UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Alberto Battistini

From Asymmetric Information to Social Knowledge: A Game Theoretic Example of Strategic vs. Bayesian Beliefs' Updating

n. 630 – Dicembre 2011
Revised, December 2013

Abstract: this paper provides a counterexample to a famous theorem of Aumann (1976) which states that common priors and common knowledge of the posteriors imply that the latter must be identical. This theorem, also known as an 'agreement theorem' after the title of the original paper, is based on the so-called 'Harsanyi doctrine', that is, on the idea that different probability assessments can only be the result of differential information. In its turn, the theorem is crucial to the epistemic conditions for Nash equilibrium, since common priors and common knowledge of the conjectures essentially mean that players already agree on how the game will be played.
Consequently, the argument is simply that when knowledge is about a conflictual phenomenon, disagreement is not only possible but also rational, though not in a Bayesian sense. More specifically, since the point is made with a game-theoretic example of the evolution of the institutional structure of production, the rationality of such disagreements is relative to the proposed solution concept, which in addition to the absence of incentives to unilateral changes of strategies requires the absence of tendencies to change the rules of the game. In particular, then, the paper shows that, when the assumptions that allow the commitment to methodological individualism are dropped, assessing the conditional probability of events taking account of the likely outcomes, that is, updating beliefs strategically rather than by a fatalistic application of the Bayes' law, is not wishful thinking but a quite logical consequence of a cognitive conflict which arises from an underlying real conflict and may lead to efficient and egalitarian institutional changes.
The basic shift in the focus of the analysis, in other words, consists of passing from the problem of asymmetric information about an observer-independent phenomenon to the problem of a different interpretation of the same observer-dependent phenomenon, a shift that in the domain of cognition mirrors the shift from Pareto-efficient exchanges among given individuals to structurally inefficient conflicts between groups in the reality domain. The contribution of the paper, in this sense, is the attempt to model the process through which agreement, rather than being ideologically assumed from the outset, turns out to be socially constructed by interested parties.

Keywords: *common knowledge, Harsanyi doctrine, mixed strategies; economic evolution, learning, institutional change; Bayesan rationality, minimax maximin, structural interdependence; methodological individualism, multi-level selection, social sciences methodology.*

JEL: *A12; C70; L10; L20*

Alberto Battistini, Dipartimento di Economia Politica e Statistica, Università di Siena
e-mail:alberto.battistini@unisi.it

1. Introduction

This paper provides a counterexample to a famous theorem of Aumann (1976) which states that common priors and common knowledge of the posteriors imply that the latter must be identical. This theorem, also known as an 'agreement' theorem after the title of the original paper, is based on the so-called 'Harsanyi doctrine', that is, on the idea that different probability assessments can only be due to differential information (Harsanyi, 1967; 1968). In its turn, that theorem is crucial to the epistemic condition for a Nash equilibrium, or to what the theorist should assume about what players know about the games they play in order that the expectation that they will play a Nash equilibrium be reasonable. Indeed, such epistemic conditions - common priors and common knowledge of the conjectures - essentially mean that players already agree on how the game will be played (Aumann and Brandeburger, 1995; Gintis, 2009, ch.8).

Consequently, the argument is simply that when knowledge is about a conflictual phenomenon, disagreement is not only possible but also rational, though not in a Bayesian sense. More specifically, the point is made with a game-theoretic example of the evolution of the institutional structure of production in the Coasean definition of firms and markets (Coase, 1991) and discussed in detail in a companion paper (Battistini, 2013). Both that paper and the present one are based on two critical assumptions. The first is a team-production-based theory of the firm, reframed in terms of the collective undertaking of non additively separable human capital investments, but stripped of the usual assumptions of no wealth effects and generalized opportunism.[1] The second assumption is that evolution of the institutional structure of production occurs within the framework of a multi-level selection process which, from the point of view of the players themselves, takes place sequentially in the event time.[2]

---

[1] As is well known, these two assumptions are the basis of the more famous interpretation of team production as costly metering (Alchian and Demsetz, 1972). As argued in the companion paper, however, such an interpretation actually amounts to a sterilization of non additive separability, as shown by the fact that, in their solution to problem of the nature of the firm, Alchian and Demsetz attribute to the entrepreneur exactly the function of separating individual contributions, thus remaining within the framework of Pareto-efficient exchanges among given individuals. By contrast, the focus on human capital investments is instrumental to a view of non additive separability in terms of a transformation of the workers' productive capacities, establishing a link with the classical issues of value and distribution. The collective element, on the other hand, emphasizes an unavoidable link among participants in the production process, so as to highlight, rather than covering, the related issues of power and conflict.

