THE ECONOMICS OF SCIENTIFIC KNOWLEDGE

Jesús P. Zamora Bonilla

Universidad Nacional de Educación a Distancia

Facultad de Filosofía. Madrid. Spain.

E-mail: jpzb@fsof.uned.es

Forthcoming in *Handbook of the Philosophy of Science. ThePhilosophy of Economics*. Ed. by Uskali Mäki. Elsevier.

ABSTRACT: The production of scientific knowledge is a social process that can be analysed with the help of conceptual tools drawn from economic science. A growing literature shows that this analysis is fruitful for the understanding of relevant cognitive aspects of the process of science and its outcomes, and can offer a middle ground between rationalist and constructivist paradigms. Three apoproaches are identified within the economics of scientific knowledge, based respectively in the following ideas: scientific research as an optimisation process, scientific research as a market for ideas, and scientific research as a collection of social mechanisms.

KEYWORDS: Economics of science; sociology of science; sociology of scientific knowledge; epistemic utility; verisimilitude; market for ideas; reflexivity; consensus; scientific norms; scientific institutions.

CONTENTS

1. The idea of an economics of scientific knowledge.

2. The optimisation paradigm.

2.1. Cost-benefit approaches.

2.2. Epistemic utility approaches.

3. The exchange paradigm.

3.1. The market metaphor.

3.2. Polany's 'Republic of Science'.

3.3. Science as a market.

3.4. The limits of the market metaphor.

4. Social mechanisms for knowledge production.

4.1. Mathematical models.

4.2. Institutionalist theories.

1. THE IDEA OF AN ECONOMICS OF SCIENTIFIC KNOWLEDGE.

The economics of scientific knowledge (ESK) is one of the youngest members in the heterogeneous field of 'Science Studies'. Being itself an example of the 'crossing of boundaries' movement that characterises a big part of recent academic activity, it is very difficult, if not impossible, to provide a comprehensive definition of ESK. However, for practical purposes we need in this survey some criteria which help to keep its content under reasonable limits, both in terms of extension and of coherence. So, one *prima facie* plausible definition of ESK, as including any piece of research having to do with 'the economic study of the production and diffusion of scientific knowledge', would force us to include in this paper such an enormous body of literature that at least a full book would be necessary to revise it.[1] On the other hand, the fact that this survey is part of a book on the philosophy of economics, belonging itself into a bigger *Handbook of Philosophy of Science*, suggests that we may select, from this immense literature, just those works dealing with questions more or less related to the traditional topics in the *philosophical* study of science, i.e., mainly topics of epistemological or methodological character. Hence, my working definition of ESK will be *the application of concepts and methods of economic analysis to the study of the epistemic nature and value of scientific knowledge*.

A little bit of history will be pertinent to better understand the implications of this definition. The expression 'economics of scientific knowledge' was first popularised by Wade Hands in a series of papers dating from the beginning of the nineties (Hands (1994a,b)), drawing on an analogy with what defenders

[1] See the two volumes in Stephan and Audretsch (2000), for a good selection of papers on the economics of science falling under this comprehensive definition. Mirowski and Sent (2002a) join also a number of important papers on the economics of science, as well as on ESK.

of the so called 'Strong Programme in the Sociology of Science' had done a couple of decades ago, i.e., to defy the traditional division of labour between sociologists of science and philosophers of science (see Bloor (1976)). According to that tradition, *philosophy of science* would study the *cognitive* aspects of scientific research ('methodology') and of science's epistemic outputs ('epistemology'), whereas *sociology of science* should be devoted to analyse the working of science as a social *institution*, and its relations with other institutions, without entering into the question of what leads researchers to accept a particular method, datum, or theory as 'right'. Without much danger of confusion, we may add to the core of that tradition the thesis that the *economics of science* should be concerned just with the 'economic' problems of scientific research, i.e., how to fund it, or how is it related to economic growth through the mediation of technological progress. Little interference would exist between these three academic disciplines (philosophy-, sociology-, and economics-of-science), for no one of them put questions for which the other two might conceivably provide a relevant answer. On the contrary, the 'new' sociologists of scientific knowledge of the seventies, inspired by the work of Thomas Kuhn and of the 'second' Wittgenstein, amongst others, endorsed the view that the construction of scientific knowledge (i.e., the constitution of a consensus, or a dissensus, about any scientific item) is essentially a *social* process, in which all the agents take one decision or another on the ground of their particular *interests*. From this fact the conclusion was drawn that the creation of scientific knowledge was as legitimate a topic for social analysis as any other process of social interaction. Hence, whereas 'sociology of science' could be taken as the study of the institutional aspects of scientific activity, 'sociology of scientific knowledge' (SSK) would have amongst its legitimate objects of study those questions traditionally reserved for methodology and epistemology.

Wade Hands' suggestion was, basically, that the same argument could be applied not only to sociology, but to economics as well:

> If we mirror the distinction between the sociology of science and the sociology of scientific knowledge, then the economics of *science* would be the application of economic theory, or ideas found in economic theory, to explaining the behaviour of scientists and/or the intellectual output of the scientific community. That is, given the goals of the individual scientists or those of the scientific community (for example, the 'pursuit of truth') the economics of science might be used to explain the behaviour of those in the scientific community or to make recommendations about how those goals might be achieved in a more efficient manner. In this way the economics of science would relate to science in precisely the way that microeconomics has typically related to the firms in the market economy (...) On the other hand, the economics *of scientific knowledge* (ESK) would involve economics in a philosophically more fundamental way. The ESK would involve economics, or at least metaphors derived from economics, in the actual characterization of scientific knowledge - that is, economics would be involved fundamentally in the epistemological discourse regarding the nature of scientific knowledge. Like the SSK argues that scientific knowledge comes to be constructed out of a social process, the ESK would argue that scientific knowledge comes to be constructed out of an economic process.
>
> (Hands, 1994a, p. 87)

The main idea behind this characterisation of ESK is that it essentially deals with the 'nature' of scientific knowledge, and that the construction of this knowledge is an 'economic process', but one may suspect that many of the topics attributed to what Hands calls here 'the economics of science' (i.e., the explanation of scientists' behaviour, or the 'recommendations' about how the epistemic goals of science can be more efficiently attained) would *exactly* refer to the questions he allotted to ESK. Perhaps in order to avoid such a confusion, in a more recent work Hands draws the distinction between the economics of science and ESK in a slightly different way:

> Economics of science analyzes (explains and/or predicts) the behavior of scientists in the same way that an economist might analyze (explain and/or predict) the behavior of firms or consumers. Like the Mertonian school of sociology, the economics of science almost always *presumes* that science produces products of high cognitive quality, but investigating whether it "really" does so is not considered to be the proper subject for economic analysis (it would be like an economist investigating whether the products of a firm "really" satisfy consumer wants). By contrast, ESK, like SSK, would address the question of whether the epistemologically right stuff is being produced in the economy of science; ESK mixes economics and normative science theory. The distinction between the economics of science and ESK mirrors not only the difference between sociology of science and SSK, but also the traditional distinction between microeconomics and welfare economics. Microeconomics, it is usually argued, predicts and/or explains the behavior of economic agents, whereas welfare economics focuses on the question of whether the social configuration produced as a result of the actions of these agents is "optimal" or "efficient" (...) The economics of science predicts and/or explains the behavior of scientists and scientific institutions, whereas ESK adds the question of whether those actions and institutions produce scientific products that are cognitively efficient or optimal (or if they are not optimal, how the institutions might be changed in order to improve epistemic efficiency).
>
> (Hands, 2001, pp. 360-1)

Although I agree that normative questions are paramount in the ESK, I think the identification of ESK with something like an '(epistemically) normative branch of the economics of science' would leave too much outside. Actually, most of the works discussed by Hands in the pages following the second quotation are not only 'normative' but 'explanatory', and, what is more relevant, these works do not consist in the application to normative problems of some (merely descriptive) economic models of science *already* existing; they are, instead, explanatory models *specifically* devised to attack those normative problems. Hence, the production of models that explain the behaviour of scientists is in itself an important part of ESK (or so will it be taken in this survey), at least as long as these models refer to scientists' *epistemic* decisions, i.e., those decisions in which what is at stake is the epistemic value that must be conferred to certain theories, hypotheses, models, research programmes, data, experiments, observations, etc. Hands, however, has wisely pointed to an important difference between the sociological and the economic understanding of social phenomena: in general, economic models are constructed in such a way that they can be given a normative interpretation almost automatically, for, after all, they show transparently the evaluations made by the agents whose interaction constitutes those models' object, and these evaluations are the raw material for any normative assessment. Contrarily to some appearances, economists are proner than sociologists to offer normative guidance, at least by telling who is going to be benefited and who is going to be damaged (and how much) if things happen in such and such a way instead of otherwise; sociologists, instead, often tend to avoid anything that may sound like an evaluative claim, fearing not to look 'impartial' enough (within the fields which are closer to ours, 'science, technology and society studies' and 'social epistemology' would be the main exception; see, e.g., Fuller (1988) and (2000)). In particular, the question of the *efficiency* of scientific institutions and practices arises much more naturally in an economic research than in a sociological one, though obviously this does not entail that the latter can not deliver some important normative conclusions.

Returning to the definition of ESK offered at the beginning ('the application of concepts and methods of economic analysis to the study of the epistemic nature and value of scientific knowledge'), it entails that *ESK will be considered here more as a branch of epistemology than as a branch of economics*: economic concepts and methods are the tools, but scientific knowledge is our object. The main questions to be asked are, hence, of the following kind: how is scientific knowledge 'socially constructed'?, i.e., how does a consensus about an item of knowledge emerge within a scientific discipline?, how do scientists determine the epistemic value of that knowledge?, how can we explain and assess the norms according to which scientists make this valuation?, in particular, how can we evaluate the cognitive efficiency of the methods employed by scientists and the objectivity of their cognitive output?, and so on. Though the works that will be commented below are very heterogeneous in many aspects, they all have in common their trying to answer some of these questions by assuming that the decisions of scientists can be analysed in more or less the same way as economic models conceptualise the decisions of entrepreneurs, consumers, and other economic agents, i.e., by assuming that scientists are trying to optimise some utility function, given all the physical, psychological, and institutional constraints they face, and given that the other scientists and the other relevant agents are also trying to do the same simultaneously.

Nevertheless, as the readers of this book will have surely noticed, there is an immense diversity of approaches within economics, and this allows for a corresponding variety of brands in the ESK. In order to simplify my exposition, I will divide them according to a couple of distinctions. In the first place, two of the more general conceptual frameworks in economics are what we could call 'the optimisation paradigm' and 'the exchange paradigm' (cf. Buchanan (1996)): according to the former, economics is all about how to maximise some important quantities (utility, profits, social welfare...), while for the latter the essential economic phenomenon is what Adam Smith identified as our 'propensity to truck, barter, and exchange', or what the Austrian economist Ludwig von Mises called 'catallaxy'. In the second place, in some branches of economics it is assumed that the proper way of doing research is by building abstract models which try to capture, usually in a strongly idealised way, the essential aspects of the portion of reality we want to study; in other branches, however, the appropriate method is taken to be that of a discursive analysis, giving more importance to the details of economic practices than to purely formal arguments. Roughly following this classification, I will divide the main body of this survey into three parts. In section 2, I will present those works that try to understand scientific research as a process of epistemic optimisation, or of rational cost-benefit decision making. Section 3 will be devoted, instead, to analyse the idea that science is basically an institution for the exchange of items of knowledge, a 'marketplace for ideas', to use the typical expression. The last section will present those works that have tried to make a more or less systematic analysis of scientific research as a set of *social mechanisms* through which different agents interact, distinguishing between those attempts of producing mathematical models of scientists' epistemic behaviour, on the one hand, and those less formal, institutionalist approaches, on the other hand. Notwithstanding all this, if we take into account that ESK has still a very short history, one can easily draw the conclusion that many of the conceivable approaches to the economic study of the constitution of scientific knowledge are still undeveloped, or have not even been envisaged at all. I hope the realisation of this fact may encourage some readers to pursue by themselves a number of these untrodden paths.[2]

2. THE OPTIMISATION PARADIGM.

In a certain sense, we might describe research on scientific methodology (both in its most abstract, philosophical variants, and in the most specific, field-dependent ones) as an attempt to find out what are the 'best' methodological practices, the 'most rational' ones, and so, the whole discipline would fall within the scope of an 'optimisation' approach. A sensible limitation in this survey is, of course, to circumvent it to only those works that have made an explicit use of optimisation concepts derived from economic theory. According to this limited definition, two basic ideas, not essentially in mutual contradiction, are particularly relevant: the idea of optimisation as a rational weighting of costs and benefits, and the idea of optimisation as the maximisation of a utility function. It is clear that the first concept can be reduced to the second one, since costs and benefits can be taken as the main variables of the relevant utility function, though I will use the difference just for expository reasons.

