

**ANARCHISM, POSTMODERNISM
AND REALISM
UNDER CONFIRMATORY BIAS**

Juan Urrutia Elejalde
Universidad Carlos III
de Madrid

1. INTRODUCTION

If the Rhetoric of Economics is conceived as an alternative approach to the understanding of the workings of the so-called Economic Science then the present paper should not be seen as an essay within the Rhetoric of Economics. I have written in this vein in two occasions though. In 1982 I contributed to the celebrations of Keynes' centenary an avant la lettre rhetorical analysis of his policy ideas mentioning explicitly how their success was due to his clever using of the emerging culture of his days (see.....). In 1985 I wrote a paper, much later published as la Moda en Economia () using McCloskey's (83) and Klammer's (83) ideas in order to clarify the sense in which the conservative receepees for the crisis were not warranted by the good rhetoric underlying the rational expectations revolution but needed for its acceptance some additional condiments hardly acceptable as good rhetoric but rather related to political power.

The present paper takes a different tack. It is rather a comment on the Rhetoric of Economics, seen now as an alternative metatheory which is concerned with acceptance and rejection of economic theories and which emphasizes values and other

rhetorical means beyond objectivity. It is written from the stand point of a practitioner economist who tries to use current theories or models in selected fields in order to clarify what this special brand of metatheory is saying, without any pretensions of constructing a kind of metamethodology. It is the second time I write in this sort of middle ground. In 1998, and in a very sketchy way, I tried to call attention to the fact that the Reflexivity typical of Economics, and other Social Sciences, makes the distinction between Realism and Rhetoric more problematic than usually thought of. This unpublished (and unfinished) article could be considered as a late follow-up of the debate about Rhetoric in Economics that took place in Economics and Philosophy in 1988. In this article I tried to confront two papers by Mäki¹ from the point of view of somebody who was aware that, in Economics, reality could be constructed from expectations, something completely unthinkable in physics, say. What I brought to bear on the discussion was standard macrotheory in the contemporary New Classical vein which includes in an essential way information problems. I talked about bubbles in the stock exchange, the Phillips curve, the Lucas critic and sun-spots or self-fulfilling prophecies.

¹ I am referring here to Mäki (88) and to Mäki (96)

In the present paper I want to contribute to the more recent debate between Mäki and McCloskey in the Journal of Economic Literature (1995), somehow summarized in a later paper by Mäki (99) and followed by a kind of addendum by him forthcoming in the Journal of Economic Issues (see Mäki (996). I want to reflect upon Mäki's diagnosis of McCloskey and his proposals for her making sense and to do so, once again, from the stand-point of a practitioner economist. But now the economics I want to rely upon is microtheoretical and not completely standard in the sense that it will be influenced by biological and psychological ideas.. As will be seen this essay can then be taken as a second best exercise in an environment exhibiting less than perfect rationality. This combination seems appropriate for the study of Rhetoric and fits rather well the ontological constraint Mäki Calls www (way the world works)²

In the net section I summarize Mäki's reconstruction diagnosis of McCloskey as well as his proposals to make sense of McCloskey. The reconstruction is clear enough and the diagnosis claims that McCloskey is not an anarchist, not a postmodernist and not a realist. These two operations, reconstruction and diagnosis, together with the

² Since in these two pieces I use economic theory they can be seen not only as comments on the Rhetoric of Economics but also as articulating an essay within the Economics of Economics which can easily be extended to the Economics of Science in general or to the Economies of Art and Culture in particular.

proposals, will allow me to establish my purpose with some precision. In the third section of the paper, and in order to legitimize anarchism (or "anything goes"), I make use of a paper by Waldman (94) on natural selection in evolution to conclude that the moral constraints usually imposed on scientific conversation might have no bite, because the emergence of a systematic error makes them indiscernible. In the fourth section I present some psychological experiments reported by Rabin in JEL () which can give some ground for taking seriously confirmatory bias (an specific kind of systematic error) as a primitive concept, and then I report on its main implication: overconfidence. In this section and the following one I make extensive use of the recent paper of Rabin and Schrag (). In the fifth section I use some of their results on the implications of confirmatory bias (wrongness and no learning) to suggest that one can be a realist and yet entertain a coherence theory of truth instead of a correspondence theory of truth just because reality might not be reached. In the sixth section I recall some conditions under which confirmatory bias could lead to discount the opinion of elites and to count only on a kind of majority rule when considering problems of aggregation of experts' or scientists' information, therefore legitimizing postmodernism. In the last section I conclude and offer some additional comments.

2. MÄKI'S RECONSTRUCTION AND DIAGNOSIS OF McCLOSKEY

For my general purposes a summary of what is to be understood as Rhetoric seems convenient. Since McCloskey is a clear writer but not very precise it is not easy to pin down what she exactly means when pushing Rhetoric as a metatheory. Therefore a certain reconstruction is necessary. I take Mäki's reconstruction of McCloskey (see Mäki (95)) as my starting point because McCloskey seems to accept it at least approximately in her reply (see McCloskey (95))

Let us start then by the notion of rhetoric, being understood from the beginning that Rhetoric (with big R) is the study of rhetoric as applied to different fields. According to Mäki "*rhetoric is the use of arguments to persuade one's audience in an honest conversation...From this perspective, rhetoric is a social process which involves (i) a persuader (speaker, writer); (ii) a persuadee or an audience (listener, reader); (iii) the aim of the persuader to persuade the persuadee; (iv) argument as the means to attain the aim (and) (v) honest conversation as the social channel of persuasion*" (p.1303).