[2] Multi-level selection models have been introduced into economics by Bowles (2004). As in related models in biology (Sober and Wilson, 1998) and anthropology (Einrich, 2004), the focus of the analysis is on the conditions under which group selection is so strong as to

Because of the first assumption, groups – which can be defined as consisting of the individuals who belong to them, of the relationships between them (intra-group relationships), and of the relationships between them and the other groups (inter-group relationships) – are inserted as an intermediate layer of analysis between the individuals and the overall economy or, more precisely, are taken as the unit of analysis. Individual players thus find themselves simultaneously playing two games, one between the groups, which determines inter-group relationships or the division of labor between firms, and the other within the group, which determine intra-group relationships, or the division of labor within the firm.[3]

Because of the second assumption, such two games are linked by a relation of recursivity, in the sense that the equilibria of one game determine the rules of the other and vice versa. The vehicle of such two-ways effect between equilibria and rules of the games, in its turn, is identified with the evolution across generations of the beliefs system, by which is meant a common understanding of the relationship between action and pay-offs. Social reality and social knowledge thus mutually determine each other in a co-evolutionary process driven by the value-maximizing mechanisms implied by the sequential operation of the multi-level selection process.

Taken together, these two assumptions imply a demise of methodological individualism which extends from the domain of reality to that of cognition. On the one hand, indeed, the phenomena of interest – value creation and distribution – are not be understood as a sum of separate, if interdependent, individual contributions, so that the conflict between individual and collective rationality that defines a standard collective action problem ceases to be the only possible problem. On the other hand, knowledge about such phenomena is not already available in a self-evident and unequivocal form, so that differential information about an observer-independent phenomenon ceases to be the only

---

overcome individual selection, thus favoring the evolutionary success of altruistic traits over selfish ones. In other words, the idea is that more cooperative groups may prevail over less cooperative groups because, under such conditions, they are better able to internalize individually-costly-but-group-beneficial traits. Besides pointing out that, with team production and wealth effects, the problem of cooperation is not necessarily a prisoners' dilemma so that competition between groups may also give rise to conflicts or 'bad cooperation', the reason that I treat the operation of multi-level selection also in sequential terms is an attempt to conceptualize the process of beliefs formation and updating in terms of the co-evolution between social reality and social knowledge. See also footnote 5.

[3]More precisely, inter-group relationships are defined as the institutional rules that govern the process through which groups form, interact with each other, and become extinct, so that they amount to the competitive structure of markets and determine the structure of property rights in the next generation. Intra-group relationships, on the other hand, are the institutional rules governing the division of labor and the distribution of the surplus, so that they amount to the organizational and distributional structure of the firm.

possible problem. On the contrary, in the first case unanimity is no longer guaranteed since, when non additive separability combines with wealth effects, there may be structurally inefficient conflicts between groups in the sense that a Pareto-efficient equilibrium within the group may lock the system in a pay-off dominated equilibrium in the game between groups, thus preventing total value maximization. In the second case, instead, unanimity is no longer guaranteed because the very existence of such conflicts implies the possibility of different interpretations of the same (observer-dependent) phenomenon.[4]

In this situation, therefore, assessing the conditional probability of events taking account of the likely outcomes, that is, updating beliefs strategically rather than by a fatalistic application of the Bayes rule, is not wishful thinking but a quite logical consequence of the cognitive conflict which arises from the underlying real conflict and may lead to efficient and egalitarian institutional changes. Naturally, since the point is made in game-theoretic terms, the rationality of such a strategic updating is relative to the solution concept proposed, which is called 'multi-level equilibrium' and, as the name suggests, refers to the 'super-game' consisting of the two games recursively played between and within the groups. As a consequence, in addition to the absence of unilateral incentives to change strategies, such a solution concept also requires the absence of tendencies to change the rules of the games. Then, to the extent that this updating process leads to a new beliefs system – technically, the intersection between the commonly known, but different, posteriors on which players agree to disagree – on its basis different groups will form, implying that different rules of the game and different equilibria will be obtained. Accordingly, the example proposed can also be seen as a first step toward modeling a form of institutional change which can be understood as part of an equilibrium refinement process eliminating the pairs of equilibria which fail to maximize total value or, in terminology familiar to game-theorists, which are not perfect in the super-game or not robust to changes in the rules of the game.