2.1. Cost-benefit approaches.

Interestingly enough, the first known application of modern economic techniques to solving epistemic problems in science was very explicit in describing the value of a scientific theory as the difference between 'costs' and 'benefits'. I'm referring to Charles Sanders Peirce's 'Note of the Theory of the Economy of Research', published in 1879, less than a decade after the introduction of marginal analysis in economic theory by Carl Menger and Stanley Jevons. In that short article, Peirce considers the problem of how much time or effort to devote to several research processes, by taking into account "the relation between the exactitude of a result and the cost of attaining it" (p. 184). The solution of Peirce's model is that total benefit (the difference between the 'utility' derived from a set of experiments and the total cost of performing them, assuming this to be a constant) is maximised by devoting to each problem an amount of resources such that the 'economic urgency' of every problem, i.e., what we would now call the 'marginal utility' derived from each experiment, is exactly the same. Besides being one of the first applications of marginal utility theory, not only to problems of epistemology, but *tout court*, Peirce's 'Note' is also remarkable by connecting the idea of 'epistemic benefits' with some fundamental concepts of statistics and probability theory, and by advancing a

[2] Two obvious research avenues, almost entirely open, would consist in applying social choice theory to the decisions of scientists (following the lines of research on 'judgement aggregation'; e.g., List and Pettit (2002)), as well as economic models of learning (e.g., Brenner (1999)). The approach developed in Goodin and Brennan (2001), in which opinions are taken as a subject for bargaining, could also be interestingly applied to the analysis of scientific consensus.

conception statistical methods close to the modern theory of confidence intervals, specially by taking into account the relevance of costs in statistical inference. Peirce's insight about the type of theoretical and practical problems his approach could open is even more surprising when we read the last paragraph of his 'Note':

> It is to be remarked that the theory here given rests on the supposition that the object of the investigation is the ascertainment of truth. When an investigation is made for the purpose of attaining personal distinction, the economics of the problem are entirely different. But that seems to be well enough understood by those engaged in that sort of investigation.
>
> (Peirce, 1879, p. 190).

Actually, and probably to Peirce's despair, the evolution of ESK in the last decades can be described as a progressive tendency to give more importance to 'the purpose of attaining personal distinction' as an essential factor in explaining the process of knowledge construction.

Peirce's work on the economics of research, however, passed almost unnoticed by as much as a century, till it was rediscovered for the philosophical world by Nicholas Rescher, who devoted a paper to it in the mid seventies (Rescher, 1976), and who has incorporated to his pragmatist vision of science Peirce's idea that economic considerations are essential to understand the rationality of research.[3] Rescher made use of this idea in a series of books, starting by *Scientific Progress*, symptomatically subtitled 'A Philosophical Essay on the Economics of Research in Natural Science'. Although that book was written in the times when the discussion about the possibility of epistemic progress was most eager, it only contains a few references to that debate (e.g., Rescher (1978a), pp. 189 ff., where he suggests that the field of practical applications of scientific knowledge can offer a common ground for 'incommensurable paradigms', in the Kuhnian sense). Instead, the book's main concern is with cultural disillusionment about science:

> Disappointment is abroad that the great promise of the first half of the century is not being kept, and there are many indications that scientists are beginning to move back towards a view of the *fin de siècle* type and to envisage an end to scientific progress. The general public is perhaps even more drastically disillusioned (...) Science has produced numerous and rapid changes in the condition of our lives, but many of them do not seem for the better. Disillusionment all too readily leads to disaffection. A great deal of sentiment is abroad that is anti-scientific and even irrationalistic in orientation (...) A realistic and dispassionate reappraisal of the future prospects of science is thus very much in order.
>
> (Rescher, 1978a, pp. 52-3).

It is in order to introduce 'realisticness' in the assessment of science that Rescher introduces the notion of costs as an essential part of his theory. This he does it by two different means. In the first place, in *Scientific Progress* he tries to show that scientific research is subject to a 'law of diminishing returns' because of the

[3] Joseph Sneed (1989), using the structuralist notion of a scientific theory (see Balzer, Moulines and Sneed (1987)) also provides a model analogous to Peirce's to determine the optimum research agenda in the development of a 'theory net'. In this model, the costs are resources, the benefits are specific solved problems (the 'elements' of the 'theory net'), and it is taken into account that solving a problem may modify the probability of solving further problems.

increasing cost of our 'knowledge-yielding interactions with nature' (*ibid.*, p. 16); Rescher makes a phenomenological description of the history of science, which empirically justifies this claim, and then proceeds to offer an 'economic' explanation of the phenomenon: important results grow in a diminishing way as the resources devoted to science increase, basically because of two reasons: first, there is a law according to which "when the total volume of findings of (at least) *routine* quality stands at Q, the volume of (at least) *important* findings stands at (...) $Q^{1/2}$" (*ibid.*, p. 98), or, in a more general way, the number of findings of merit m decreases exponentially as m grows; second, the increasing technical difficulties in the creation and implementation of scientific instruments and experiments makes it grow the economic cost of each *unit* of new relevant scientific information. This diagnosis allows Rescher to conclude that, on the one hand, science is going to face economic and technical limits much earlier than the limits derived from the finiteness of human ingenuity, but, on the other hand, there is really no end in the prospects of the quantity, quality and variety of forthcoming discoveries (see also Rescher (1996)).

Most important for the evolution of ESK is the second type of use Rescher makes of the Peircean insights, first in his book *Peirce's Philosophy of Science* (1978b), and more decisively in *Cognitive Economy. The Economic Dimension of the Theory of Knowledge* (1989). In the latter book, Rescher generalises the identification of rationality with cost-benefit calculation to other aspects of scientific research besides that of the selection of problems (what, as Hands (1994a, p. 87) rightly points, would belong more to the 'economics of science' than to ESK, as we defined them in the first section). In particular, Rescher employs the cost-benefit idea to argue for the following points:

a) scepticism is wrong in only considering the cost of mistakenly accepting wrong hypotheses, but does not take into account the cost or rejecting right ones (i.e., the benefits we would have had by accepting them); Cartesian rationalism commits just the opposite error; an intermediate attitude towards epistemic risks is better than both extremes;

b) scientific communication is also organised according to the principle of minimising costs, in this case, the cost of searching for relevant information;

c) rational theory choice is grounded on the principle of induction, which is given an economic interpretation: we tend to infer from our data those theories which have more simplicity, uniformity, regularity, coherence, and so on, because these virtues are cognitive labour saving; only when the most economic theory does not give us enough benefits, is it rational to look for a not so simple hypothesis;

d) from a falsifications point of view, the same applies to the decisions about what theories are going to be proposed and tested in the first place.

Other author that has tried to apply 'cost-benefit rationality' to problems in philosophy of science, and in particular to its falsificationist brand, has been Gerard Radnitzky (Radnitzky (1986, 1987)). The problem Radnitzky initially considers is the following: according to Popper, scientists' decisions of accepting or rejecting a theory or a 'basic statement' (i.e., an empirical proposition employed to test a more general theory) are not reducible to any kind of algorithm (like deductive or inductive logic, say), but always rest on conventional decisions; Popper, however, did not explain what precise criteria scientists can or must employ in taking those decisions, in particular, what kind of arguments can be used to decide when to *stop testing* a theory or a basic statement. Radnitzky's idea is to apply Popper's 'situational logic' to that problem, reconstructing the scientist's situation as one in which account must be taken of the 'costs' and 'benefits' of every possible decision: for example, rejecting a particular basic statement can demand offering an alternative explanation of how scientific instruments work, and the latter can be 'more expensive' than the former because it not only demands to invent that explanation, but also to lose some of the *other* things which had been explained with the old theory.

Regarding the merits and demerits of Rescher's and Radnitzky's approaches, one general criticism of their 'cost-benefit' metaphor is that they hardly offer any hint about how costs and benefits can be made commensurable, and this leaves all scientific decisions basically underdetermined. In particular, I think Radnitzky's approach is still less fruitful, mainly because of the following reasons:

a) In the first place, being a true Popperian, Radnitzky attempts to separate his theory as much as he can from the 'sociologist' approaches he attempts to criticise (i.e., those according to which scientists are really moved by 'social', 'non-epistemic' interests, which usually make them more or less blind to rational arguments); this forces him to present his approach as a description of 'ideal type' researchers, motivated only by 'scientific progress', and so he does not help us to know whether *real* scientists behave according to that ideal, or, still worse, whether real human beings *can* behave that way. Rescher attaches more importance to the actual cognitive capacities of people, though without entering into many details.

b) In the second place, describing the goal of science just as 'scientific progress', as Radnitzky does, is little more than a tautology, for 'progress' means 'approximation to some goals', and this is empty until we know what the goals are. We need either a *philosophical* explanation of these goals ('certainty', 'predictive success', or 'problem solving'; but, by which criteria do we decide what problems are 'important', and what an 'interesting' solution consists in?), or we simply identify the aims of science with the goals of *real* scientists (but this seems precluded by point a). Obviously, Radnitzky's own option is the first one, and he takes the goal of science to be 'interesting truth' (but why not 'interesting' from the point of view of the 'social interests' of scientists?); unfortunately, in spite of Popper's efforts, we lack a convincing logical argument showing that falsificationism is the optimal methodology to follow in our attempt to approach to the truth. Rescher's view, on the contrary, seems to be closer to the idea that the relevant values are those of real scientists, but he does not offer a model of how these different aims are mutually interconnected or integrated.

c) In the third place, Radnitzky's use of the terms 'cost' and 'benefit' is too naive from an economic point of view. This is basically due to the fact that, having deprived himself of an operative description of the goals of science (e.g., a coherent theory about how do some methodological decisions *exactly* help in approaching the goals of science), it is always indeterminate *why* is something a 'cost' or a 'benefit', and, in a similar way, it is always unclear how can we decide that some costs are 'higher' or 'lower' than some benefits, since we lack something like a *measure* of them.

d) Lastly, we can criticise the very choice of cost-benefit analysis as an appropriate tool to understand the process of scientific research. In Radnitzky's case, this choice is probably due to his desire of going directly to an explanation of *ideal* scientists' rational decision making, without assuming that something like induction can really exist, and without giving relevance to the 'social' aspects of scientific research. An analytical tool more coherent with standard economic practice is *rational choice theory* (i.e., the hypotheses of expected utility maximisation), which Rescher does not hesitate in using, but this option requires a detailed description of the agent's goals (her 'utility function', even if it contained only epistemic factors), as well as a hypothesis about the agent's beliefs, expressed in terms of *quantitative probabilities*, and this is inconsistent with the Popperian rejection of induction. A still more appropriate analytical tool is *game theory*, which adds to rational choice considerations the idea that the results of an individual's behaviour also depend on the actions of others; but Radnitzky seems to be looking for a theory which can give methodological advises that could in principle be followed by a completely isolated researcher (a 'Robinson Crusoe', so to say), and ignores everything that could take into account the essentially *collective* nature of the scientific enterprise.

2.2. Epistemic utility approaches.

The second route that has been followed within the optimisation approach has consisted into trying to define a specific ('cognitive', or 'epistemic') utility function which rational scientific research should maximise. This has been the strategy of what is usually called *cognitive decision theory*, which is basically an adaptation of the Bayesian theory of rational choice to the case when the decisions to be made are those of accepting some propositions or hypotheses instead of others.[4] Hence, in the case of scientific research, it is assumed that scientists decide (or should decide, if we give this approach a normative interpretation) to accept a particular solution to a scientific problem, instead of an alternative solution, if and only if the expected utility they derive from accepting the former is higher than the expected utility they would attain from accepting any other solution to that problem. The expected utility of accepting the hypothesis h given the 'evidence' e is defined as:

$$(1) \qquad EU(h,e) = \Sigma_{s \in X}\, u(h,s)p(s,e)$$

where the s's are the possible states of the world, $u(h,s)$ is the *epistemic* utility of accepting h if the true state of the world is s, and $p(s,e)$ is the probability of s being the true state given the evidence e. One fundamental problem for a cognitive utility theory is, of course, that of defining an 'appropriate' epistemic utility function u; but, before discussing this problem, there is a still more basic conceptual difficulty that has to be mentioned: standard decision theory is a theory about what *actions* an agent will perform, given her options, her preferences, and the knowledge, beliefs, or information she has about how the relevant things are. It may sound even absurd to say that one can choose what to know, or what to believe. Of course, one can do things in order to gain more or less information, and one can as well allocate more effort to look for information about some topics than about others, but, once the results of this search are in front of you, you usually do not 'choose' what to believe: you just happen to have certain beliefs. Indeed, the fact that a person's beliefs have been 'chosen' by her is frequently a very strong reason to doubt of their truth, or at least, to doubt of the epistemic rationality of that person. Cognitive decision theorists counterargue that the object of an epistemic utility function is not really an agent's system of beliefs: these are represented in (1) by the (subjective) probability function p. The 'acts' whose cognitive utility is relevant are, rather, those of *accepting* or *rejecting* (or suspending judgement on) a given proposition (the hypothesis h). As it has been cogently defended by Patrick Maher (1993, pp. 133 ff.), the acceptance of a scientific hypothesis is logically independent of our belief in its truth: attaching probability 1, or any other 'high' level of probability, to a theory is neither a sufficient nor a necessary condition for its acceptance (for example, most scientific theories are accepted even though scientists actually believe they are not literally true). We may add that, as it will be evident in the next sections, scientists usually have ('social') reasons to accept a hypothesis that have nothing to do with how confident they are about its truth.