There is little doubt that every scientific discipline and every artistic or cultural activity uses rhetoric. Therefore Rhetoric as the study of rhetoric has a broad scope and a wide range of application. I'll focus here only on Economic Rhetoric, that is on the study rhetoric as it is used in Economics. This Rhetoric seem to me very interesting and important for two reasons. In the first place because one way (perhaps the only way) we have to select theories is by choosing the most plausible one, that is by choosing that theory which we find most coherent with our present set of beliefs, and secondly (and complementarily) because our beliefs are changed by (and perhaps only by) rhetoric. The point is that coherence and plausibility are open notions quite prone to be influenced by rhetorical devices. Consequently Rhetoric is closely related to a coherence theory of justification which according to Mäki (95) "suggests that all beliefs are justified by their relations to other beliefs with which they cohere".

Let us now turn to Mäki's reconstruction of McCloskey's theory of truth. Mäki writes: *the truth (with small t) of a statement consists of its coherence with a certain set of beliefs that a privileged set of humans, obeying the cannons of sprachetik end up*

with in an ongoing conversation... (p.1306). This is a coherence theory of truth which clashes brutally against any (intuitive) correspondence theory of truth. However McCloskey qualifies this coherence theory of truth with social and moral constraints giving rise to what he calls *an elite and angel theory of truth*: only utterances coming from "good economists" matter and, among those, attention should be paid only to the ones reached in the process of an "honest" conversation.

Given this reconstruction Mäki's diagnosis follows immediately at least in negative terms McCloskey is not *"an intellectual anarchist subscribing to a literal understanding of the credo 'anything goes' with respect to economics" (p.1311)* because certainly she proposes and tries to further moral and social constraints upon scientific conversation. Neither she is a *postmodernist philosopher* for whom *"all participants in any conversation are on equal footing" (p.1312)* because she relies on elite's opinions. Finally McCloskey is not a *"realist about truth" (p.1312)* even if she believes in the existence of the independent reality of the world.

Once reconstruction and diagnosis have been elaborated, Mäki indulges, quite naturally, several proposals quite attuned to his own way of looking at this matters. Mäki is a scientific realist (see Mäki 96) and therefore he dislikes a coherence theory of truth, but quite independently of this, and just to make sense of McCloskey's apparently arbitrary claim that Economics is in good shape, he proposes to stitch to a correspondence theory of truth and reserve angels and elites for the coherence theory of justification at most keeping them apart from any theory of truth and from Rhetoric it self. Let us look briefly at each of this proposals .

As far as Rhetoric is concened the proposal implies that anarchism and postmodernism cannot be excluded from Rhetoric something quite natural specially if me recognizes that it will be polish to let the absence of angels or elites to carry the absence of Rhetoric. I will have something to add to this proposal. As for keeping angels and elites away from the theory of truth seems quite obvious since they have nothing to do with truth when one entertains a correspondence theory of truth. Nothing to add here. Finally Mäki proposes that at most elites and angels could be included in the coherence theory of justification. I will comment on that in the signal.

My peruse propose in this paper can now be stated. I want to study the reconstruction , diagnosis and proposals I have just expounded in an environment (allowing second best analysis and bounded rationality) which exhibits confirmatory bias and seems to fit well the undertones of Rhetoric. In this environment I want to make three broad comments. In the first place I want to reinforce Mäki's proposal of admitting anarchism and postmodernism as part of Rhetoric . However this suggestion also makes reasonable a coherence theory of truth even if one is a realist; something that Mäki presumable wouldn't tike. In the second place I want to suggest that in the environment studied angels and elites are, surprisingly, compatible with anarchism and postmodernism casting doubts into Mäki's diagnosis of McCloskey. In the third place, the compatibility of anarchism, postmodernism and realism in the environment studied makes the use of angels and elites in the theory of justification something more doubt full than Mäki seems to admit. I will briefly express my doubts about the possibility that elites and angles could make the market of ideas run smoothly.

3. MORAL CONSTRAINTS (SPRACHETIK AND HERRSCHAFTFREIKEIT)

These two notions imply surely that scientific conversations are honest attempts at persuasion which exclude both lying about findings or personal characteristics and the formation of coalitions to exercise power in matters of justification. However, as we will now see these two traits might not exclude systematic errors when judging the merits of a particular theory or when trying to weigh one theory against another. The important point is that these systematic errors might look as dishonest or mafia-like behavior.

In the sequel Waldman's (94) ideas will be summarized showing that an evolution might lead to an equilibrium in which the above-mentioned systematic error will appear. Let us then assume that at a point in time t of the evolutionary process there are a continuum of scientists of unit mass and a continuum of scientific institutions of unit mass. This is a technical assumption that should not bother as here. Both scientists and institutions are characterized by a pair of parameters δ and γ , $\delta \in \{\delta_1, \dots, \delta_N\}$, $\gamma \in \{\gamma_1, \dots, \gamma_N\}$. Now, for a scientist $\delta \geq 0$ stands for the desutility of effort and, for an institution, it stands for the desutility of effort it inoculates in any of its members. Similarly we take γ to stand for the bias of

a scientist in evaluating its own performance and for the self-confidence an institution endows any of its members with. The number and / or relevance of the contributions of scientist i is given by $F_i = f_i(e_i, I_i)$ a function of the effort $s(he)$ chooses to make and of the quality of the institution $s(he)$ belongs to. Where quality is a continuous variable. However the subjective estimation of these contributions is given by $F_i^E = \gamma_i F_i$, where indeed γ_i is the bias or self-confidence. The utility of this scientist is given by $U_i = \gamma_i F_i - \delta_i e_i$. Given some standard characteristics of the function f , which make effort and quality of institution complementary, it is easy to understand that with any $\delta > 0$, the effort which maximizes utility is always smaller than the maximum possible effort \bar{e} .