---

[4] Observer-dependence and independence are clearly related to the distinction between subjectivity and objectivity, a distinction famously at the core of Western thought and equally famously a matter of degree. Following Searle (2005), however, additional insights may be gained when such a distinction is combined with that between ontology and epistemology, where the former term refers to the nature of the phenomena and the latter to the statements about them. It thus becomes clear that, while natural and social sciences are respectively concerned with ontologically objective and ontologically subjective phenomena, both must consist of epistemologically objective statements, that is, statements whose degree of truth does not depend on the individuals who make them. The complication, however, is that for the latter it is often the case that explanation of the phenomena of interest must include an explanation of how the epistemologically subjective beliefs of the individuals who give rise to them end up by being inter-subjectively shared.

2. Multi-level equilibrium at t=1: an example

More formally, let $\Gamma = \left\{\gamma_b^t, \gamma_w^t\right\}$ be the super-game, where $\gamma_b^t = \left\{g, \sigma_g, \Pi_g / \beta^t\right\}$ and $\gamma_w^t = \left\{i, \sigma_i, \Pi_i / \beta^t\right\}$ are respectively the game between groups and the game within the group at time *t*, with $t = 0,1$ indicating event time, that is, the three periods, each of them associated with a different generation, in which the beliefs' system may change.[5]

As usual, omitting the time index for simplicity, $g = 1,...,n/m$ ($i = 1,...,n$), σ_g (σ_i), and Π_g (Π_i), respectively indicate players, strategies and pay-offs, while *m* is the group size. In their simplest form – that is, with four players, two groups and two players per group in each generation (*n*=4 and *m*=2) – at *t=0* the two games can be represented as in fig. 1.

A \ B	UNREG.	REG.
UNREG.	D, D	Y, X
REG.	X, Y	B, B

Fig. 1a.

1 \ 2	HIER.	EG.
HIER.	a, c	0, 0
EG.	0, 0	b, b

Fig. 1b.

[5] In effect, the notion of recursivity, implying that the equilibria of one game determine the rules of the other through the effect on the beliefs system, which in its turn determine players' types and groups, delineates a learning process by which players pass from deduction (or from the 'general' to the 'particular', that is, from inter-group relationships to intra-group relationships) to induction (or from the 'particular' to the 'general', that is, from intra-group relationships to inter-group relationships). This kind of logical sequence in the learning process can be fruitfully linked to the institutional and evolutionary view of the process of capitalistic development in terms of recurrent cycles. The real problem which has motivated the formulation of the present example, indeed, is the life-cycle of a technological system whose introduction is taken as exogenous (i.e., not due to economic factors), whereas its successive development until exhaustion depends on a mutually determining interdependence with institutions, as it happens in that tradition. Though obviously less rich than historically minded analysis, a game-theoretic approach provides the opportunity of modelling the process through which this interaction between technology and institutions is accompanied, both as a cause and an effect, by the evolution of beliefs system. See Maddison (1982), Mokyr (2002), and Perez (2002).

In the game between groups (fig. 1a), the strategies available to the two players (group A and group B, or firm A and firm B) are regulated or unregulated inter-group relationships. Recalling the definition given in the preceding section, the difference between the two strategies is whether or not groups form, interact with each other, and become extinct through some form of (explicit or implicit) coordination.

In the game within the group (fig.1b), the strategies available to the four players (individuals 1 and 2, and 1' and 2', or entrepreneurs 1 and 1', and workers 2 and 2') are hierarchical or egalitarian intra-group relationships and, recalling the definition given in the preceding section, the difference between the two strategies is whether the division of labor and the distribution of the surplus within the firm are determined, respectively, by authority or consensus, and by appropriating or sharing.

The pay-offs in γ_b, with $B > D$, and $X < Y \leq D$ or $Y < X \leq D$ depending on the dominant technological paradigm, represent the profit associated with the two just mentioned strategies. As shown in the companion paper, this profit arises from realizing the surplus-value created by the collective undertaking of non-additively separable human capital investment and is measured by the difference between the cost of the investments and the cost of the next best type of investments (the equilibrium price or opportunity cost). Remarkably, as also shown in the just quoted paper, this kind of profit is not eliminated by free-entry and competition, the intuition being that, since surplus-value does not arise from some form of scarcity but from the production process, opponents or potential entrants have no incentives to undercut because they can earn the same profit by doing exactly the same thing.