Other possible objection is that, even assuming that acceptance and belief are

[4] See Niiniluoto (1987), ch. 12, for an exposition of the first relevant contributions to cognitive decision theory, and Weintraub (1990) for a sceptical argument.

not the same thing, the only relevant thing from the point of view of a sound epistemology is the latter, and not the former; for example, van Fraasen (1980) made precisely this point in discussing the 'inference to the best explanation' approach: once you has concluded that *h* has a higher probability (but less than 1) than any rival theory, accepting *h* would entail to go further than what your evidence allows. This criticism, however, seems to be based on the assumption that accepting a theory is identical with attaching probability 1 to it, what is not the case, as Maher has argued. Nevertheless, the idea that once you have subjective probabilities you don't need acceptance may still have a point, particularly for Bayesian epistemologists. Maher's answer is to point to the fact that scientists (and ordinary people as well) do *actually* accept and reject theories and other types of propositions (an empirical phenomenon that calls for some explanation), and even more importantly:

> Much of what is recorded in the history of science is categorical assertions by scientists of one or another hypothesis, together with reasons adduced in support of those hypotheses and against competing hypotheses. It is much less common for history to record scientists' probabilities. Thus philosophers of science without a theory of acceptance lack the theoretical resources to discuss the rationality (or irrationality) or most of the judgements recorded in the history of science (...) Without a theory of acceptance, it is also impossible to infer anything about scientists' subjective probabilities from their categorical assertions. Thus for a philosophy of science without a theory of acceptance, the subjective probabilities of most scientists must be largely inscrutable. This severely restricts the degree to which Bayesian confirmation theory can be shown to agree with pretheoretically correct judgements of confirmation that scientists have made.
>
> (Maher, 1993, pp. 162f.)

Once we have seen some of the reasons to take acceptance as an act scientists can perform, we can turn to the question of what is the utility function they are assumed to be maximising when they decide to accept some propositions instead of others. Cognitive decision theory is grounded on the idea that this utility function is of an epistemic nature, i.e., the utility of accepting *h* only depends on the 'epistemic virtues' *h* may have. Or, as the first author in using the epistemic utility concept stated:

> the utilities should reflect the value or disvalue which the outcomes have from the point of view of pure scientific research, rather than the practical advantages or disadvantages that might result from the application of an accepted hypotheses, according as the latter is true or false. Let me refer to the kind of utilities thus vaguely characterized as *purely scientific*, or *epistemic*, *utilities*.
>
> (Hempel, 1960, p. 465).

Of course, it was not assumed by Hempel, nor by other cognitive decision theorists, that a *real* scientist's utility function was affected only by epistemic factors; after all, researchers are human beings with preferences over a very wide range of things and events. But most of these authors assume that scientists, *qua* scientists, *should* base their decisions on purely epistemic considerations (and perhaps often do it). So, what are the cognitive virtues an epistemic utility function must contain as its arguments?[5] One obvious answer is 'truth':

[5] Thomas Kuhn's famous discussion about the fundamental values of science (precision, coherence, scope, simplicity, and fecundity), and

coeteris paribus, it is better to accept a theory if it is true, than the same theory if it is false. This does not necessarily entail that accepting a true proposition is always better than accepting a false one (although some authors have defended this, as Levi (1967)), for other qualities, which some false theories may have in a higher degree that some true theories, are also valuable for scientists, as, e.g., the informative content of a proposition (recall Rescher's argument against scepticism). So, one sensible proposal for defining the expected epistemic utility of h is to take it as a weighted average of the probability h has of being true, given the evidence e, and the amount of information h provides. This leads to a measure of expected cognitive utility like the following (Levi (1967), Hilpinen (1968)):

(2) $$EU(h,e) = p(h,e) - qp(h)$$

where the parameter q is a measure of the scientist's attitude towards risk: the lower q is in the epistemic utility function of a researcher, the more risk averse she is, for she will prefer theories with a higher degree of confirmation ($p(h,e)$) to theories with a high degree of content ($1 - p(h)$). If formula (2) reflects the real cognitive preferences of scientists, it entails that, in order to be accepted, a theory must be strongly confirmed by the empirical evidence, but must also be highly informative. Scientific research is a difficult task because, usually, contentful propositions become disconfirmed sooner than later, while it is easy to verify statements that convey little information. One may doubt, however, that these are the only two cognitive requisites of 'good' scientific theories. For example, (2) leads to undesirable conclusions when all the theories scientists must choose among have been empirically falsified (and hence $p(h,e)$ is zero): in this case, the cognitive value of a theory will be proportional to its content, what means that, in order to find a theory better than the already refuted h, you can simply join to it any proposition (it does not matter whether true or false) which does not follow from h. For example, Newtonian mechanics joined with the story of Greek gods would have a higher scientific value than Newtonian mechanics alone.

In order to solve this difficulty, one interesting suggestion has been to introduce as an additional epistemic virtue the notion of closeness to the truth, or verisimilitude (cf. Niiniluoto (1987) and (1998), Maher (1993)), a notion that was introduced in the philosophy of science as a technical concept in Popper (1963): amongst false or falsified theories (and perhaps also amongst true ones) the epistemic value does not only depend on the theories' content, but also on how 'far from the full truth' they are. The main difference between Niiniluoto's and Maher's approaches is that the former is 'objective', in the sense that it assumes that there exists some objective measure of 'distance' or '(di)similarity' between the different possible states of nature, and the value of accepting a theory is then defined as an inverse function of the distance between those states of nature that make the theory true and the state which is actually the true one. Maher's proposal, instead, is 'subjective' in the sense that it starts assuming that there is an undefined epistemic utility function with the form $u(h,s)$, perhaps a different one for each individual scientists, and the verisimilitude of a hypothesis is then introduced as a normalised difference between the utility of accepting h given what the true state is, and the utility of accepting a tautology. In Maher's approach, then, epistemic utility is a primitive notion, which is only assumed to obey a short list of simple axioms: (i) accepting a theory is better when it is true than when it is false, (ii) the utility of accepting a given true theory does not depend on what the true state is, (iii) accepting a true theory is better than accepting any proposition derivable from it, (iv) there is at least a true theory accepting which is better than accepting a tautology,

about how they can be given different weight by different scientists (Kuhn, 1977), can easily be translated into the language of utility theory.

and (v) the utility of accepting a full description of a true state of nature is a constant and higher than the utility of accepting a logical contradiction. Maher assumes that different scientists may have different cognitive utility functions, and hence, they can give different verisimilitude values to the same theories, even if the principles listed above are fulfilled. Actually, Niiniluoto's approach is not completely objective, because the definitions of distance between states depend on what factors of similarity each scientist values more or less. This is not a bad thing: after all, cognitive preferences are *preferences*, and these are always the preferences of some particular agent.

One general criticism that can be made to the proposals examined in this subsection is that the arguments their authors give in favour of one definition of cognitive utility or another, are always grounded on our 'intuitions' about what is better or worse in the epistemic sense. With the exception of Maher, they seldom discuss whether those functions do actually represent the cognitive preferences of flesh-and-bone scientists. The absence of such a discussion also deprives us of an answer to the question of what would happen if the real preferences would not coincide with the ones defended by cognitive decision theorists: should we then criticise scientists for not being 'scientific' enough?, or could we take this disagreement as an argument against the utility functions defended by those philosophers? Furthermore, from the proposed definitions it is frequently impossible to derive any behavioural prediction (besides some generalities like the ones commented in connection to formula (2)) about what decisions will be made in a minimally realistic scenario by a scientist who happend to have such cognitive preferences. A different problem is that there are too many definitions of epistemic utility, and it seems reasonable to ask whether the criteria to *prefer* one definition over the rest are also derivable from some ('higher level'?) epistemic preferences. At the very least, we should demand from a candidate definition that accepting it as an appropriate representation of the 'right' epistemic preferences is an optimum decision according to that very same definition. Regarding to this criticism, I think Maher's approach is more appropriate than the others, for, even if it could seem that by not offering an explicit definition of epistemic utility he had left this notion unexplained, I prefer to interpret his strategy as a 'liberal' one, in the sense that it allows to take *any* function that satisfies the five principles listed in the preceding paragraph as an acceptable epistemic utility. This strategy amounts to denying the philosopher the right to determine what epistemic preferences are 'appropriate', besides indicating some minimal requisites that make these preferences deserve to be called 'cognitive'.

An alternative but related approach has been defended by me in a series of papers about 'methodological truthlikeness' (e.g. Zamora Bonilla (1996, 2000)): instead of searching for a definition of cognitive utility which satisfies some intuitive requirements, or instead of just acknowledging the right of scientists to have the epistemic preferences they may have, the philosopher of science should try to *discover* what these cognitive values are. The suggested strategy is an abductive one: in the first place, we have to look for the methodological patterns scientists actually follow, i.e., their 'revealed' criteria for theory preference (patterns which are not supposed to be universal); in the second place, we must try to find a definition of cognitive utility from which those patterns could be mathematically derived as 'empirical predictions'. Both things can obviously be made with a bigger or lesser level of detail: we can look for very general methodological patterns, which would be taken as 'stylised facts' about theory choice, empirical testing, and so on, or we can alternatively inquiry about detailed methodological decisions in specific case studies; and we can also employ very simple hypothetical utility functions, with just a few arguments within them, or develop more complicated functions. My impression is that, the more we concentrate on the specificities of particular cases, the less likely it is that actual scientific decisions depend only on cognitive factors; the most general methodological patterns, on the other hand, are those defining the kind of activity scientific research is, and the type of output that society can expect to derive from it, and this allows to guess that scientific institutions and practices will have probably evolved in such a way that those patterns are coherent with scientists' more general epistemic preferences. Now that the methodology of case studies has become a kind of orthodoxy in the philosophy and the sociology of science, some may have doubts about the mere possibility of finding in the history of science any regularities regarding the scientists' methodological decisions (for, do not researchers use all methods 'opportunistically'?), and even more about the viability of deriving an explanation of these practices just from a few simplistic formulae (e.g., from the assumption that certain cognitive utility function is maximised through those decisions). To the first objection we can answer that the patterns that have to be searched are not of the type 'scientists always employ method X', but rather of the form 'under circumstances Z, scientists tend to employ method X; under circumstances Z´, scientists tend to employ method X´, and so on'. Regarding the second objection, the proof of the cake is obviously in the eating.

The main definition of epistemic value I have proposed in the papers mentioned above asserts that the verisimilitude of a hypothesis h for a scientist given a set E of empirical regularities or data, is identical to the product of the 'similarity' between h and a subset F of E (measured as $p(h\&F)/p(h\vee F)$, where p stands for the scientist's subjective probabilities) and the 'rigour' of F (measured as $1/p(F)$), for that subset F for which this product is maximum. These definitions allow to derive a wider set of 'methodological theorems' than other existing measures of epistemic utility (cf. Zamora Bonilla (1996)), but also to explain some 'stylised facts' about the development of scientific research programmes or about the differences between the role of theoretical and empirical arguments in economic theory as compared to natural science (cf. Zamora Bonilla (2003) and (1999a), respectively). I recognise that the number of 'facts' so explained is not too big, but what this indicates is not that formal models are useless (after all, other types of explanations of scientific practice haven't got a much better success record), but that they must be enriched to take into account other relevant factors in scientific decision making. This is exactly what the contributions to be examined in the next sections try to do.

3. THE EXCHANGE PARADIGM.

3.1. The market metaphor.

Traditionally, economics is not only about the optimisation of some magnitudes, be they utility, profits, wealth, or social welfare. Beyond the assumption that economic agents are rational beings who always try to make the best possible choice, there is the indisputable fact that the objects of economic research are *social* phenomena, that have to do with the *interrelation* between numerous agents. Classical economists introduced the idea that there is a specific realm ('the economy') the discipline they were creating dealt all about, a complex entity all of whose elements depended on the rest. Later on, more limited social fields (i.e., single markets) attracted the systematic attention of economic theorists, but the fundamental idea was still that these fields were systems composed by interdependent parts. In economics, contrarily to the majority of the other social sciences, the most fundamental type of connection existing between the elements of those systems is assumed to be that of *exchange relationships*, and economic science can then be understood as the study of those social phenomena in which the basic 'bond' consists in (more or less free) exchanges. Of course, the market is the paradigm of such social phenomena, and hence, trying to describe science as an exchange system leads us almost automatically to interpret it like something akin to a 'market'. Due to the common identification of the economic efficiency of markets with another classical metaphor (the 'invisible hand'), the thesis that 'science is like a market' has often been taken as an assumption about the working of some 'epistemic invisible hand' mechanism behind the process of scientific research. This vision of science as a 'marketplace for ideas' was not originally a technical notion in the analysis of scientific research,[6] but rather a common metaphor 'floating in the air', and probably having a basic ideological role: that of justifying the autonomy of scientific opinions from external social pressures. The following quotations, the first one by a Popperian philosopher, and the second one by a prominent economist, are illustrative of that opinion:

> I was taught as a student that the university is a marketplace of ideas where new ideas are welcome and falsehoods can be challenged without recrimination.
>
> Bartley (1990, p. xvi)

> I do not believe that (the) distinction between the market for goods and the market for ideas is valid. There is no fundamental difference between these two markets.
>
> Coase (1974, p. 389)

Nevertheless, it is one thing to express the metaphor that 'science is a market', and it is a very different thing to try to use it as an analogy to illuminate in a detailed way the essential features of scientific research (by the way, the analogy can also be used in the opposite direction, to understand the market as a knowledge generating mechanism; cf. Hayek (1948)). Once we begin to employ the metaphor as an analytical device, an obvious problem is to make it explicit what we are really understanding by the market concept, for if we take buyings and sellings as the fundamental bricks of the market, it seems clear that 'scientific ideas' or 'opinions' are not *really* bought nor sold by researchers, save in a very metaphorical sense (perhaps with the main exception of technological research; this difference between ideas researchers decide to made public -and then not sellable- and ideas kept secret -and then commodifiable- has even been proposed as the basic distinction between 'science' and 'technology', cf. Dasgupta and David (1994)). So, in order to understand ('pure') science as a market, we need to concentrate on some *abstract* features of markets, some that 'real' markets may reflect in some definite way, but that other social institutions (which can be taken as markets just by means of an analogy) materialise in different ways. As we could expect, proponents and critics of the idea of 'science as a market' have tended to concentrate on different aspects of the analogy. On the other hand, the market concept (and particularly the concept of a 'free' market) is by no means a *neutral* idea normatively speaking; the debate between pro-market and anti-market activists is longstanding, and it also reflects in the analysis of science: some authors argue that 'science is a market' and that this is a good thing both from a social and from a cognitive point of view, whereas other authors have employed the same thesis to justify scepticism about the epistemic objectivity of scientific knowledge, or to assert that, the more science 'becomes' a market, the less advantageous it is for the

[6] Though it underlied several sociological approaches like those of R. K. Merton, W. O. Hagstrom, or J. Cole, cf. the "competition model" in Callon (1995), as well as Ziman (1968) and (2002).

common citizen (cf. Fuller (2000), Mirowski and Sent (2002b); for a more positive view, see Goldman and Cox (1996), where the idea of a 'free market for ideas' is applied to all public communication, and not only to science). In the remaining of this section, I will present in the first place Polany's pioneering understanding of science as a self-organising system; in subsection 3.3 I will discuss some of the proposals to explicitly analyse science as a kind of market, and lastly I will present some of the criticisms that these proposals have received.