We now turn to the exploration of natural selection and the We now turn our attention to the evolutionary process. Let us assume that at period t the data are the following. There is a proportion λ of scientists characterized by $(\hat{\delta}, \hat{\gamma})$ and a proportion $1 - \lambda$ characterized by (δ', γ') . Similarly there is a proportion λ of institution characterized by $(\hat{\delta}, \hat{\gamma})$ and a proportion $(1 - \lambda)$ characterized by (δ', γ') . At t there are random pairings of scientists and institutions and these pairings will determine the proportions of scientists characteristics in the

next generation. Let $k(\delta, \gamma)$ be the number of "off-springs" which has a scientist with characteristics (δ, γ) under the forces of natural selection. Note that the number of off-springs depends only on the characteristics of the scientist although the characteristics these off-springs might have depend also on the institution the scientist is paired with. Let $K = \lambda k(\hat{\delta}, \hat{\gamma}) + (1 - \lambda)k(\delta', \gamma')$. Then $\hat{\lambda} = \lambda k(\hat{\delta}, \hat{\gamma}) / K$ is the proportion of scientist having characteristics $(\hat{\delta}, \hat{\gamma})$ in $t + 1$. When a scientist $(\hat{\delta}, \hat{\gamma})$ "mates" with an institution (δ', γ') the inheritance yields $\hat{\lambda}/4$ scientists of each of the following pairs of characteristics : $(\hat{\delta}, \hat{\gamma}), (\hat{\delta}, \gamma'), (\delta', \hat{\gamma}), (\delta', \gamma')$.

We now turn to exploration of natural selection and the corresponding notion of equilibrium. We say that $(\hat{\delta}, \hat{\gamma})$ is a first best adaptation if and only if $k(\hat{\delta}, \hat{\gamma}) \geq k(\delta_i, \gamma_i)$, $\forall i, j$. Analogously $(\hat{\delta}, \hat{\gamma})$ is a second best adaptation if and only if (i) $k(\hat{\delta}, \hat{\gamma}) \geq k(\delta_i, \hat{\gamma}) \forall \delta_i \neq \hat{\delta}$ and (ii) $k(\hat{\delta}, \hat{\gamma}) \geq k(\hat{\delta}, \gamma_i) \forall \gamma_i \neq \hat{\gamma}$. It is very easy to find conditions under

which the first best adaptation is unique and implies $\hat{\delta} = 0$ and $\hat{\gamma} = 1$ with no discounting and no bias in the judgment of own merit. In this context the natural notion of equilibrium is Evolutionary Stable Strategy (ESS) which occurs when an strategy, say (δ, γ) is such that, if adopted by all members of the population, no mutant strategy could invade the population under natural selection.

In order to formalize this notion of equilibrium in our context let $(1 - \lambda)$ be the proportion of scientists and institutions with characteristics $(\hat{\delta}, \hat{\gamma})$ and let λ be the corresponding proportions with characteristics (δ', γ') . Then $(\hat{\delta}, \hat{\gamma})$ is ESS if $\exists \lambda$ small enough s. t. , given any (δ', γ') , the $(1 - \lambda)$ proportion of scientists characterized by $(\hat{\delta}, \hat{\gamma})$ goes to 1 as t goes to ∞ . The interesting point is that $(\hat{\delta} > 0, \hat{\gamma} > 1)$ can be a second best adaptation (Waldman

lemma 2) and that this second best adaptation can also be ESS provided $3k(\hat{\delta}, \hat{\gamma}) > k$

$(\delta_i, \gamma_i) \forall_{i,j}$ (Waldman Proposition 4).³

The very general intuition of these results (the formal proof of which is indeed in Waldman) is rather obvious. There are initial conditions in the form of population proportions in the space of characteristics which are reinforced by the forces of natural selection even when they are not the first best. As has been suggested the first best is given by a level of effort $e < \bar{e}$ and by absence of bias, $\gamma = 1$, which occurs when $\delta = 0$. As in any theory of second best if $\delta \neq 0$ the second best solution does not necessary involve $\gamma = 1$ but in general requests $\gamma \neq 1$.⁴

As far as Mäki's first proposal is concerned the point of the exercise is to suggest that the moral constraints of *Sprachetik* and *Herrschaftfreiheit* might not be binding in a

³ Note in passing that $\hat{\gamma} > 1$ means systematic error. The possibility of such systematicity in committing errors contradicts the underlying logic of the rational expectation hypothesis that is, systematic errors (in forecasting par example) can be detected and yet not to induce any revision.

⁴ The relationship between second best and evolution is interesting in its elf bus is also related to boulder rationality. Evolution is blind, it proceeds, for instance, ramdon pairings, and not by organizing them in any optimal way. And it is preusly because it proceeds like this that evolution may and generate first best selection. However the solutions it does generate are persistent if they are evolutionary stable (just as the QWERTY key board is persistent. But this persistence does not mean that strategies cannot change in extreme circumstances like uniting two different population of approximately the same size. These

very particular way. In the example examined the situation is free of herrschaft because the pairings are random and no coalitions are allowed to form for the mutual promotion of a particular scientific theory or approach. The situation is also such that Sprachetik is the rule since nobody hides its characteristics with some strategic goal in mind. However some systematic error or overvaluation of own merit might emerge through the forces of natural selection. If it does emerge we encounter a social situation which is informationally equivalent to a situation in which scientists strategically misrepresent their merits or connive with others to impose certain theories.

Under this informational equivalence is not completely foolish to entertain an anarchistic attitude. Therefore this possibility reinforces Mäki's proposal of admitting anarchism in Rhetoric, not because it is realist to do so but because even if it want's. He situation might look as such and there might be no way of make believe that what is going on is an honest and open discussion.

4. CONFIRMATORY BIAS AND OVERCONFIDENCE

apparently stable strategies can also change in expectedly when expectation have to be coordinated because there are multiple second but solutions, a case which is quite normal.

That there might be systematic biases in the agents perceptions might be taken as a theoretical curiosity. However psychological experiments reported by Rabin (98) seem to show quite clearly that in fact those biases are present in real life. Among these experiments the ones which seem most congenial to my purpose here are the ones referred to as confirmatory bias , a kind of bias which produces a certain belief perseverance.