Though the exact relationships between the two games have yet to be specified, the pay-offs in γ_w, with $a > b > c = \overline{w}$, $2b=B$, or X, and $a+c=D$, or Y, reflect the associated theory of distribution, with $\overline{w}$ indicating the amount that workers can earn by independent participation in the production process or, again, the opportunity cost or equilibrium wage. An important condition for what follows is that, because of non additive separability, joint production is always at least as profitable as independent production, so that $Min\{X, Y\} \geq \hat{w} + \overline{w} > 2c$, where $\hat{w} > \overline{w}$ represents the (higher) amount entrepreneurs can command by independent participation to the production process and is assumed to be higher than $Min\{X/2, Y/2\}$ but lower than b. The off diagonal pay-offs, then, are set to zero because the mismatch between strategies is assumed to 'eat' profit.[6]

[6] Basically, the reasoning behind such pay-offs determination is as follows: technological conditions, of which the type of investments are a stylized representation, determine inter-group relationships by minimizing the organizational costs of the division of labor between firm,

In sum, being different from both the conflict between individual and collective rationality defining a standard collective action problem and the mere distributional conflict typically considered 'political' rather than 'economic', the structural conflict between groups considered here is not to be represented by a Prisoner Dilemma or a zero-sum game but by two coordination games, one symmetric with only one pay-off dominant Nash equilibrium and one asymmetric with two Pareto-efficient Nash equilibria (all in pure strategies, as will be clearer in a moment).[7]

Finally, β^t is the beliefs system at time *t*, now defined, as is usual in epistemic game theory, as a common probability distribution over players' types within the group, conventionally called (*Upper* and *lower*) or, both, (*same same but different*). Remarkably, because of the advanced interpretation of team production, in the present context such types are relative rather than absolute, in the sense that they do not exist in isolation but only acquire meaning in the relation of one to the other. Consequently, the problem is not that players know their type but not that of their opponents; rather, it is that there may be different views about which type they are with respect to each other. Accordingly, there is no reason to play mixed strategies, and indeed players do not use them.[8]

---

subject to an individual participation constraint ensuring that investments are actually undertaken. Such participation constraint, which is in some sense the organizational counterpart of a technological feasibility constraint, can be satisfied at minimum costs either when expressed in terms of amount workers can earn by independent participation to the production process, or in terms of the value of the product. The first case, which is the normal case, occurs when wealth effects are binding in sense that, while entrepreneurs can pay workers that amount, the converse is not true. The second case, which is by far much rarer, occurs when wealth effects are not binding since the net value of a (non additively separable) investment independently realized is lower than the amount workers can earn by directly and independently participating to the production process. Consequently, in the former case, profit is appropriated by entrepreneurs, property rights are correspondingly concentrated, and the subsequent development of the technological paradigm is inhibited. In the latter case, by contrast, profit is shared, property rights are correspondingly distributed and the subsequent development of the technological paradigm is favoured.

[7] To familiarize himself with the games of the text, the reader may find it useful to use the following numbers, which refer to the benchmark case of binding wealth effects: B=5; D=Y=4; X=3; a=3; c= $\overline{w} = 1$; b=2.5; $\hat{w} = 2$.

[8] As is well known, the most accepted justification for the existence of a mixed strategies equilibrium is that such strategies –though never really played- represent the uncertainty of the *other players*, so that in fact they should be the opponents' conjectures about the player's choices (see Aumann and Brandenburger, 1995, and Gintis, 2009, ch. 6). Notably, even the more convincing evolutionary interpretation as a proportion of a given population playing a given strategy does not apply in the present context because the strategic structure of the problem is such that it is in the interest of each as well as of all the players of a given type (row

As a result, as the following definition makes clear, since γ_w is asymmetric, it is possible to distinguish between an equilibrium beliefs system from a disequilibrium beliefs system depending on whether or not they induce a pure strategy equilibrium within the group. Hence we have:

DEFINITION 1

(1.a) $\beta_x^* = \left\{\min \beta / \Pi_i(\sigma_i^x, \sigma_{-i}^x / \beta_x^*) \geq \Pi_i(\sigma_i^y, \sigma_{-i}^x / \beta_x^*)\right\}$ for every player *i* and at least one strategy *x*.

(1.b) $\beta_x^{\#} = \left\{\min \beta / \Pi_i(\sigma_i^x, \sigma_{-i}^x / \beta_x^{\#}) < \Pi_i(\sigma_i^y, \sigma_{-i}^x / \beta_x^{\#})\right\}$ for every strategy *x* and at least one player *i*.