3.2. Polany's 'Republic of Science'.

The first serious attempt to analyse the working of science by means of the market analogy was Michael Polany's classic article 'The Republic of Science. Its Political and Economic Theory' (1962), although in that paper he explicitly avoided to assimilate science with a market, but tried instead to show that both institutions are examples of self-co-ordination processes:

> (The) highest possible co-ordination of individual scientific efforts by a process of self-co-ordination may recall the self-co-ordination achieved by producers and consumers operating in a market. It was, indeed, with this in mind that I spoke of "the invisible hand" guiding the co-ordination of independent initiatives to a maximum advancement of science, just as Adam Smith invoked "the invisible hand" to describe the achievement of greatest joint material satisfaction when independent producers and consumers are guided by the prices of goods in a market. I am suggesting, in fact, that the co-ordinating functions of the market are but a special case of co-ordination by mutual adjustment. In the case of science, adjustment takes place by taking note of the published results of other scientists; while in the case of the market, mutual adjustment is mediated by a system of prices broadcasting current exchange relations, which make supply meet demand (...)
>
> Polany (1962, pp. 467-8).

With the help of this analogy, based on an Austrian conception of the economy as a self-regulating system, Polany's analysis proceeds by indicating the ways in which the choices of an individual scientist are constrained by the professional standards of her discipline, and how these standards emerge as a solution to the problems of co-ordination that are faced in scientific research. On the one hand, individual scientists try to attain the maximum possible 'merit' with their stock of intellectual and material resources. On the other hand, the scientific merit attached to a result depend on a number of factors, the most important ones being the result's plausibility, its accuracy, its relevance, and its originality. The three first criteria tend to enforce conformity, whereas originality encourages dissent, and this tension is essential both in guiding the decisions of individual researchers, and in explaining the tremendous cognitive success of science. Actually, Polany's claim seems to be that these are exactly the criteria employed in science *because* they have proved to be efficient in the production of knowledge: the 'invisible hand' argument refers not only to the attainment of efficient *results* in the decisions of individual researchers (who maximally exploit the gains from epistemic trade thanks to competition), but to the establishing of the most appropriate *rules* within the scientific community. Unfortunately, no detailed empirical analysis to justify these conclusions are offered in the article. In particular (and this is a problem of most of the contributions that will be surveyed in this section), Polany does not even recognise the possibility that norms which are efficient in the pursuit of knowledge may be not so efficient in the pursuit of merit, and viceversa, and it is not clear what type of efficiency has more weight in guiding the evolution of scientific standards.

Other quasi economic argument offered by Polany refers to what he calls "the uniformity of scientific standards throughout science", something which allows the commensuration of the values of very different discoveries in completely disparate parts of science:

> This possibility is of great value for the rational distribution of efforts and material resources throughout the various branches of science. If the minimum merit by which a contribution would be qualified for acceptance by journals were much lower in one branch of science than in another, this would clearly cause too much effort to be spent on the former branch as compared with the latter. Such is in fact the principle which underlies the rational distribution of grants for the pursuit of research. Subsidies should be curtailed in areas where their yields in terms of scientific merit tend to be low, and should be channelled instead to the growing points of science, where increased financial means may be expected to produce a work of higher scientific value (...) So long as each allocation follows the guidance of scientific opinion, by giving preference to the most promising scientists and subjects, the distribution of grants will automatically yield the maximum advantage for the advancement of science as a whole.
>
> (*ibid.*, p. 472)

Again, this is just a transposition to the case of science of the Austrian economics thesis that prices are the instrument for the co-ordination of individual decisions in the market. But, without a systematic analysis of how 'scientific value' is constituted by the interconnected decisions of different individuals, the argument lacks any logical cogency, not to talk about its *prima facie* plausibility, for, as a matter of fact, the quality standards for acceptance of contributions in different disciplines, as well as in different journals within the same discipline, are very far from uniform. A last important economic metaphor in Polany's analysis of 'the Republic of Science', also analogous to a Hayekian view of the market as an epistemic co-ordination mechanism, is his view of 'scientific opinion' as a single collective authority, what is constituted by myriads of single judgements of individual scientists, each one having competence on just a tiny fraction of all scientific knowledge overlapping more or less with the areas of competence of other colleagues: [7]

> Each scientist who is a member of a group of overlapping competences will also be a member of other groups of the same kind, so that the whole of science will be covered by chains and networks of overlapping neighbourhoods (...) Through these overlapping neighbourhoods uniform standards of scientific merit will prevail over the entire range of science (...) This network is the seat of scientific opinion. Scientific opinion is an opinion not held by any single human mid, but one which, split into thousands of fragments, is held by a multitude of individuals, each of whom endorses the others' opinion at

[7] For a comparison of Polany's vision of science with Hayek's general approach to mind and society, see Wible (1998, ch. 8) and Mirowski (2004, chs. 2 and 3), where the differences between both approaches are discussed.

> second hand, by relying on the consensual chains which link him to all the others through a sequence of overlapping neighbourhoods.
> (ibid., p. 471).

I think this view of scientific opinion (which is in some sense similar to Philip Kitcher's notion of 'virtual consensus', which we will see in section 4.1) may lead in a natural way to develop analytical models and empirical studies in which scientific interactions are understood as the elements of a *network*, but this is mostly work that is still to be done.

3.3. Science as a market.

Curiously enough, it was not economists, but sociologists, the first ones in taking over the analogy between science and markets (an analogy that, as we have just seen, Polany explicitly presented as only working at the level of abstract mechanisms), in particular Pierre Bourdieu, in his pathbreaking article 'The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason' (1975). According to Bourdieu, scientific research consists in the competition for scientific authority, which is a kind of monopoly power, "the socially recognised capacity to speak and act legitimately (i.e., in an authorised and authoritative way) in scientific matters" (*op.cit*, p. 31). The distribution of this authority within a scientific discipline at a given moment is what constitutes it as a 'field', in the Bourdiean sense of a structure of interrelations and capacities which determine the interests and strategies of each actor. Authority is seen as a kind of 'social capital', which can be "accumulated, transmitted, and even reconverted into other kinds of capital" (p. 34). In very explicitly economic terms, Bourdieu asserts that these 'investments'

> are organized by reference to - conscious or unconscious- anticipation of the average chances of profit (which are themselves specified in terms of the capital already held). Thus researchers' tendency to concentrate on those problems regarded as the most important ones (e.g., because they have been constituted as such by producers endowed with a high degree of legitimacy) is explained by the fact that a contribution or discovery relating to those questions will tend to yield greater symbolic profit. The intense competition which is then triggered off is likely to brig about a fall in average rates of symbolic profit, and hence the departure of a fraction of researchers towards other objects which are less prestigious but around which the competition is less intense, so that they offer profits of at least as great.
> (Bourdieu, 1975/1999, p. 33)

Perhaps the most characteristic feature of this type of competition as compared to others (entrepreneurial, political, artistic, and so on), is that the scientific field is highly autonomous, in the sense that "a particular producer cannot expect recognition of the value of his products (...) from anyone except other producers, who, being his competitors too, are those least inclined to grant recognition without discussion and scrutiny". Bourdieu argues that this autonomy is what has created the false impression of scientific research being a 'disinterested' activity, but he also declares that the existence of specific social interests pushing the strategies of scientists do not entail that the cognitive products of these strategies lack epistemic objectivity. Rather on the contrary, the specificity of the scientific field consists in the fact that the competition that takes place within it under an "inherent logic" which brings about, "under certain conditions, a systematic diversion of ends whereby the pursuit of private scientific interests (...) continuously operates to the advantage of the progress of science" (p. 39). This "transmutation of the anarchic antagonism of particular interests into a scientific dialectic" (i.e., one which is based on the observance of "scientific method") is effected thanks to the need of each individual scientist to fit his arguments to a set of methodological practices whose most competent performers are precisely his own competitors, a process that usually leads all competitors to a "forced agreement" (p. 41) that rarely occurs outside the natural sciences (save for the violent imposition of a dogma, as it is the case in religions and in totalitarian regimes).

Bourdieu's vision of scientific research as a market for scientific credit was transformed by Bruno Latour and Steve Woolgar (1979) into what we might call a Marxist theory of the scientific market. According to these authors, the essential aspect of the research process is that the 'capital' won by a scientists is always re-invested, generating a 'cycle of credit'. One important implication of this view is that no single element of the cycle is more fundamental than the rest, but this seems to lead Latour and Woolgar to the strange conclusion that the motivation of scientists is not the pursuit of credit, nor of any other of the elements of the cycle (access to scientific facilities, production of reliable information, publication of research results, and so on), but "the acceleration and expansion" of the cycle by itself (op.cit., p. 208), an idea which is difficult to implement in a standard economic analysis. Other important insight in Latour and Woolgar's approach is the relevance they attach to an aspect of the interdependence of researchers which is usually lacking in other sociological theories: the fact that the *value* of the information provided by a scientists depend on the *demand* of that information by other scientists, who need that information in order to produce in their turn further information who can be transformed in credibility, and so on (op. cit., p. 206). For economically oriented readers, Latour and Woolgar's avoidance to discuss an obvious question can be disappointing : to what extent the working of the credibility cycle favours the production of *reliable* information (i.e., information that *is* useful), and not only that of 'credible' information (i.e., information that is *taken to be* useful). Of course, the very same question is precluded by the fact that their chapter on the credibility cycle is a continuation of the one where they famously argue that scientific statements can not be taken as objective representations of an independent reality, for this 'external' reality is 'constructed' in the same process that leads to the collective acceptance of the statement presumably describing it. A possible answer to this unposed question is that of David Hull (1988), who asserts that the basic factor enhancing a scientist's credit is not the recognition of his results by other researchers, but the *use* they make of them, and this provides an incentive to produce results that are epistemically reliable, for, in general, wrong statements will easily lead to failed predictions and actions. Thomas Leonard (2002) also explains how some common scientific rules (particularly peer review and replication) evolved historically as mechanisms guaranteeing that individual scientists have an interest in using and producing reliable ideas.

The most systematic attempt to illuminate the process of science in terms of the market concept has been made Allan Walstad (a physicist), especially in his paper 'Science as a Market Process' (2002). In line with Polany's contribution, Walstad develops an Austrian approach, avoiding to use mathematical models, basically because of the absence of numerical data about the relevant facts, and because of the essential instability and complexity of social interactions. Instead, he presents a list of similarities, as well as differences, between 'traditional' and 'scientific markets'. The more important similarities are the existence in both cases of a high level of *specialisation*, *exchange* ('recognition' in payment for the use of information), acquisition and *investment* of cognitive capital (but, contrarily to Latour and Woolgar, allowing for the possibility of some 'final ends', either cognitive or practical, serving as an explanation of the production cycle), *entrepreneurial activity* (both in the Schumpeterian sense of a disequilibrating force -novelty creating-, and in the Kirznerian sense of an equilibrating factor -e.g., arbitraging-), *institutional organisation* (e.g., research teams or groups, analogous to firms in traditional markets), and *self-regulation* (with evolved mechanisms that discourage inefficient activities, facilitate co-ordination, and avoid market failures; e.g., citation records performing a similar informational role, to prices in traditional markets; see also Butos and Boettke (2002)). The main differences between science and common markets is the absence of money as a means of exchange in the former; this entails that scientists can not charge different prices for citations, nor carry out indirect exchanges, nor transfer to others the right to be cited (as it is the case for patents and other property rights, for example).