Experiment 1 (Lord, Ross and Lepper).

They asked 151 undergraduates to complete a questionnaire that included three questions on capital punishment. Later, 48 of these students were recruited to participate in another experiment. Twenty-four of them were selected because their answers to the earlier questionnaire indicated that they were " 'proponents' who favored capital punishment, believed it to have a deterrent effect, and thought most of the relevant research supported their own beliefs. Twenty-four were opponents of capital punishment, doubted its deterrent effect and thought that the relevant research supported their views". These subjects were then asked to judge the merits of randomly selected studies on the deterrent efficacy of the death penalty, and to state whether a given study (along with criticisms of that study) provided evidence for or against the deterrence hypothesis. Subjects were then asked to rate, on 16 point scales ranging from -8 to +8, how the studies they had read moved their attitudes toward the death

penalty, and how they had changed their beliefs regarding its deterrent efficacy. At confidence levels of $p < .01$ or stronger, Lord, Ross, and Lepper found that proponents of the death penalty became on average more in favor of the death penalty and believed more in its deterrent efficacy, while opponents became even less in favor of the death penalty and believed even less in its deterrent efficacy. (Rabin, p.27)

Experiment 2 (Darley and Gross)

Seventy undergraduates were asked to assess a nine-year-old girl's academic skills in several different academic areas. Before completing this task, the students received information about the girl and her family and viewed a video tape of the girl playing in a playground. One group of subjects was given a fact sheet that described the girl's parents as college graduates who held white-collar jobs; these students viewed a video of the girl playing in what appeared to be a well-to-do suburban neighborhood. The other group of subjects was given a fact sheet that described the girl's parents as high school graduates who held blue-collar jobs; these students viewed a video of the same girl playing in what appeared to be an impoverished inner city neighborhood. Without being supplied any more information, half of each group of subjects was then asked to evaluate the girl's reading level, measured in terms of equivalent grade level. There was a small difference in the two groups' estimates -those subjects who had viewed the "inner-city" video rated the girl's skill level at an average of 3.90 (i.e., 9/10 through 3rd grade) while those who had viewed the "suburban" video rated the girl's

skill level at an average of 4.29. The remaining subjects in each group were shown a second video of the girl answering (with mixed success) a series of questions. Afterwards, they were asked to evaluate the girl's reading level. The inner-city video group rated the girl's skill level at an average of 3.71, significantly below the 3.90 estimate of the inner-city subjects who did not view the question-answer video. Meanwhile, the suburban video group rated the girl's skill level at an average of 4.67, significantly above the 4.29 estimate of the suburban subjects who did not view the second video. Even though the two groups viewed the identical question- and-answer video, the additional information further polarized their assessments of the girl's skill level. Darley and Gross (1983) interpret this result as evidence of confirmatory bias- subjects were influenced by the girl's background in their initial judgments, but their beliefs were evidently influenced even more strongly by the effect their initial hypotheses had on their interpretation of further evidence. (Rabin pp 27-28)

These two experiments show a psychological phenomenon called polarization which occurs according to the first impressions received by agents. This polarization might turn into confirmatory bias, especially well described by Lord, Ross and Lepper:

" With confirming evidence, we suspect that both lay and professional scientists rapidly reduce the complexity of the information and remember only a few well-chosen supportive impressions. With disconfirming evidence, they continue to reflect upon any

information that suggests less damaging "alternative interpretations". Indeed, they may even come to regard the ambiguities and conceptual flaws in the data opposing their hypotheses as somehow suggestive of the fundamental correctness of those hypotheses" (Rabin p.28).

It should also be reported that experience and learning do not necessarily eliminate confirmatory bias , it may even exacerbate the problem as several experiments reported by Rabin (98) show. This is particularly so among "experts who have rich models of the system in question". Rabin continues: "indeed many authors have hypothesized the role of reasoning process itself in exacerbating the confirmatory bias".

This confirmatory bias seem to be specially important under the following condition: (i) ambiguity of the evidence, (ii) abstract character of the situation, (iii) necessity of interpretation of evidence of situation and (iv) previous reasoning. These conditions are dearly present in the Social Sciences in general and in Economics in particular. They are all persuasive in art and culture but clearly absent in the natural sciences (or are they?). Where only the forth is aplicable. Note that this confirmatory bias does not violate sprachetik : it simply casts doubts on its efficacy to solve discrepancies. It is then interesting to use this caution when reading the article by Mäki,

forthcoming in the Journal of Economics Issues , in which he expresses his irritation, and the reasons for it, prompted by McCloskey's response in JEL to Mäki's Diagnosis in the same Journal.

Let us now formalize the consequences of confirmatory bias. In this section we focus attention on how it leads to overconfidence. In the next two sections we will try to understand its eventual consequences for the pursuit of truth (however defined).

At this point we just prepare the stage for further analysis. Let us focus then on a particular agent, a scientist say, and let x be a particular theory. This theory $x \in \{A, B\}$ and $x=A$ means " x is true" and $x=B$ means " x is false". For the moment it is not necessary to specify what kind of theory of truth we stick to. Our scientist has an a priori belief over the set $\{A, B\}$ which can be written as $\text{prob}(x=A) = \text{prob}(x=B) = 0.5$. At each t , $t = 1, 2, 3, \dots$, the scientist receives a signal s_t , independent and identically distributed (i.i.d), which is correlated with the true state of the world, A or B. This signal $s_t \in \{a, b\}$. Since we do not specify whether "true" means "coherent" with other theories or "faithful" to the real world the interpretation of S_t can go from the reaction to the presentation of x to an educated audience at a seminar to the result of an experiment specially designed to ascertain the correspondence of the content of x with some trait of reality. But now, the signal s_t is not perfect and its "correctedness" can be measured by $\vartheta = \text{prob}(s_t = a/A) = \text{prob}(s_t = b/B) \in (0.5, 1)$. Note that since $\vartheta \neq 1$ we are in a

context which can be called in very broad terms as one of second best. Furthermore at each t the scientist does not perceive s_t but rather $\sigma_t \in \{\alpha, \beta\}$ and this perception may misinterpret the signal more or less depending on the "severity" of the confirmatory bias. For simplicity let us assume that only those signals which conflict with a priori beliefs are misinterpreted and let us denote by q the probability of misreading a conflicting signal.