For example, in the game above, $\beta_H^* \geq \frac{b}{b+c}$ and $\beta_E^* \leq \frac{b}{a+b}$, confirming the intuition by which, if the two players are almost sure that they are (*Upper* and *lower*), then the H-equilibrium will obtain, while if they are almost sure they are (*same same but different*), then the E-equilibrium will obtain. Conversely, $\frac{b}{a+b} < \beta_x^{\#} < \frac{b}{b+c}$, confirming the intuition by which, when it is not so clear who is who, row players (the entrepreneurs) find it optimal to play the H-strategy while column players (the workers) prefer the E-strategy. This is shown in fig. 2, where β indicates the probability that players are (*Upper and lower*), while p(β) and q(β) respectively are the probability that player 1 and player 2 play the H-strategy.

---

or column) to play the same strategy. If anything, the strategies in γ_w are more similar to (perfectly) correlated strategies which, not surprisingly, have had a more skeptical reception among game theorists (see Aumann, 1987, Aoki, 2007, and Gintis, 2009, ch. 7)

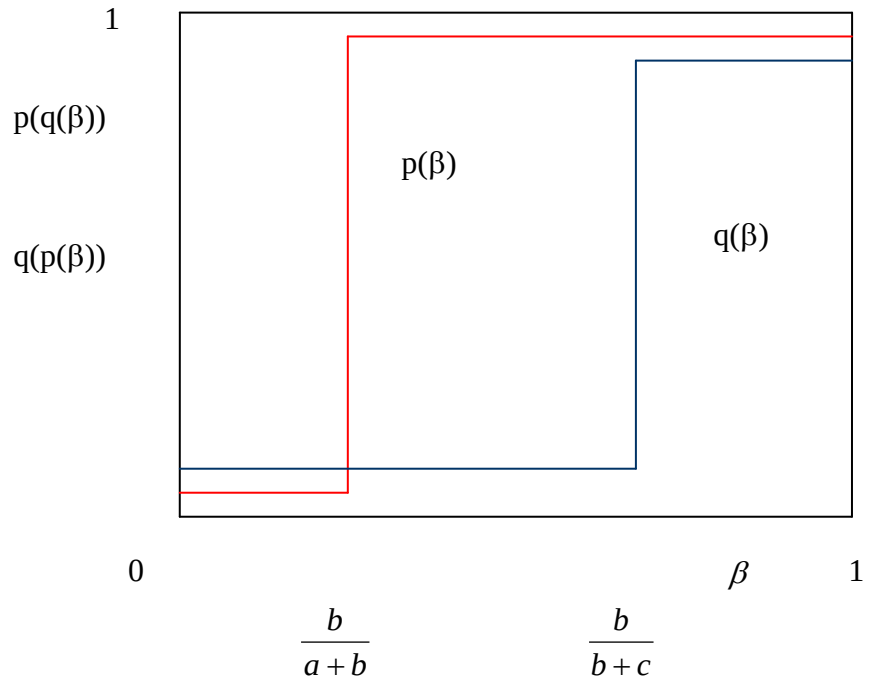

FIG. 2

The last thing to be noted about this picture is that, though at this point it does not make a great deal of difference, not only does β not represent the individual probability assessment about the opponent's type, but it may refer to the actual opponent as well as to the players of the other games which are being played; that is, it may refer to player 2 as well as to player 1' or 2'.

The next and most important step concerns the evolution of the beliefs system or, more precisely, its co-evolution with the strategies profiles, which in formal terms respectively stand for reality and knowledge. The key notion here is that of recursivity, which, as already noted, means that the equilibria of one game determine the rules of the other and vice versa through their effect on the beliefs system.

Probably the best way to familiarize oneself with this kind of circular causation mechanism is to see it as a combination of two approaches that, without renouncing the use of game theory to analyse institutions, have nevertheless found it impossible to accept the epistemic assumption of classical

game theory.[9] The first is the evolutionary game theoretical approach, where beliefs and strategies are assumed to co-evolve in a feedback loop starting from precedents determining beliefs (or expectations) which, in their turn, determine actions that then become precedents for the next round of play (Young, 1998). The difference here is that precedents come from γ_b while the resulting beliefs determine strategy profiles in γ_w. Moreover, since we are interested in the robustness of the pairs of equilibria in the two games, rather than in the process through which an equilibrium is reached in a single game, we limit our attention to strategy profiles that are an equilibrium for their game. This reasoning can be therefore visualized by means of the following expression: $[\hat{\sigma}_i^0 = g(\beta_0^*(\hat{\sigma}_g^0))]$, where $\hat{\sigma}$ indicates an equilibrium strategy profile.

