans, M. and Morgan, M.S. (2001) "Ceteris Paribus Conditions: Materiality and the Application of Economic Theories", Journal of Economic Methodology 8, pp. 11-26.

Cartwright, Nancy (1989) Nature's Capacities and Their Measurement, Oxford: Oxford University Press.

Cartwright, Nancy (1991) "Replicability, reproducibility, and robustness: Comments on Harry Collins", History of Political Economy, 23: 143-155.

Collins, H.M. (1985) Changing Order, Beverly Hills: Sage.

Engle, Robert F., David Hendry, and Jean-Francois Richard (1983) "Exogeneity", Econometrica, 51: 283-308.

Friedman, M. (1953) "The Methodology of Positive Economics", in Essays in Positive Economics, Chicago, University of Chicago Press.

Gabszewicz, Jean J. and Thisse, Jacques-François (eds) (1999) Microeconomic Theories of Imperfect Competition, Cheltenham: E. Elgar.

Gibbard, Alan and Varian, Hal (1978) "Economic models", Journal of Philosophy, 75(11):664-677 [Reprinted in Bruce Caldwell (ed.) The Philosophy and Methodology of Economics, Cheltenham: E. Elgar, 1993. Vol. III: 401- 414].

Granger, Clive W.J. (1999) Empirical Modeling in Economics, Cambridge: Cambridge University Press.

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

Guala, Francesco (2001) "Building economic machines: the FCC auctions", Studies in History and Philosophy of Science 32, pp. 453-477.

Guala, Francesco (2002) "Models, simulations and experiments", in L. Magnani and N.J. Nersessian (eds.) Model-Based Reasoning: Science, Technology, Values, New York, Kluwer.

Guala, Francesco and Salanti, Andrea (2001) “Theory, experiments, and explanation in economics”, Revue internationale de philosophie, 217: 327-49.
Hamminga, Bert and de Marchi, Neil (eds) (1994) Idealization VI: Idealization in Economics (Poznan' Studies in the Philosphy of Sciences and Humanities, 28), Amsterdam and Atlanta: Rodopi.
Hausman, Daniel M. (ed.) (1984) The Philosophy of Economics. An Anthology, Cambridge and New York: Cambridge University Press.
Hausman, Daniel M. (1990) “Supply and demand explanations and their ceteris paribus clauses”, Review of Political Economy, 2:168-87.
Hausman, Daniel M. (1992) The Inexact and Separate Science of Economics, Cambridge and New York: Cambridge University Press.
Hausman, Daniel M. (1998a) Causal Asymmetries, Cambridge: Cambridge University Press.
Hausman, Daniel M. (1998b) “Problems with Realism in Economics”, Economics and Philosophy 14, pp. 185-213.
Hempel, C.G. (1966) Philosophy of Natural Science, Prentice Hall.
Hicks, John (1979) Causality in Economics, Oxford: Basil Blackwell.
Hoover, Kevin D. (2001a) Causality in Macroeconomics, Cambridge: Cambridge University Press.
Hoover, Kevin D. (2001b) The Methodology of Empirical Macroeconomics, Cambridge: Cambridge University Press.
Humphreys, P. (1989) The Chances of Explanation, Princeton: Princeton University Press.
Kreps, David M. (1990) A Course in Microeconomic Theory, London: Harvester Wheatsheaf.
Kreps, David M. and Scheinkman, José A. (1983) “Quantity precommitment and Bertrand competition yield Cournot outcomes”, Bell Journal of Economics, 14(2): 326-337 [Reprinted in Gabszewicz and Thisse (1999), 105-116]
McAfee, R.P. and J. McMillan (1987) ‘Auctions and bidding’, Journal of Economic Literature 25, 699-738.
Machlup, Fritz (1955) “The problem of verification in economics”, Southern Economic Journal, 22(1): 1-21.
Mäki, Uskali (1996) “Scientific Realism and Some Peculiarities of Economics”, Boston Studies in the Philosophy of Science 69, pp. 424-465.
Mäki, Uskali (2000) “Kinds of Assumptions and Their Truth: Shaking an Untwisted F-Twist”, Kyklos, 53(3): 317-36.
Milgrom, P. (1989) ‘Auctions and bidding: A primer’, Journal of Economic Perspectives 3, 3-22.
Morgan, Mary (2001) “Models, stories and the economic world”, Journal of Economic Methodology, 8: 361-384.
Morgan, Mary (2002) “Experiments without material intervention: model experiments, virtual experiments and virtually experiments”, in H. Radder (ed.) Scientific Experimentation and Its Philosophical Significance, Pittsburgh, University of Pittsburgh Press.
Pearl, Judea (2000) Causality, Cambridge, Cambridge University Press.
Plott, Charles (1995) “Rational Individual Behaviour in Markets and Social Choice Processes: The Discovered Preference Hypothesis”, in K.J. Arrow, E. Colombatto, M. Perlman and C. Schmidt (eds.) The Rational Foundations of Economic Behaviour, London, Macmillan.

Popper, Karl R. (1934) Logik der Forschung, Wien: Springer (Engl. trans.: The Logic of Scientific Discovery, 1959, London: Hutchinson)
Popper, Karl R. (1957) “The Aim of Science”, Ratio 1, pp. 24-35, reprinted in Objective Knowledge, Oxford, Clarendon Press, 1972.
Popper, Karl R. (1963), Conjectures and Refutations. The Growth of Scientific Knowledge, London: Routledge & Kegan Paul.
Salanti, Andrea (1998) “Falsificationism”, in John B. Davis, D. Wade Hands and Uskali Mäki (eds) The Handbook of Economic Methodology, Cheltenham: E. Elgar, 183-189.
Schumpeter, J.A. (1992) [1942] Capitalism, Socialism and Democracy, London and New York: Routledge.
Simon, Herbert (1969) The Sciences of the Artificial, Boston, MIT Press.
Sugden, Robert (2000) “Credible words: The status of theoretical models in economics”, Journal of Economic Methodology, 7(1): 1-31.
Titmuss, R.M. (1970) The Gift Relationship: From Human Blood to Social Policy, London: Allen and Unwin.
Vercelli, Alessandro and Dimitri, Nicola (Eds) (1992) Macroeconomics. A Survey of Research Strategies, Oxford: Oxford University Press.
Vickrey, W. (1961) “Counterspeculation, auctions, and competitive sealed tenders”, Journal of Finance, 16: 8-37.
Woodward, J. (1997) “Explanation, invariance and intervention”, Philosophy of Science, 64: S26-S41.
Woodward, J. (2000) “Explanation and invariance in the special sciences”, British Journal for the Philosophy of Science, 51: 197-254.