At each point in time the scientist updates his a priori beliefs according to the signals perceived. It is not very difficult to establish now a certain relationship between the "correctness" of the signal, ϑ , and the "severity" of the confirmatory bias, q , at stage t of the updating process. For no doing let $s^{t-1} = (s_1, s_2, \dots, s_{t-1})$ and let $\sigma^{t-1} = (\sigma_1, \sigma_2, \dots, \sigma_{t-1})$ and let us introduce two simple notions:

$$\begin{aligned} \vartheta^* &= \text{prob}(\sigma_t = \alpha / \text{prob}(x = A / \sigma^{t-1}) > 0.5, x = B) \\ &= \text{prob}(\sigma_t = B / \text{prob}(x = B / \sigma^{t-1}) > 0.5, x = A) \end{aligned}$$

i.e. the probability that the scientist misreads the evidence at t given that the previous evidence (may be also misread) has led him to entertain the wrong believe, and

$$\begin{aligned} \vartheta^{**} &= \text{prob}(\sigma_t = \alpha / \text{prob}(x = A / \sigma^{t-1}) > 0.5, x = A) \\ &= \text{prob}(\sigma_t = \beta / \text{prob}(x = B / \sigma^{t-1}) > 0.5, x = B) \end{aligned}$$

i.e. the probability that the scientist reads correctly the evidence at t given that the previous evidence (may be read correctly or may be misread) has led him to entertain the right belief. It is now very simple to establish that ⁵

$$\vartheta^* = (1 - \vartheta) + q\vartheta$$

$$\vartheta^{**} = \vartheta + q(1 - \vartheta)$$

which are two relationships between ϑ and q . If $q=0$ there is no confirmatory bias and $\vartheta^{**} = \vartheta$ and $\vartheta^* = 1 - \vartheta$. In this case the scientist has no confirmatory bias, s(he) is a perfect Bayesian statistician and, according to Harsanyi, perfectly rational. However if $q=1$, $\vartheta^* = \vartheta^{**} = 1$ the scientist beliefs are completely determined by the first signal s(he) perceives, rightly or wrongly. First impression matter a lot and beliefs cannot be changed by any posterior evidence conflicting with this first impression. We could say that when ever $q>0$ but $q \neq 1$ rationality is not perfect and this justifies my calling this paper an exercise in bounded rationality.

Let us then examine the intermediate cases. Suppose the scientist has received n_α α signals and n_β β signals, $n_\alpha > n_\beta$. S(he) thinks s(he) has received n_α α signals and n_β β signals. Therefore starting from the a priori belief $\text{prob}(X = A) = \text{prob}(X = B) = 0.5$ s(he) has correctly updated to

⁵ For example let θ^* be the probability of my misreading the signal α when B is the case but I take A to be the case because it is supported by previous evidence. It is given by the probability that the signal is correct, ϑ , times the probability that I misread it because it conflict with my previous belief, q , plus the probability of the signal being incorrect, $(1 - \vartheta)$ which I always believe because it agrees with my previous beliefs. θ^{**} can be interpreted analogously.

$$\text{prob}(x = A/n_\alpha, n_\beta) = \frac{\vartheta^{n_\alpha - n_\beta}}{\vartheta^{n_\alpha - n_\beta} + (1 - \vartheta)^{n_\alpha - n_\beta}}$$

and to

$$\text{prob}(x = B/n_\alpha, n_\beta) = \frac{(1 - \vartheta)^{n_\alpha - n_\beta}}{\vartheta^{n_\alpha - n_\beta} + (1 - \vartheta)^{n_\alpha - n_\beta}}$$

such that we can define

$$\Lambda(n_\alpha, n_\beta) = \frac{\text{prob}(x = A/n_\alpha, n_\beta)}{\text{prob}(x = B/n_\alpha, n_\beta)} = \frac{\vartheta^{n_\alpha - n_\beta}}{(1 - \vartheta)^{n_\alpha - n_\beta}} < > 1$$

Λ is a kind of likelihood ratio. If $\Lambda > 1$ the scientist thinks $x = A$ is more likely than $x = B$. If $\Lambda < 1$ the scientist thinks $x = B$ is more likely than $x = A$ in spite of the fact s(he) thinks s(he) has received more signals correlated with A than signals correlated with B. Note now that $\Lambda(n_\alpha - n_\beta)$ depends on $m = n_\alpha - n_\beta$ and that this number does not coincide with $m^* = n_a - n_b$. It does however when $q=0$. In this case the perfect Bayesian statistician will calculate the likelihood ratio as $\Lambda(n_a, n_b) = \Lambda^*(n_\alpha - n_\beta)$. The following result can now be established

Fact 1 (Overconfidence): If $n_\alpha > n_\beta$ and $n_\alpha + n_\beta > 1$, then

$$\Lambda(n_\alpha, n_\beta) < \Lambda(n_\alpha, n_\beta)$$

Proof: See Rabin and Schrag

Interpretation: Whatever is the correctness of the signal, ϑ , if there is confirmatory bias ($q > 0$) the scientist will be overconfident in his belief about which state of the world, A or B, is more likely. This or her belief in favor of $x=A$ (in favor of x being true) is stronger than is justified by the available evidence.