3.4. The limits of the market metaphor.

The analysis of the scientific process in terms of market concepts can be criticised in a number of ways. For example, some may argue that the vision of researchers as 'scientific entrepreneurs' distorts the essential aspects of scientists' motivations. This criticism can be raised both by rationalist philosophers who think that scientists basically pursue epistemic goals, and by sociologists who think that scientific ideas are just rhetorical strategies to defend the power of some social classes. Fortunately, the 'entrepreneurial' character of many scientific decisions is backed enough by a huge amount of case studies (e.g., Latour (1987), Pickering (1995)), independently of whether their authors employ an 'economic' approach or not. On the other hand, the market metaphor might be criticised because it puts too much emphasis on *voluntary* exchanges, and not so much in the *institutional* mechanisms of science (e.g., Ylikoski (1995), Wray (2000)). I think this criticism is more relevant, but it surely ignores the complexity of modern economic analysis: as it will be obvious for anyone acquainted to contemporary microeconomics, there is no such a thing as 'the' market concept; what there is, instead, is a varied set of market notions linked by family resemblances, rather than united under a single definition, and this diversity reflects indeed a still wider variety of *types* of markets in the real world. One thing which is clear for almost all economists is that the differences between these types of markets essentially depend on the *institutions*, *norms*, and *habits* those markets are embedded into (although there is deep disagreement about how beneficial this institutional embedding actually is). Just to put a compelling example: any market transaction presupposes some property *rights*, as well as some *legal* mechanism to punish the infringement of those rights; it also presupposes some *procedures*: shirts in a warehouse are not bought in the same way as bonds in the stock market. So, any serious analysis of science 'as a market' must make it clear what are the institutions allowing or constraining 'scientific transactions' (cf. Polany's approach). Actually, the contributions I will examine in section 4.2 are 'institutionalist', not in the sense that they deny that the market metaphor is appropriate, but because they explicitly consider the role of institutions in the 'scientific market'. So, I think the most important criticisms to the idea that 'science is a market' are those coming from inside, i.e., those problems the very application of the idea has allowed to disclose: first, do some serious 'market failures' exist within science?, and second, is the idea of a 'scientific market' when applied to itself logically coherent?[8]

The expression 'market failure' is employed to refer to those circumstances under which the free decisions of sellers and buyers would lead to an inefficient result. Monopolies (absence of competition), externalities (decisions affecting to third parties), public goods (non divisible amongst private consumers), informational asymmetries (one party knowing more about the good than the other), and transaction costs (the ones incurred in carrying out agreements) are typical situations where suboptimal solutions to co-ordination problems may emerge if the agents are left to decide by themselves in the pursuit of their private interest (see, e.g., Salanié (2000)). The problem for market theories of science is that in the case of scientific research all these circumstances are the norm rather than the exception. For example, cognitive monopolies, i.e., the neglection of 'heterodox' theories or methods, not only arise very frequently (cf. Wible (1998), chs. 6 and 7), but, as we saw when reporting Bourdieu's approach, the striving for the monopolisation of epistemic authority is the basic force driving scientific competition; indeed, we can interpret Thomas Kuhn's 'paradigms' as examples of scientific monopolies (cf. Oomes (1997)), and Popper's (1970) complaints about the epistemic inferiority of 'normal science' as an argument demanding more competition in research. Oomes also explain monopolies as caused by 'historical lock-in', when an idea or method becomes accepted just because of having gained a temporal advantage over its competitors.

To a high extent, epistemic competition naturally leads to monopolies because, with the exception of some relativist philosophers, it is assumed that scientific problems have only one 'right' answer, or, at least, that the more correct answers 'displace' the worse ones. Stated in economic terms, this means that knowledge is a 'public good', one that, when 'produced' for a single individual, all the other agents can freely make use of it. The public nature of knowledge was employed by many authors to justify its production through governmental funding (e.g., Nelson (1959)), although more recently it has been put into question, at least as a universal fact about scientific knowledge: Dasgupta and David (1994) explain the publicity or privacity of knowledge as an endogenous variable of the research process, and Callon (1994) argues that it is only the production of heterodox ideas what must be publicly financed. Little analysis has been done, instead, of other questions more relevant from the epistemological point of view, as whether it is scientific knowledge's being *information*, or its being *true*, or its being *collectively accepted*, what makes of it a public good, and what epistemic consequences each one of these options has, i.e., how does exactly the public or private nature of knowledge affect the *cognitive efficiency* of the research process. One plausible avenue for research on these questions would be to apply the ample literature about knowledge and information in non-co-operative games. Lastly, informational asymmetries and transaction costs are other phenomena that clearly manifest in the case of science. I will comment more about the former in the section 4.1. The case of transaction costs has been much less studied, the main exceptions being Mäki (1999), which will be commented below in this section, and Turner (2002), who describes science as a market system of public certification processes, which attempt to reduce the costs the evaluation of information would have for single individuals.

Finally, reflexivity can also pose problems for market theories of science, and, more generally, to the very idea of an economics of science, as it has been stated by several authors. The most obvious difficulty is that an economic analysis of science applied to economics itself *might*

[8] See also McQuade and Butos (2003) for an Austrian explanation of science which concentrates on the *differences* between the market process and the science process.

show that economics is not 'good enough' as a science, perhaps due to the existence of some 'market failures'. For example, Thomas Mayer (1993) argues that economics as a market for ideas fails because it is basically an activity dominated by producers: professional economists are the only agents controlling what must count as 'good' economics, whereas in other branches of science successful technological application is the final arbiter. This, according to Mayer, has lead to the dominance of mathematical over empirical research, and to putting the goal of precision before the goal of truth. A similar argument is presented in Zamora Bonilla (2002). But, if these arguments are right, then the very intellectual tool with which they are produced might be flawed, and economics could be alright after all! This conclusion is paradoxical; in fact, it is an example of the 'Liar's paradox' ("What I'm saying is false"), that philosophers of logic have examined for centuries. Of course, this is not only a problem for an economics of science, but, in general, for any other scientific explanation of science, particularly for those approaches concluding that scientific knowledge is not 'objective' (cf. Hands (1994a), pp. 91 ff., as well as Mäki (2004), pp. 217 ff. for criticisms of the credibility of economics in particular). Wible (1998, chs. 11 and 12) takes this paradox as an opportunity to criticise what he calls "the architecture of economic theory and method", a criticism he hopes would justify the use of an evolutionary conception of rationality instead of classical equilibrium concepts (cf. section 4.2 below). Another possible solution to the paradox would be to argue that the particular branch (or tool within this branch) of science one is employing is not in a bad epistemic state, after all. This is what some sociologists of scientific knowledge did (e.g., Bloor (1976)), by suggesting that sociological knowledge is actually better grounded than the theories of natural science (a position advanced centuries ago by Giambattista Vico). As far as market theories of science are concerned, they have usually tended to conclude, through some kind of invisible hand argument, that science works more or less well in general, and so the paradox would not arise. But, if this were the case, a possible criticism could be made by showing that the specific strand of economics that is being used in developing such a market theory of science is actually *not* a dominant one within the economics profession (i.e., it is not very successful in the market for economic theories). Uskali Mäki (1999) has levied this criticism towards Ronald Coase's defence of a free market of ideas, but it could also be applied to the more recent, and much more developed Austrian approaches we have examined in this section. Furthermore, Mäki (2004, pp. 219-20) argues that the application of market-like models to our understanding of science might have self-destructive effects, by shaping and channelling the most selfish predispositions of scientists towards epistemically inefficient patterns of behaviour, in a similar fashion to the way exposure to neo-classical economic concepts seems to enhance the self-interestedness of economics students. Perhaps science is really a market, but it would work better if we didn't know it. Mäki's suggestion, nevertheless, is in line with that of most critics of market theories of science: we must just be careful in the application of economic concepts, being conscious of their limitations as well as of their diversity and analytical power.

4. SOCIAL MECHANISMS FOR KNOWLEDGE PRODUCTION.

4.1. Mathematical models.

The most distinctive feature of modern economics is probably its reliance on the methodology of mathematical model building. As with any other powerful tool, the danger exists of people being excited by it and using it much more intensely that what would be sensible or necessary. In the case of economics, Mayer (1993) has argued that the difficulty in deciding whether economic theories are *right or wrong*, and hence in ranking economists according to their ability to discover the *truth* about economic facts, has led to take mastery in the invention and manipulation of mathematical models as a paramount criterion of scientific excellence; the mathematisation of economics would have been, hence, more an example of academic pedantry than a real epistemological virtue. Though I do not share such an extreme position, I think it can serve nevertheless to remind us that the final aim of scientific model building is that of illuminating *real* phenomena, and it has not to be carried out for its own sake. On the other hand, mathematical models are basically *logical arguments*, whose main virtue is that they allow us to see very clearly what follows, and also what does not follow, from a definite set of premises. These premises describe an imaginary world, and mathematical analysis just allows to decide in an unambiguous way what would happen in that world under some conceivable circumstances. The most important question is, hence, to what extent that imaginary world represents well enough the relevant aspects of the actual (or counterfactual) way things are, so that our conclusions are transportable to the real world (cf. Sugden (2002)), but this question is decided through a 'dialectical' process, for after all, how empirically accurate our conclusions are is usually (or it should be) the most compelling reason to judge whether our model is appropriate enough (cf. Friedman (1953)).

Mathematical models of scientific knowledge production are not common, however. This is due to a combination of circumstances: *sociologists of science* would have the main motivation to use those models, but most of them may think that it would be an example of 'economics imperialism', and that it tends to hide the qualitative complexity of social processes; most *philosophers* don't know too much about economics for even considering seriously the possibility of engaging into an economics of scientific knowledge, perhaps beyond producing some very general, qualitative arguments; *economists* have been too busy developing complicated mathematical techniques and applying them to proper 'economic' problems, for losing their time investigating such a minor question; and lastly, *philosophers and methodologists of economics* may have been the scholars where the right interests and the right resources were combined in an optimal way, but most of them are either too much critical of standard economic theory for considering worthy the effort, or fear that an economics of scientific knowledge would be dangerously close to social constructivism and relativism (cf. Davis (1998)). As a result, until now only a fistful of authors have tried to develop some formal economic analyses of the social mechanisms of scientific knowledge production.

As far as I know, the first application of an economic model to a problem clearly falling within the philosophy of science was due to Michael Blais (1987), who employed Robert Axelrod's evolutionary 'solution' to the Prisoner's Dilemma game to illustrate how *trust* in the work of scientists may emerge just as a result of self-interested behaviour, without requiring a 'moral' explanation. According to Blais, the interaction between researchers and journal editors can be represented as a Prisoner's Dilemma: the former can either 'cooperate' (i.e., perform good research) or not, and the latter can also 'cooperate' (i.e., publish good results) or not; cooperation is more costly than defection for both types of agents, although both benefit more if everyone cooperates than if everyone defects. As it is well known, Axelrod (1984) showed that, when a game like this is repeatedly played by automata which have been programmed to follow always the same strategy (although not all of them the same one), the most successful strategy amongst a high number of them which had been proposed was the one called 'tit-for-tat' (cooperate the first time, and then cooperate if and only if the other player has cooperated the last time). Using this strategy is, then, in the interest of every agent, what also makes cooperation to become the most frequent decision. Blais argues that trust in the results of scientists by part of journal editors, and in the quality of published papers by part of researchers, may have evolved in a similar way. The mechanism of interaction between scientists and editors is decisive, nevertheless, and it may lead to suboptimal 'solutions' in the Prisoner's Dilemma game, like, for example, letting pseudo-scientific papers become published (cf. Bracanovic (2002); see also Hardwig (1991) for a criticism of Blais' thesis).

One year after the publication of Blais' paper Cristina Bicchieri published a different application of game-theoretic ideas to the understanding of scientific knowledge. In her article "Methodological Rules as Conventions" (1988), she tried to illuminate the establishment of specific scientific methods from the point of view of David Lewis' theory of conventions. Opposing both functionalist explanations of scientific rules, either epistemic (that would justify them by their ability to sustain preordained, usually cognitive goals), or sociological (that would explain the adoption of rules by individual scientists by resource to the pressure of some relevant social groups), she offers an account according to which is based on the coordination of strategic choicess of individual scientists, a coordination that can lead in principle to different sets of rules. In particular, the main element in the decision of a scientist to adopt a method, criterion, or procedure is the fact that he expects that his colleagues also obey it. In this sense, the choice of a rule or systems of rules instead of other is *arbitrary*, for individual scientists would have been equally happy if some different norms were collectively adopted. Being an equilibrium of a coordination game explains the (relative) *stability* of scientific methods, for once the individuals expect others to comply, it is costly for each agent to follow a different rule. But the rationality of the agents cannot explain why some norms *emerge* instead of others. Actually, if, as Bicchieri assumes, several norms would lead to different coordination equilibria which are exactly as good as the other for all the agents, then it is true that the 'social choice' of a rule instead of other would be arbitrary, but it is reasonable to assume that the game that individual scientists face is not a game of *pure* coordination: different rules may have more value for different and also for the samescientists, even taking into account the gains from coordination (stated differently, it is possile that in many cases the resulting game is of the 'Battle of the Sexes' type), a fact that Bicchieri marginally acknowledge in one footnote (Bicchieri (1988), p. 488).