As mentioned before the persistence of beliefs which may occur under confirmatory bias might be underneath the apparently not very honest response of McCloskey to Mäki's Diagnosis. In fact Mäki's sequel in the Journal of Economic Issues can be read as accusing McCloskey of violating sprachetik. However if we think that confirmatory bias is a fact of we can not infer dishonesty in McCloskey response but perhaps some obvious overconfidence completely coherent with sprachetik.

In any case the groundwork elaborated in this section will yield its fruit in the next two, when confronting the issue of realism and elites.

5. ON REALISM

If we want to discuss McCloskey's claim of being a realist, and we want to discuss it in terms of truth theory we should move to a correspondence theory of truth. In this section we continue to make use of Rabin and Schrag (99) results in order to show that

confirmatory bias might lead to entertain false theories at least in probabilistic terms and to eliminate possibilities of learning.

Let us first discuss the possibility of our scientist being wrong. As we already know this scientist may think that the probability of $X=A$ is $\mu > 0.5$ when in fact it is smaller than μ . S(he) is overconfident but s(he) can be said to be right if the true probability of $X=A$ is greater than 0.5. If in fact this true probability is less than 0.5 s(he) can be said to be wrong. Can s(he) be wrong in this sense?. Yes if ϑ and q are sufficiently close to one.

Fact 2 (Wrongness): There are values of ϑ and q close to one which, given $n_\alpha - n_\beta \geq 2$,

$$\text{yield } \Lambda^*(n_\alpha, n_\beta; \vartheta q) < 1$$

Proof: see Rabin and Schrag

Interpretation: When confirmatory bias is very severe, $q \approx 1$, and the signal is very informative, $\vartheta \approx 1$, and the scientist has perceived one or none β signals s(he) is probably correct in believing $x = A$. However when s(he) has received two or more β signals s(he) is probably in correct in believing $x = A$ because s(he) does so only because. The confirmatory bias is very severe. Therefore having received two or more β signals is something very informative while receiving additional α signals is not very informative because, with $q \approx 1$, the very first α signal is the one that explains why the scientist believes $x = A$.

Let us now turn to the possibility of learning. One might hope that, even when having received a certain amount of evidence you are still wrong according to a correspondence theory of truth, the observation of many additional signals will dissipate your ignorance and you will end up learning the truth. Rabin and Schrag have shown that this might not be the case.

Let P_w be the probability that although the true state is $x=A$, the scientist comes to believe irreversibly with near certainty that $x = B$, starting from any belief. The following fact means that there is a positive probability that this may happen.

Fact 3 (No learning): If $q > 1 - \frac{1}{2\vartheta}$, then $P_w > 0$

Proof: See Rabin and Schrag

Interpretation. Despite receiving an infinite number of signals it may occur that the scientist becomes almost sure that $x = B$ when in fact $x = A$. This happens when the quantitative relationships between q and ϑ (recall $\vartheta \in (0.5, 1)$) is such that $\vartheta^* = (1 - \vartheta) + q\vartheta > 0.5$. When this is the case we already know that, once the scientist has come to believe $x=A$, he is more likely to perceive confirmatory evidence in favor of his or her incorrect beliefs. Than to perceive signals which conflict with this incorrect belief.

These two last facts are direct consequences of the confirmatory bias a trait of the psychological personality which could very well be an evolutionary stable second

best adaptation under the force of natural selection . Be as it may these results about wrongers and about learning problematize once more the issue of realism. Now, it is not that expectations might in fact construct reality, something I have discussed in my 1998 already mentioned article, but rather that the correspondence theory of truth is at stake in a certain sense. One can have a realist conception of the world in the sense of believing in a world outside language (as McCloskey confesses she really does: "I'm a realist "answers to Mäki) and one can even, and correspondingly, like to have a correspondence theory of truth (as Mäki advises McCloskey to have) and yet face the impossibility of entertaining such a theory of truth just because the outside world may be completely inaccessible, a possibility that the two results above appear to sustain.

Under these circumstances we might easily say that "correspondence" is an interesting notion that however has to be considered inoperative. Then why should one not hold a coherence theory of truth?. For one thing it is accessible precisely because confirmatory bias has no bite against it. One might even use a more forceful argument, namely that coherence is the best strategy one may follow in an open society so as to discover the real underlying world. This argument deserves scrutiny but it has to wait to another occasion. However some additional comments will be offered in the last section. In any case it should be clear that this argument is not immune to confirmatory bias.

6. ON ELITES

In order to complete my aim in this article I have to show that the social constraints represented by elites might not be binding in scientific conversations. For so doing I continue to explore the consequences of confirmatory bias now on how to aggregate the information provided by experts each of which judges whether $x = A$ or $x = B$, i.e. whether theory x is true or false (without restricting these notions to any correspondence with the underlying reality). Mäki notices at this respect that McCloskey puts more strength on the opinions of members of a scientific elite which in his case could be the Chicago type of economist or, more in general, the one belonging to the neoclassical tradition. However, as we will see presently, it can be shown that, under confirmatory bias, it is quite reasonable to pay attention to the majority of scientists or experts without weighing their opinions according to belong-ness to any particular elite.

Just to be more precise, and because we are interested in Rhetoric as metatheory, let us think of a Principal (be it the metatheorist or any scientific policy board) who must collect information regarding some particular theory from a set of scientist who may be considered his Agents. Because these agents are subject to confirmatory bias the optimal contract between Principal and Agents must not only take into account the usual

incentive compatible constraint but must also "prevent decisions based on good faith overconfidence "(Rabin and Schrag (99) p.63) In particular incentives that increase the collection of a lot of information might not be desirable.