The most extensive and systematic application of economic modelling to the analysis of epistemological problems has been performed by the philosopher Philip Kitcher in a paper entitled "The Division of Cognitive Labour" (1990), and later extended to constitute the last chapter of his book *The Advancement of Science* (1993).[9] Kitcher's main goal in those contributions was to develop an economic branch of the field of 'social epistemology', based on a methodologically individualist conception of social processes (that sees social outcomes as deriving from the interconnected decisions of individuals) and simultaneously on a reliabilist conception of knowledge (which takes progress towards the truth as the epistemic goal of science). The role for social epistemologists would be "to identify the properties of epistemically well-designed social systems, that is, to specify the conditions under which a group of individuals, operating according to various rules for modifying their individual practices, succeed, through their interactions, in generating a progressive sequence of consensus practices" (Kitcher,

[9] Some early criticisms of Kitcher's game-theoretic approach were Fuller (1994), Hands (1995), Levi (1995), Mirowski (1995), and Solomon (1995).

1993, p. 303). The most idiosyncratic concept in this quotation is that of a *consensus practice*, which is one of the central ideas in Kitcher's book: by a 'practice', Kitcher refers to the collection of cognitive resources an individual or a group has, including such things as valid concepts, relevant questions, accepted answers, and exemplars of good experimental or argumentative procedures (op.cit., p. 74); a group's consensus practice will contain, so to say, the intersection of the elements included within its members' individual practices, together with the social mechanisms of deferred cognitive authority which allow to create a 'virtual' consensus, i.e., those things every member *would* accept if she happened to follow all the threads of deferred authority (op. cit., pp. 87f.). This idea of a virtual consensus represents what a scientific *community* knows (or pretends to know), in spite of no one of its single members being able of getting all that information in her own mind. A series of consensus practices is epistemically progressive if it develops an increasingly more accurate set of concepts (i.e., concepts better and better fitting real natural types) and an increasing amount of right solutions to scientific problems (i.e., objectively true descriptions of actual events and causal mechanisms). Hence, for Kitcher the task for social epistemology is, basically, to analyse those mechanisms according to which scientists interact and produce a consensus practice, in order to critically assess whether we can expect they lead to cognitive progress, or to what extent they do it.

Kitcher's strategy is divided into two stages. In the first place, he discusses how individual scientists act when taking into account their colleagues' actions, in particular, how they take their decisions about how much authority to confer those colleagues, as well as about how to compete or cooperate with them. In the second place, Kitcher analyses what epistemic consequences different distributions of researchers' efforts may have. In order to elaborate this strategy, Kitcher employs models from standard and evolutionary game theory, as well as from Bayesian decision theory. Although this strategy is grounded on methodological individualism, when it goes to normative problems it finally rests on the idea that there is some kind of collective (or 'objective') standard of epistemic value against which to measure the actual performances of a scientific community:

> There are two types of inquiry that are worth pursuing: first, we want to know what, given the range of possibilities, is the *best* approach to the problem situation in which we are interested; second, we should scrutinize which of the available combinations of individual decision procedures and sets of social relations would move the community closer to or further away from the *optimal* approach.
>
> (op. cit., p. 305; my italics).

With this way of describing the normative ambition of his project, Kitcher's models could perhaps be catalogued within the optimisation paradigm we studied in section 2; however, the relevance he gives to the interaction processes and to the decisions of individual scientists when competing against each other (what Kitcher calls 'the entrepreneurial predicament'), connects equally well his approach with the exchange paradigm. Actually, this combination is what makes of the model building approach a much more systematic way to analyse science as an economic process, as compared to the contributions discussed in the past sections. Nevertheless, in my opinion Kitcher does not explain in a satisfactory way what is the connection between the 'objective' evaluation he refers to and those of each individual scientist; for example, does the former emerge from the latter through some kind of aggregation procedure, as if it were a kind of '(cognitive) social welfare function'? The difficulty of providing such a derivation was soon indicated by Wade Hands (1995, p. 617f.). Or is there really no connection between both types of valuation, and the former is just assumed by the philosopher while pretending to act as a (fictitious) benevolent planner of scientific research? More probably, Kitcher's view about this problem seems to be, first, that individual scientists' utility functions can be decomposed into a 'social' factor, basically related to the Bourdieu's 'credit' concept, and an 'epistemic' factor, which would basically be the pursuit of objective truth; and second, that this epistemic part of scientists' preferences will accord, in principle, with the 'objective' valuation assumed by the philosopher. Many of Kitcher's assertions substantiate this interpretation, for example:

> We can think of the problems that concern me as including those that would face a philosopher-monarch, interested in organizing the scientific work force so as to promote the collective achievement of significant truth. Science, of course, has no such benevolent dictator. In consequence, individual scientists face coordination problems. If we suppose that they internalize the (fictitious) monarch's values, how will they fare? If we assume instead that they are motivated in baser ways or that they are locked into systems of authority and deference, will they necessarily do worse than a society of unrelated individuals, each of whom is pure of heart?
> (op. cit., p. 305).

We can divide the models elaborated in Kitcher (1993, ch. 8) into two different groups. The first group (sections 2 to 12) relates to what the author call *'the entrepreneurial predicament'*: how do researchers pursuing scientific recognition are expected to behave in response to their colleagues' behaviour. The problems Kitcher attacks here are basically the following ones:

a) during a research process carried out by a scientist, some intermediate parts can either be directly performed by her, or she can borrow the result announced by another colleague; Kitcher analyses how this decision depends on the probability that both the scientist an her colleagues have of finding a right solution to the intermediate problem, and on the probability the former has of being the first solver of the main problem, depending on her previous choice, as well as on the different resources each researcher may have;

b) the previous discussion depends on the assumption that individual scientists may 'calibrate' the reliability of their colleagues' assertions; the next group of models Kitcher develops is directed, then, to analyse how this 'calibration' may take place, particularly when *A* has to use *B*'s assessments of the reliability of *C*, and when *A*'s assessment of *B* depend on *B*'s assessment of *A*; different distributions of 'authority' may also lead to different decisions about when to borrow other's result or when to do the job by oneself;

c) finally, scientists don't have to decide only whether to accept the result announced by a colleague or not: they also have the option of *replicating* it, postponing the other decision until the replication has confirmed or disconfirmed the result; here the main economic problem is that of determining the optimal number of replications (taking into account that they are costly), and of knowing whether this number can be reached by the independent decisions of individual scientists.

The second set of models relate to what Kitcher calls *'the division of cognitive labour'* (a label which could also have been applied to the former group of models, nevertheless), by which he refers not to the decisions about what problems to attack, but about what methods to employ to do it, as well as what solutions to those problems to accept (i.e., the old methodological problem of theory choice). In the latter case, by 'choosing' a theory Kitcher distinguishes two successive stages: first, scientists may choose a theory 'to work with', as a kind of exploration, so to say; and second, at some moment a consensus about what is 'the' right theory must emerge. The main question that concerns Kitcher is the difference between the distribution of efforts which is optimal from a cognitive point of view, and the distribution that would arise if each researcher were individually pursuing her own interest, and hence the problem is basically one of *coordination*. Kitcher considers several cases, according to whether individual scientists are motivated just by the pursuit of truth, or by the pursuit of success, or by a mix of both goals, and also according to whether all scientists are assumed to have the same utility preferences and estimations about the probability of each theory being right, or there is some motivational or cognitive diversity. Perhaps the most relevant conclusion Kitcher offers (pp. 364 f.) is that in a community of researchers whose members were just motivated by professional glory (i.e., they didn't mind about whether the finally accepted theory is true or false), they would always choose that theory with the highest probability of being finally accepted, and so no one would pursue other theories or methods, but, just with a slight weight attached to the goal of truth in the scientists' utility function, a distribution of efforts close to the optimum will be attained.[10]

I pass now to describe the Bayesian model offered in Goldman and Shaked (1991). In this model, researchers gain recognition for their success in modifying the subjective probabilities that their colleagues attach to each possible solution of a scientific problem. In order to do this, researchers perform two different types of acts: investigative acts (ranging from observation to the formulation of new hypotheses), and speech acts (which can be either the presentation of evidence in favour of a solution, or the criticism of a previous speech act). The structure of the research process can be described hence as a sequential game: in the first stage, a researcher decides what investigative act to perform; second, nature 'chooses' an outcome thereof; third, the researcher decides how to interpret the outcome, i.e., what solution she thinks the outcome supports (we can understand this move as the proposal of a particular solution); fourth, the other researchers change their subjective probabilities; and lastly, the first scientist gets a recognition level determined by the application of some socially sanctioned recognition-allocating mechanism. Goldman and Shaked prove that, in a scenario like this, researchers whose utility function depends uniquely on the attainment of such a type of recognition will perform *on the average* those experiments and those speech acts more conductive to an increment in the possession of truth by the members of the scientific community (i.e., an increment in the probabilities attached to the true solutions). Some criticisms can be made to this model: in the first place, it assumes that researchers will not misrepresent their subjective probabilities (after all, if each scientist wants her own theory to be the winning one, she will not assert that she thinks another theory is better); in the second place, in the Goldman-Shaked model the proposed recognition-allocating mechanism is simply imposed to the scientists, but perhaps these would prefer to play the game according to different rules.

A couple of mathematical models of scientific knowledge production were developed by Philip Mirowski, one of them in collaboration with Steve Sklivas (Mirowski and Sklivas, 1991; Mirowski, 1996; both reprinted in Mirowski (2004)). In the first of this papers, an attempt is made of explaining the observation made by sociologists of science, according to which researchers almost never perform *exact replications* of the experiments made by

[10] Rueger (1996) extends Kitcher's argument in order to take into account scientists' cognitive risk aversion.

others (contrarily to what positivist expositions of the scientific method prescribed), though the results of those experiments are actually *employed* in ensuing research practice (and in this sense, they are 'reproduced'). Mirowski and Sklivas develop a game theoretic account of this behaviour: the first performers of an experiment gain nothing from independent replications if these are successful, and loose if they fail, but they gain from the further use of the experiment by others; use (or 'reproduction') is costly for those researchers who perform it, though exact replication is even more costly; on the other hand, only failed replication gives a positive payoff to replicators; lastly, the more information is conveyed in the report of the original experiment, the less costly both use and replication become. From these assumptions, Mirowski and Sklivas derive the conclusion that the optimum strategy of the scientist performing an original experiment is to provide such an amount of information that is enough to incentivate use, but not enough to make replication worthwhile; replication will only have a chance if the editors of the journals command to provide still more information in the experimental reports. In the second of the models referred to above, Mirowski compares the process of measuring a physical constant to the process that determines prices in the markets: as differences in the price of the same good at different places create an opportunity to arbitrage, inconsistencies in the measured values of a constant (derivable from the use of some accepted formulae -e.g., physical laws- and the values of other constants) create an opportunity to make further relevant measurements. Graph theory is employed at this point to describe the interconnection between the measured values of several constants (the 'nodes' of the graph) and the formulae connecting them (the 'edges'), and to suggest an index measuring the degree of mutual inconsistency the existing values display. Interestingly enough, the application of this index to several branches of science shows that economics has been much less efficient in the construction of consistent sets of measures, a fact Mirowski explains by the reluctance of neo-classical economists to create an institutional mechanism capable of recognising and confronting this shortcoming (an explanation which could be tested by comparing the measures assembled by economic agencies with, say, a neo-classical or a Keynesian orientation).

Other of the authors who has contributed more to the development of an economics of scientific activity has been Paul David, partly in close collaboration with Partha Dasgupta (e.g., Dasgupta and David (1994)). Most of their work on the topic, however, fits better the category of 'economics of science' than that of 'economics of scientific knowledge' (as I defined them in sec. 1), for they have basically tried to look for an economic explanation of the social institutions of research and development. In some more recent papers, nevertheless, David has articulated some ideas which definitely belong into the ESK field, particularly in David (1998a), where he presents some models about the behaviour of reputation-seeking scientists when deciding what opinion to express (see also David (1994) for an analysis of cumulative advantage in science, and (1998b) for a hypothesis about the origin of peer review). In one of these models (David (1998a), pp. 138 ff.), researchers have to adopt or reject a theory, T, which *in the long run* will be either accepted as right by the community, or rejected as wrong, but which it is now under discussion, so that some researchers are currently accepting it, and others rejecting it; if a scientist thinks with probability p that the theory will be collectively adopted in the end (considering her own private knowledge and the opinion expressed by her neighbour colleagues), the utility she expects to get will also depend on whether now there is a majority or a minority of colleagues accepting T, in the following way: let a be the utility of adopting T if it is now majoritarily rejected, but collectively adopted in the future ('being right with the few'); let b be the utility of rejecting T under the same conditions ('being wrong with the crowd'); let c the utility of adopting T if it is now majoritarily accepted, and collectively adopted in the future ('being right with the crowd'), and let d the utility of rejecting T under these conditions ('being wrong with the few'); lastly, let us assume that $a > c > b > d$. It follows that the scientist will adopt the majority opinion if and only if $(1-p)/p < (c - d)/(a - b)$. This entails that, if the difference between being eventually right having defended a minority opinion, and being eventually wrong but having defended the majority opinion, is low enough in reputation terms, then conformity to the majoritarian opinion will be a dominant strategy. David also formulates a stochastic graph-theoretic model of consensus formation (*op. cit.*, pp. 145 ff.), closely related to the one just described. Imagine that at each time one single scientist is randomly selected to give her the chance of modifying her opinion about T; she will do it depending on the opinions that the colleagues to which she is 'connected' have expressed in the moment immediately before. This 'voting' mechanism generates a Markovian process that has unanimous acceptance of T and unanimous rejection as the only absorbing states, states that are reached with a probability equal, respectively, to the proportion of individual scientists accepting or rejecting T at the beginning of the process.

Other economic explanations of the 'construction' of scientific consensus were developed in the late nineties. In 1997, within a conference organised at Notre Dame by Philip Mirowski and Esther Mirjam Sent on "The Need for a New Economics of Science", a couple of papers were presented (Oomes (1997), Brock and Durlauf (1999)) in which the choice of a theory by each individual scientist depends on two factors: an individual effect, which takes into account the researcher's 'private' assessment of the theory, and a conformity effect, which takes into account the (expected) choices of her colleagues. A similar approach, though on much weaker mathematical assumptions, was independently elaborated in Zamora Bonilla (1999b). The main conclusions of all these models were in part related to those of David's paper, but went further than it in some respects: first, more than one social equilibrium (i.e., a distribution of individual choices such that nobody has an interest in making a different choice, given the choices of her colleagues) are possible; second, path-dependence is a significant factor in the attainment of an equilibrium (e.g., two scientific communities having *the same* empirical evidence about a couple of alternative theories might end making different choices, if their data had just been accumulated in a different *order*); but, third, contrarily to what happened in David's model, some equilibrium states can correspond to a non unanimous theory choice (i.e., diversity of individual judgements can take place in the equilibrium). Another important conclusion of these models is that, as the factors influencing the individual effects change (e.g., by finding new empirical or theoretical arguments which affect the assessment each scientist makes), the number of scientists accepting a theory can suffer a sizeable change at some point, even though those influencing factors have accumulated by small marginal increments (i.e., the dynamics of scientific consensus is not necessarily linear). Zamora Bonilla (forthcoming a) generalises the analysis for cases where individual choices depend not only on how many colleagues are accepting a theory, but on *which ones* are doing it. The last two referred papers additionally consider the possible effects of *collective choices*, i.e., the forming of (not necessarily universal) coalitions in which every member would be interested in adopting the theory if and only if the other members did the same; the paper shows that, if coalitions are feasible, in most of the cases where two equilibria existed, only one of them becomes stable under collective choice (i.e. no other coalition can force a move to the other equilibrium), and, if it happened that one of the equilibria was Pareto-superior with respect to the other, the former one will be coalition proof. This last conclusion suggests that there is a middle ground between 'free market' and 'social planning' approaches to the economics of scientific knowledge: the epistemic efficiency of science perhaps would not mainly come from the unintended coordination of individual choices, nor from the calculations of a single planner, but from *the free constitution of groups*.

In a similar vein, Zamora Bonilla (2002) presents a model in which the members of a scientific community can choose the 'confirmation level' (or other measure of scientific quality) a theory must surpass in order to become acceptable, under the assumption that scientists are motivated not only by the quality of the theories, but mainly by being recognised as proponents of an accepted theory. The chances of getting recognition are too small both if the chosen level is very low (for then too many successful theories will exist to compete with), and if the level is very high (for then it would be very difficult to discover an acceptable theory). Zamora Bonilla (forthcoming b) offers a game theoretic analysis of how the *interpretation* of an experimental result is chosen (in a way similar to Goldman and Shaked (1991), but taking into account strategic considerations), which illuminates the role of social mechanisms in scientific communication, understood as constraints in the way the 'game' between authors and readers is played. That the choice of scientific norms is an appropriate subject for economic analysis has being recognised in other works; for example, Kahn, Landsburg and Stockman (1996) also analyse from an economic viewpoint the choice of an empirically versus a theoretically oriented strategy by scientists; Slembeck (2000) designs a betting mechanism to compel scientists to agree on empirical facts; Michael Strevens (2003) employs optimality analysis to justify the use of the priority rule in allocating scientific resources; lastly, Max Albert (2005) offers a dynamic game model, according to which scientists select a methodological rule depending on the choices of their colleagues.

4.2. Institutionalist theories.

I will end this survey of the main contributions to the economics of scientific knowledge by discussing three works which attempt to offer a more or less systematic conception of the process of scientific discovery, but that abstain to employ mathematical models as their main analytical tool. The first of the contributions I will discuss is Wible (1998), which also was the first full length book on the topic. The most important parts of it in connection to ESK are its five last chapters, where Wible presents an economic view that he explicitly opposes to the idea of science as a 'market for ideas'. According to Wible (*op. cit*., ch. 8-9), automatic self-corrective mechanisms like those assumed in 'perfect markets' do not exist, or are not too powerful, in the process of scientific knowledge production; he goes even further and considers science as an economic institution *opposite* to perfectly competitive markets, and the organisational structure of scientific decision making (research programmes, university departments, peer review, public funding, and so on) is explained in a new-institutionalist way, as ways to minimise 'transaction costs', and it is contrasted with the organisation of the market-governed aspects of science (i.e., those related to science as a consumer of economic inputs - labour, commodities...-, and as a producer of commercial applications). The last three chapters are the most philosophical of the book. Here Wible offers an evolutionary account of economic and scientific rationality, understanding scientific decisions as essentially constrained by scarcity of resources, and, in general, presenting all decision making as an evolving, multidimensional and non-mechanical process of problem solving. Under this interpretation, actual scientific rationality would not be 'justificationist', in the same way that economic rationality would not really be 'maximising'. Wible also discusses the problem of self-reference (recall section 3.4): an economics of economics will necessarily be self-referential, and this may lead to detect internal inconsistencies or infinite regresses within some economic arguments; but, from an economic point of view, an infinite regress must end at some point, since our cognitive resources are limited; Wible suggests that a 'critical', 'non-justificationist' rationality, is necessary in order to handle these logical problems. This issue of reflexivity is also applied to the tacit 'economics of science' that most mainstream economists seem to assume. They tend to see their own activity (scientific research) as a competitive process similar to those which prevail in the economy, a process which tends naturally to some efficient equilibrium; but, Wible argues, the features of scientific knowledge entail that science is an imperfect and slow evolutionary process, not explainable in terms of mechanistic equilibrium concepts. One possible criticism that can be made to Wible's approach is that he gives almost no insight about how such an evolutionary account of science would look like once it were elaborated in detail, for one fears that he is not thinking in something like evolutionary *mathematical* models (e.g., those of evolutionary game theory, or those of the Nelson-Winter type). On the other hand, in the first chapters of the book Wible does not hesitate in offering some simplified equilibrium models of some types of scientists' 'misconduct' (namely, fraud and replication failure -which, according to the model by Mirowski we examined in the past subsection, needs not be consider a 'failure' at all), and so, one wonders why the same explanatory strategy would not be appropriate to analyse 'normal' research behaviour.

Another institutionalist approach to the study of scientific knowledge production has been developed by Christoph Lütge (2001) and (2004), who uses constitutional political economy as a complement to a naturalised philosophy of science (see Brennan and Buchanan (1975) for a clear exposition of constitutional economics). The main elements of this approach are the following: first, the essential object of economic analysis is not individual decision making, but *interactions*, where considerations of strategic decision arise; second, the main purpose of an economic analysis is not theoretical, but *practical*, i.e., promoting the design of institutions; and third, there are no external normative criteria, for *consensus* amongst the interacting agents is the only source of normativity. Lütge's proceeds, then, to identify *social dilemma situations* in scientific research, situations that create the opportunity for introducing institutional mechanisms allowing agents to overcome those problems. Drawing on Martin Rudwick's case study of the 'great Devonian controversy', Lütge identifies three different dilemma situations (which are not necessarily unique): the priority dilemma (i.e., how to allocate research efforts and scientific recognition), the property rights dilemma (i.e., whether, or how much, to plagiarise the works of others), and dilemma of access to objects of study (i.e., whether to 'monopolise' these objects or not). Many of these cases can lead to a Prisoner's Dilemma situation, that can only be 'solved' by collectively changing the rules according to which the interactions between researchers take place. Unfortunately, Lütge does not analyse how this collective decision could be taken. Another significant Buchanian thesis stressed by Lütge is that, though consensus is the fundamental normative criterion to evaluate scientific items, it can be applied to very different levels: scientists may agree (or disagree) about specific cognitive *achievements* (theory, data, and so on), but consensus on *norms* is much more important from a normative point of view, particularly if we distinguish amongst different *levels* of methodological or institutional rules (more on this below).

The last work I am going to discuss is surely the until now most systematic attempt to develop an economic institutionalist theory of science, Yangfei Shi's book *The Economics of Scientific Knowledge*. In the first chapters, Shi analyses the behaviour of scientists, both as a producers of scientific information, and as a *consumers* of the information produced by others, and also the *exchange* relationships which emerge between self-interested researchers, each one playing simultaneously the roles of producer and consumer. The rest of the book is devoted to analyse the *institutional structure* of science, i.e., the system of norms which regulate the behaviour of scientists (norms that are interpreted as solutions to collective action problems). These norms are classified into three different groups: 'distributive rules' (i.e., rules about the allocation of resources), 'constitutional rules' (cognitive paradigms and authority structures), and 'aggregative rules' (the mechanisms creating 'scientific order'). In spite of the fact that Shi's analysis of scientific norms is more complete than other ones, it can be subjected to several criticisms. In the first place, three very different kinds of things are conflated in his classification: individual routines, social norms, and what we could call 'equilibrating mechanisms' (more or less equivalent to Shi's 'aggregative rules'). By individual routines I mean all types of regular practices which can be observed in the behaviour of a single scientist; these practices can spread towards other scientists by imitation, but even if they are universally adopted, this does not make of them a social norm automatically, for what is characteristic of norms is that they are *taken as compulsory* by the individuals (even if everybody disobeys them). On the other hand, 'equilibrating mechanisms' are neither individual routines nor social norms, and they can even be hidden for the individuals, as in the case of an 'invisible hand mechanism', in which case they can hardly be taken as 'norms', but rather as the 'natural laws' of a social structure. In the second place, Shi's notion of 'constitutional rules' is at odds with what is ordinarily meant by these expression in standard constitutional political economy, where the idea is, first, that a sharp distinction can be made between choices made *under* rules and the choice *of* the rules themselves, and second, that social norms can be divided into 'constitutional' and 'non-constitutional' ones, the former being those which establish (ideally, by unanimity) the collective choice mechanisms by which the latter are going to be chosen (not necessarily by unanimity). On the other hand, a constitutional economists would have expected to find in Shi's discussion a distinction between cognitive norms of different *levels*, so that, in case of disagreement among scientists about a particular theory or a certain methodological practice, researchers could appeal to some 'higher level rules' to reach a common decision, if this is possible, and to rules of still a higher level if there is also disagreement about the latter, and so on. From this point of view, the 'constitutional rules' of a scientific community would only be those fundamental principles to which its members can resort in case of discursive conflict.

ACKNOWLEDGEMENTS

This paper was prepared while its author was taking part in Spanish Government's research projects BFF2002-03353 ('Cognitive roots in the assessment of new information technologies'), SEJ2004-20076-E ('Science, democracy, and economics') and HUM-2005-01686/FISO ('The emergence of technoscientific norms'). Generous support from Urrutia Elejalde Foundation is also acknowledged.

REFERENCES

Albert, M., 2005, 'Product Quality in Scientific Competition', mimeo.

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

Balzer, W., C. U. Moulines, and J. Sneed, 1987, *An Architectonic for Science: The Structuralist Program*, Dordrecht, D. Reidel.

Bartley, W. W., 1990, *Unfathomed Knowledge, Unmesured Wealth: On Universities and the World of Nations*, La Salle, Ill., Open Court.

Bicchieri, C., 1988, "Methodological Rules as Conventions", *Philosophy of the Social Sciences*, 18:477-95.

Blais, M. J., 1987, 'Epistemic Tit for Tat', *The Journal of Philosophy*, 84:363-75.

Bloor, D., 1976, *Knowledge and Social Imagery*, London, Routledge.

Bourdieu, P., 1975, "The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason", *Social Science Information*, 14.6:19-47 (quoted from its reprint in M. Biagioli (ed.), 1999, *The Science Studies Reader*, London, Routledge).

Bracanovic, T., 2002, 'The Referee's Dilemma: The Ethics of Scientific Communities and Game Theory', *Prolegomena*, 1:55-74.

Brennan, G., and J. Buchanan, 1985, *The Reason of Rules*, Cambridge, Cambridge University Press.

Brenner, Th., 1999, *Modelling Learning in Economics*, Celtenham, Edward Elgar.

Brock, W. A., and S. N. Durlauf, 1999, 'A Formal Model of Theory Choice in Science', *Economic Theory*, 14:113-30.

Buchanan, J., 1996, *The Economics and the Ethics of Constitutional Order*, Ann Arbor, University of Michigan Press.

Butos. W. N., and P. J. Boettke, 2002, 'Kirznerian Entrepreneurship and the Economics of Science', *Journal des Economistes et des Etudes Humaines*, 12:119-30.

Callon, M., 1994, 'Is Science a Public Good?', *Science, Technology, and Human Values*, 19:395-424.