Suppose the Principal tries to allocate a given amount of money, $W = 1$, among three different theories, T_A , T_B and T_C . Theory T_C is conventional wisdom and we can say that the "returns" of "investing" W on it are given by $r(T_C) = 1$ with complete certainty. However the other two theories are "risky" and their "returns" depend on the state of nature $x \in \{A, B\}$ in the following way: $r(T_A / A) = r(T_B / B) \in (1, 2)$ and $r(T_A / B) = r(T_B / A) = 0$. The Principal has a conventional Von-Neuman-Morgenstern utility exhibiting risk aversion and considers the possibility of asking an expert or a scientist about the most convenient allocation of $W = 1$ among the three theories. As before Principal and Agent share the common prior belief that $\text{prob}(X=A) = \text{prob}(X=B) = 0.5$. It is clear that if the Agent has no further information he will recommend T_C ; but if $s(\text{he})$ gathers additional information that changes his or her prior belief about the state of nature $s(\text{he})$ will recommend one of the risky theories.

The problem is how to incentivize the Agent. To be concrete suppose that the Agent has no cost of gathering information and that $s(\text{he})$ is completely risk averse. In this conditions the Principal has to pay the Agent a fixed amount and, however small, if it is positive the agent will gather an infinite number of observations. Therefore we only have two possibilities. Either the Principal pays no incentive and then, as mentioned, T_C is selected. Or the Principal pays a

certain amount and the agent gather lots of information. In this latter case we now from Fact 3 that the Agent might be wrong even after gathering all the new observations. It has been shown by Rabin and Schrag that the agent, after information gathering, will identify the true state of the world with probability:

$$\mu^*(\vartheta, q) = \frac{\vartheta[2(\vartheta + q(1 - \vartheta) - 1)](1 - \vartheta + q\vartheta)}{q[1 - 2(1 - q)\vartheta(1 - \vartheta)]}$$

which is increasing in ϑ and decreasing in q . Therefore the Principal's expected pay-off is

$$Eu^* = \mu^*(\vartheta, q)u(R) + (1 - \mu^*(\vartheta, q))u(o)$$

and s (he) does not want the Agent to gather information when $u(1) \geq Eu^*$. Define now $\underline{\mu}$ as

the value of μ^* satisfying $u(1) = \underline{\mu}u(R) + (1 - \underline{\mu})u(o)$. Since $\mu^*(\vartheta, q) \geq 0$ the Principal

offers an incentive for the Agent to get informed if and only if $\vartheta \geq \underline{\mu}$. That is there is always a

high enough q and a small enough ϑ which make not paying any incentive the optimal choice.

The intuition is quite obvious, in any of these two cases the probability of being wrong is very high.

Given this intuition it should also be intuitive that if the Principal has to allocate a given amount of signals he wants to gather among several agents the more Agents the better, since each agent will be less likely to be carried away by his or her confirmatory bias. The next step is to think about how to weigh the information of each agent once the number of agents has been decided but the number of signals gathered by each agent is

unknown . Rabin and Schrag (99) give the following example. Suppose the correctness of the signal is given by $\vartheta = 0.6$. There are three Agents. Suppose first that they were perfect Bayesian statisticians . Two of three agents report believing in T_A with probability 0.6 (meaning that each of them has received 2 more times signals a than signals b) and one agent reports believing in T_B with probability 0.77 (meaning that he has received three more b signals than a signals). The Principal in this case should believe in T_B with probability 0.6. That is s(he) should pay attention to the "strength" of the agents beliefs perhaps weighing more heavily the believes of agents who are more experienced or belong to any kind of elite. What if the Principal knows that the agents are subject to confirmatory bias? In this case Rabin and Schrag discuss the case and come up with the following statement *"If confirmatory bias is no severe that only an agents first signal is very informative, then the Principal may wish to discount the strength of agents' beliefs and basically aggregate according to a majority rules criterion."*

It seems quite safe to say that under the conditions of this section the number of scientist consulted is the greeter the better and that there are no elites the opinion of which is more heavily weighted than that of non members of the elite. Scientific conversations might not be socially constrained and hence postmodernism might not be

excluded from Rhetoric since all participants in the conversation might be on an equal footing. This reinforces Mäki's first proposal⁶

⁶ In this sense it is worth mentioning that Klammer's Conversation with Economists is rather postmodernist in the sense that he gives equal voice to all sorts of economists. However postmodernism is a larger issue that would deserve additional distinctions between relativism and pluralism for instance. The distinction between dialogue and performance made by Mäki (99) has also bearing on the issue.

7. CONCLUSIONS AND FINAL COMMENTS

This paper has articulated a second best analysis of a situation quite far from perfect rationality . Confirmatory bias can be thought of as a second best adaptation to the forces of natural selection and can also be an evolutionary stable strategy so that it is here to stay as seems to be supported by several psychological experiments. But once confirmatory bias is at work it is quite clear that economic agents in general or scientists in particular do not act as perfectly rational in the sense that they do not mimic the behavior of a Bayesian statistician. Combining second best theory and not perfect rationality is not staple stock in Economics. However it seems appropriate to the field of Rhetoric. Quite intuitively it does not appear to be any room for rhetoric when rationality is perfect and first best is attainable. Persuasion should be automatic and directed in such a world.

This kind of analysis has yielded three main results. First honesty and open, power-free, conversations may not preclude systematic error in appreciation of theories. Therefore the moral constraint supposedly operating on the opinions of scientists might not be binding in the sense that their opinions might look completely anarchistic. Second the social constraint might also not be binding because each scientist opinion carries the same weight regardless of fame or honor, a very postmodern situation. Third,

one can be a supporter of the correspondence theory of truth, one can have no doubts about the existence of an independent underlying real world and yet one has to accept that an honest and informed conversation may lead to the acceptance of false theories.