Callon, M., 1995, 'Four Models for the Dynamics of Science', in S. Jasanoff, G. E. Markle, J. C. Petersen, and T. Pinch (eds.), *Handbook os Science and Technology Studies*, Thousand Oaks, Sage, pp. 29-63.

Coase, R. H., 1974, 'The Market for Goods and the Market for Ideas', *Amercian Economic Review Papers and Proceedings*, 64:384-91.

Dasgupta, P., and P. David, 1994, 'Toward a New Economics of Science', *Research Policy*, 23:487-521.

David, P., 1994, 'Positive Feedbacks and Research Productivity in Science: Reopening Another Black Box', in G. Granstrand (ed.), *Economics of Technology*, Amsterdam, Elsevier, pp. 65-89.

David, P., 1998a, 'Communication Norms and the Collective Cognitive Performance of 'Invisible Colleges'', in G. Barba *et al.* (eds.), *Creation and Transfer of Knowledge Institutions and Incentives*, Berlin, Springer, pp. 115-63.

David, P., 1998b, 'Clio and the Economic Organization of Science: Common Agency Contracting and the Emergence of 'Open Science' Institutions', *American Economic Review Papers and Proceedings*, 88:15-21.

Davis, J. B., 1998, 'The Fox and the Henhouses: The Economics of Scientific Knowledge', *History of Political Economy*, 29:741-6.

Friedman, M., 1953, 'The Methodology of Positive Economics', in *Essays in Positive Economics*, Chicago, The University of Chicago Press., pp. 3-43.

Fuller, S., 1988, *Social Epistemology*, Bloomington, Indiana University Press.

Fuller, S., 1994, 'Mortgaging the Farm to Save teh (Sacred) Cow', *Studies in History and Philosophy of Science*, 25:251-61.

Fuller, S. 2000, *The Governance of Science*, Philadelphia, The Open University Press.

Goldman, A., and J. C. Cox, 1996, 'Speech, Truth, and the Free Market for Ideas', *Legal Theory*, 2:1-32.

Goldman, A., and M. Shaked, 1991, 'An Economic Model of Scientific Activity and Truth Acquisition', *Philosophical Studies*, 63:31-55.

Goodin, R. E., and G. Brennan, 2001, 'Bargaining over Beliefs', *Ethics*, 111:256-77.

Hands, D. W., 1994a, 'The Sociology of Scientific Knowledge and Economics: Some Thoughts on the Possibilities', in Robert Backhouse (ed.), *New Perspectives in Economic Methodology*, London: Routledge, pp. 75-106.

Hands, D. W., 1994b, 'Blurred Boundaries: Recent Changes in teh Relationship Between Economics and The Philosophy of the Natural Science', *Studies in History and Philosophy of Science*, 25:751-72.

Hands, D. W., 1995, 'Social Epistemology Meets the Invisible Hand: Kitcher on the Advancement of Science', *Dialogue*, 34:605-21.

Hands, D. W., 2001, *Reflection without Rules. Economic Methodology and Contemporary Science Theory*, Cambridge, Cambridge University Press.

Hardwig, J., 1991, 'The Role of Trust in Knowledge', *The Journal of Philosophy*, 88:693-700.

Hayek, F. A., 1948, 'The Use of Knoewledge in Society', in *Individualism and Economic Order*, Chicago, The University of Chicago Press, pp. 77-91.

Hempel, C. G., 1960, 'Deductive Inconsistencies', *Synthese*, 12:439-69 (reprinted in *Aspects of Scientific Explanations*, New York, The Free Press, pp. 53-79).

Hilpinen, R., 1968, *Rules of Acceptance and Inductive Logic*, Amsterdam, North-Holland.

Hull, D., L., 1988, *Science as a Process. An Evolutionary Account of the Social and Conceptual Development of Science*, Chicago, The University of Chicago Press.

Kahn, J. A., S. E. Landsburg, and A. C. Stockman, 1996, 'The Positive Economics of Methodology', *Journal of Economic Theory*, 68:64-76.

Kitcher, Ph., 1990, 'The Division of Cognitive Labor', *The Journal of Philosophy*, 87:5-22.

Kitcher, Ph., 1993, *The Advancement of Science: Science without Legend, Objectivity without Illusions*, Oxford, Oxford University Press.

Kuhn, Th.S., 1977, 'Objetivity, Value Judgments, and Theory Choice', in *The essential tension*, Chicago, The University of Chicago Press.

Latour, B., 1987, *Science in Action*, Cambridge, Ma., Harvard University Press.

Latour, B., and S. Woolgar, 1979, *Laboratory Life. The Social Construction of Scientific Facts*, Beverly Hills, Sage.

Leonard, Th. C., 2002, 'Reflection on Rules in Science: An Invisible-Hand Perspective', *Journal of Economic Methodology*, 9:141-168.

Levi, I., 1967, *Gambling with Truth*, New York, Knopf.

Levi, I., 1995, 'Cognitive Value and the Advancement of Science', *Philosophy and Phenomenological Research*, 55:619-625.

List, Ch., and Ph. Pettit, 2002, 'Aggregating Sets of Judgements: An Impossibility Result', *Economics and Philosophy*, 18:89-110.

Lütge, Ch., 2001, *Ökonomische Wissenschaftstheorie*, Würtzburg: Königshausen & Neumann.

Luetge, Ch., 2004, 'Economics in Philosophy of Science: A Dismal Contribution?', *Synthese*, 140:279-305.

Maher, P., 1993, *Betting on Theories*, Cambridge, Cambridge University Press.

Mäki, U., 1999, 'Science as a Free Market: A Reflexivity Test in an Economics of Economics', *Perspectives on Science*, 7:486-509.

Mäki, U., (ed.), 2002, *Fact and Fiction in Economics*, Cambridge, Cambridge University Press.

Mäki, U., 2004, 'Economic Epistemology: Hopes and Horrors', *Epistemce*, 1:211-22.

Mayer, Th., 1993, *Truth versus Precision in Economics*, Aldershot, Edward Elgar.

McQuade, Th. J., and W. N. Butos, 2003, 'Order-Dependent Knowledge and the Eocnomics of Science', *The Review of Austrian Economics*, 16:133-152.

Mirowski, Ph., 1995, 'Philip Kitcher's *Advancement of Science*: A Review Article', *Review of Political Economy*, 7:227-241.

Mirowski, Ph., 1996, 'A Visible Hand in the Marketplace for Ideas', in M. Power (ed.), *Accounting and Science*, Cambridge, Cambridge University Press.

Mirowski, Ph., 2004, *The Effortless Economy of Science?*, Durham, Duke University Press.

Mirowski, Ph., and E.-M. Sent (eds.), 2002a, *Science Bought and Sold: Essays in the Economics of Science*, Chicago and London: The Chicago University Press.

Mirowski, Ph., and E.-M. Sent, 2002b, 'Introduction', in Mirowski and Sent (2002a), pp. 1-66.

Mirowski, Ph., and S. Sklivas, 1991, 'Why Econometricians Don't Replicate (Although They Do Reproduce)?, *Review of Political Economy*, 31:146-63

Mueller, D. C., 2003, *Public Choice III*, Cambridge, Cambridge University Press.

Nelson, R. R., 1959, 'The Simple Economics of Basic Scientific Research', *Journal of Political Economy*, 67:297-306.

Niiniluoto, I., 1987, *Truthlikeness*, Dordrecht, D. Reidel.

Niiniluoto, I., 1998, 'Verisimilitude: The Third Period', *British Journal for the Philosophy of Science*, 49:1-29.

Oomes, N. A., 1997, 'Market Failures in the Economics of Science: Barriers to Entry, Increasing Returns, and Lock-In by Historical Events', paper presented to the conference on *The Need for a New Economics of Science*, Notre Dame.

Peirce, Ch. S., 1879, 'A Note on the Theory of the Economy of Research', *United Stated Coast Survey for the Fiscal Year Ending June 1876*, Washington D.C. (quoted from the reprint in Mirowski and Sent (2002a), pp:183-190).

Pickering, A., 1995, *The Mangle of Practice: Time, Agency, and Science*, Chicago, The University of Chicago Press.

Polany, M., 1962, 'The Republic of Science: Its Political and Economic Theory', *Minerva*, 1:54-73 (quoted from the reprint in Mirowski and Sent (2002a), pp. 465-85.

Popper, K. R., 1963, *Conjectures and Refutations. The Growth of Scientific Knowledge*, London, Routledge and Keagan Paul.

Popper, K. R., 1970, 'Normal Science and Its Dangers', in I. Lakatos and A. Musgrave, *Criticism and the Growth of Knowledge*, Cambridge, Cambridge University Press.

Radnitzky, G., 1986, 'Towards an 'Economic' Theory of Methodology', *Methodology and Science*, 19:124-47.

Radnitzky, G., 1987, 'Cost-Benefit Thinking in the Methodology of Research: the 'Economic Approach' Applied to Key Problems of the Philosophy of Science', in G. Radnitzky and P. Bernholz (eds.) *Economic imperialism: the economic approach applied outside the field of economics*, Paragon House, New York.

Rescher, N., 1976, 'Peirce and the Economy of Research', *Philosophy of Science*, 43:71-98.

Rescher, N., 1978a, *Scientific Progress. A philosophical essay on the economics of research in natural science*, Pittsburg, The University of Pittsburgh Press - Basil Blackwell.

Rescher, N., 1978b, *Peirce's Philosophy of Science*, Notre Dame, Notre Dame University Press.

Rescher, N., 1989, *Cognitive Economy. The Economic Dimension of the Theory of Knowledge*, Pittsburg, The University of Pittsburgh Press.

Rescher, N., 1996, *Priceless Knowledge. Natural science in economic perspective*, Lanham, Rowman & Littlefield.

Rueger, A., 1996, 'Risk and Diversification in Theory Choice', *Synthese*, 109:263-80.

Salanié, B., 2000, *Micoreconomics of Market Failure*, Cambridge, Ma., The MIT Press.

Schuessler, A. A., 2000, *A Logic of Expressive Choice*, Princeton, Princeton University Press.

Shi, Y., 2001, *The Economics of Scientific Knowledge: A Rational-Choice Neoinstitutionalist Theory of Science*, Cheltenham, Edward Elgar.

Slembeck, T., 2000, 'How to Make Scientists Agree: An Evolutionary Betting Mechanism', *Kyklos*, 53:587-92.

Sneed, J., 1989, 'MicroEconomic Models of Problem Choice in Basic Science', *Erkenntnis*, 30:207-24.

Solomon, M., 1995, 'Legend Naturalism and Scientific Progress: An Essay on Philip Kitcher's *The Advancement of Science*', *Studies in History and Philosophy of Science*, 26:205-18.

Stephan, P., and D. B. Audretsch (eds.), 2000, *The Economics of Science and of Innovation* (2 vols.), Cheltenham, Edward Elgar.

Strevens, M., 2003, 'The Role of the Priority Rule in Science', *The Journal of Philosophy*, 100:55-79.

Sugden, R. 2002, 'Credible Worlds: The Status of Theoretical Models in Economics', in Mäki (2002), pp. 107-36.

Turner, S., 2002, 'Scientists as Agents', in Mirowski and Sent (2002a), pp. 362-84.

van Fraasen, B., *The Scientific Image*, Oxford, Clarendon Press.

Walstad, A., 2001, 'On Science as a Free Market', *Perspectives on Science*, 9:324-40.

Walstad. A., 2002, 'Science as a Market Process', *The Independent Review*, 7:5-45.

Weintraub, R., 1990, ‘Decision-Theoretic Epistemology’, *Synthese*, 83:159-77.

Wible, J. R., 1998, *The Economics of Science: Methodology and Epistemology as if Economics Really Mattered*, London, Routledge.

Wray, K. B., 2000, ‘Invisible Hands and the Success of Science’, *Philosophy of Science*, 67:163-75.

Ylikoski, P., 1995, ‘The Invisible Hand and Science’, *Science Studies*, 8:32-43.

Zamora Bonilla, J. P., 1996, ‘Verisimilitude, Structuralism and Scientific Progress’, *Erkenntnis*, 44:25-47.

Zamora Bonilla, J. P., 1999a, ‘Verisimilitude and the Scientific Strategy of Economic Theory’, *Journal of Economic Methodology*, 6:331-50.

Zamora Bonilla, J. P., 1999b, ‘The Elementary Economics of Scientific Consensus’, *Theoria*, 14:461-88.

Zamora Bonilla, J. P., 2000, ‘Truthlikeness, Rationality and Scientific Method’, *Synthese*, 122:321-35.

Zamora Bonilla, J. P., 2002, ‘Economists: Truth-Seekers or Rent-Seekers?’, in U. Mäki (2002), pp. 356-75.

Zamora Bonilla, J. P., forthcoming a, ‘Science Studies and the Theory of Games’, *Perspectives on Science*.

Zamora Bonilla, J. P., forthcoming b, ‘Rhetoric, Induction, and the Free Speech Dilemma’, *Philosophy of Science*.

Ziman, J. M., 1968, *Public Knowledge: The Social Dimension of Science*, Cambridge, Cambridge University Press.

Ziman, J. M., 2002, ‘The Microeconomics of Academic Science’, in Mirowski and Sent (2002), pp. 318-40.