These three results obtained in a certain environment sustain the three broad comments I advanced before. First, they given Mäki additional reasons to seclude angels and elites from Rhetoric although these reasons also support a coherence theory of truth something presumably Mäki will not like. Second the compatibility of angels and elites with anarchism, postmodernism (and of course realism). Can yield an alternative diagnosis of McCloskey (more anarchism and postmodernism than Mäki would concede) more attuned to her proclaimed realism. Third the compatibility of anarchism, postmodernism and realism among themselves makes room from wrongnes and absence of learning and these in terms last doubts on the necessity of maintaining angels and elites in the coherence theory of justification.⁷

These conclusions may appear as sustaining a rather severe semantic pessimism (or at least skepticism) in Economics. With some final comments about this I close this essay. Semantic matters would look even gloomy if we add strategic considerations in

⁷ This compatibility may deprive Mäki of any consolation to his irritation about two of McCloskey's presumed inconsistencies . May be McCloskey when not really answering Mäki is not being dishonest, she might be subject to a very simple and common confirmatory bias. When McCloskey indulges in an unwarranted defense of neoclassical or Chicago Economics she might be misreading the evidence.

relation to the behavior of scientists. My two colleagues Zamora and Garcia-Bermejo have written on this. Zamora (99) has described how consensus might be reached among scientists and how this consensus might be objective in the sense that scientist will not misrepresent opinions but will tell the truth. However, additional considerations of information cascades or rumors might lead to clusters of opinion completely unrelated to truth as correspondence of opinions with underlying reality. Garcia Bermejo (99) on the other hand has been exploring the possibility of implementing in Nash equilibrium some "epistemologically interesting" situation without clearest results. What is indeed needed is a strategy proof mechanism that has as its equilibrium a semantically acceptable situation. However this is easier said than done.

Even if we assume away strategic considerations we have to confront the issue of whether the "market for ideas" might lead to the discovery of truth under confirmatory bias. The point is, however, that even if the severity of the problem were nil and even if moral constraints would underlay the market and social constraints would facilitate its functioning, the market for ideas might not work because increasing returns to scale are very common when ideas are involved. Furthermore if we took into account strategic behavior incentive compatibility and efficiency would require perfect competition understood as Makowski and Ostroy (99) understand it, that is requiring

full appropriation something almost incompatible with the well known inappropriability of ideas. If in addition we include now confirmatory bias it appears quite impossible to rely on the "market" as a way of reaching truth, given wrongness for instance.

In fact strategic behavior and confirmatory bias together would pose serious puzzles. Consider the following. Confirmatory bias might lead to every scientist stiding to its own theory and we might not discover any majority. Since majority is the why way we have to justify our beliefs under confirmatory bias we may want to penalize nengalenty somehow. But this incentive scheme might generate estrange dynamics specially if information flows according to rumors or cascades. (see Benerjee (93) and Bikhchanderi, Hisshlefer y Welch (92). There are chances that clusters of false opinions might dictate what is true.

In conclusion, one cannot be very optimistic about the possibility that a coherence theory of truth, even including elites and angles, could be a good strategy to attain truth in a correspondence sense. The promising strategy is to study alternative imperfect social decision arrangements en for their relative characteristics in relation to errors of acceptance of take theories or to errors in the rejection of true theories. What I think is worth exploring is the literature on the architecture of imperfect economic

systems initiated by Sah and Stightz (88) under the possibility of confirmatory bias. But this exploration will have to wait to some other occasion.

REFERENCES

- Banerjee, A (93): "The Economics of Prumours" , Review of Economic Studies, 60, pp. 309-327.
- Bikhchandani, S.,D.Hirshleifer and I. Welch (92): A Theory of Fads, Fashim, Lustom, and Cultural Change as Information Cascadas", Journal of Political Economy, 100, pp. 992 - 1026.
- Garcia Bermejo, J.C. (98): " Compatibilidad de incentivos en la investigación científica: una condición mínima"
- Klamer, A (83) Conversations with Economists. Rowman, Littelefield, Maryland
- Mäki, U. (88) : " How to combine Rhetoric and Ralism in the Methodology of Economics", Economics and Philosophy, 4, pp 89-109.
- (95) : "Diagnosing McCloskey" , Journal of Economic Literature, vol.XXXIII, pp. 1300-1318
 - (96) : "Scientific Realism and some Peculiaritis of Economics" in R.S. Cohen, R. Hilpinen and Qiu Renzons (eds.) Realism and Anti-Realism in the Philosoplay of Science, pp. 427-447
 - (98) : "The way the world works (www).Towards an ontology of theory of choice", unpublished.
 - (99 a) : "Representation Repressed : Two Types of Semantic Scepticism in Economics", in R. Rossini Fauretti, G. Sandri and R. Scazzieri. Incommensurability and Translation, Ruhnian Perspectives on Scientific Communication and Theory Change, pp. 307-321
 - (99 b): " Performance against dialogue, or answering and really answering:A participant observer's reflections on the McCloskey conversation", forthcoming in the Journal of Economic Issues
- Makowski, L y J. Ostroy (98): Perfect competition and the creativity of the Market", forthcoming in the Journal of Economic Literature.
- McCloskey,D (83): "The Rhetoric of Economics", Journal of Economic Literature 21(2) pp. 481-57
- (95) " Modern Epistemology against Analytic Philosophy: a Reply to Mäki", Journal of Economics Literature, vol 33, p.1319-23

Rabin, M. (98): "Psychology and Economics", Journal of Economic Literature, vol.36, pp. 11-46.

- and J.L. Schrag (99): "First Impressions Matter: A Model of Confirmatory Biase", Quarterly Journal of Economics, pp.37-81

Sah, R.k y J.E. Stiglitz (1988) : " Commiteas, Hierarchies and Polyarchies", Economic Journal, pp....

Urrutia, J. (9e): " La Moda en Economia" in "Perspectivas actuales de lógica y filosofía de la ciencia"

- (98): Realism, Rhetoric and Reflexitivity in Economics (no publicado)
- y R. Botas (82) : ¿ Necesitamos otro keynes"

Waldman, M (94) : Systematic Errors and the theory of Natural Selection, American Economics Review, vol. 84, pp. 482-97.

Zamora, J. (99) : " The Elementary Economics of Scientific Conversus", Theory, 14,pp 461-488.