The second approach is called Historical and Comparative Institutional Analysis (Greif, 2006) and, on the contrary, starts from beliefs determining actions which in turn confirm or otherwise the initial beliefs depending on whether or not they are self-enforcing. The difference here concerns the restrictions imposed by the equilibrium concept with respect to which self-enforcement is defined, that is, the absence of tendencies to change the rules of the game rather than confirmed expectations (as in the Nash equilibrium) or credible threats (as in sub-game perfection). Accordingly, one can write: $[\beta_1 = h(\hat{\sigma}_i^0(\beta_0^*))]$.

The first component of this combination –the 'deductive' part of the assumed learning process- is formalized in the following assumption, while the second –the'inductive' part- is formalized in the following proposition, which contains the proposed updating rule:

ASSUMPTION 1

$$(2)\ \beta_0 = f(\hat{\sigma}_g^0) = \hat{\sigma}_g^0 = \beta^*.$$

This assumption, which finds its justification in the focus on the stability of the pairs of equilibria in the two games, simply makes it possible to start from one of these pairs of equilibria at time t=0. For instance, since inter-group relationships minimize the participation constraint at the macro-level, , so as to determine the structure of property rights within the group, if almost all groups choose the U-strategy in γ_b, a strict separation between decision-makers and subordinates will obtain. Accordingly, players will be almost sure

[9] The problem for this kind of analysis, in effect, is not only that such assumptions are patently unrealistic but also that, since players are assumed to know everything of interest from the outset, they preclude an understanding of both institutional change and institutional variety.

that they are, respectively, (*Upper* and *lower*), so that the beliefs system will be $\beta_H^* \geq \frac{b}{b+c}$ and consequently the H-equilibrium will be obtained in γ_w.

As usual in these circumstances, it should be noted that it is a simplifying assumption that does not entail generality losses. For one thing, since, as noted in footnote 5, the real problem being considered is the life cycle of a technological system, it may well be the case that, in the phase of introduction, the pay-offs in γ_b are such that the strategies corresponding to the new paradigm actually imply an equilibrium in dominant strategies. Even if this is not the case, and even if a more realistic interpretation with *n* groups is considered, the structure of the game implies that a disequilibrium configuration will not be long-lasting even in real time.[10]

PROPOSITION 1

Assume $\Pi_i(\hat{\sigma}_g^t) \geq w_i = \begin{cases} \overline{w} \\ \hat{w} \end{cases}$, with i=1, 1', 2, 2' and t=0,1.

Then,

$$(3)\quad \beta_1 = g(\hat{\sigma}_i^0(\beta_o^*)) = \begin{cases} \beta^* \text{ if } \hat{\sigma}_{\mathrm{i}}^0 = \mathrm{argmax}\Pi_{\mathrm{g}}(\hat{\sigma}_{\mathrm{i}}^0, \hat{\sigma}_{\mathrm{i}}^{'0}), \forall \hat{\sigma}_{\mathrm{i}}^0, \hat{\sigma}_{\mathrm{i}}^{'0} \in \hat{\sigma}_{\mathrm{i}} \\ \beta^{\#} \text{ if } \hat{\sigma}_{\mathrm{i}}^0 \neq \mathrm{argmax}\Pi_{\mathrm{g}}(\hat{\sigma}_{\mathrm{i}}^0, \hat{\sigma}_{\mathrm{i}}^{'0}), \forall \hat{\sigma}_{\mathrm{i}}^0, \hat{\sigma}_{\mathrm{i}}^{'0} \in \hat{\sigma}_{\mathrm{i}} \end{cases}$$

To prove this proposition, which states that the maximization of the value of the group as a function of the equilibrium strategy profiles within the group is a necessary condition for an equilibrium beliefs system, first note that, if that is not the case, that is, if $\hat{\sigma}_i^0$ is the H-equilibrium, there is an alternative equilibrium strategy profile within the group (the E-equilibrium) which maximize the value of the group and is realizable as an alternative equilibrium between groups which, as a consequence, maximizes total value (the R-equilibrium). However, since the original equilibrium within the group is

[10] The point is that a disequilibrium strategy combination in γ_b, or even a mixed strategy equilibrium, is associated with a disequilibrium strategy combination in γ_w - and so with the off diagonal 0-pay-offs - for at least some groups, so as to accelerate the convergence to one of the two pure strategy equilibria.

Pareto-efficient, this alternative equilibrium between groups cannot be implemented by a simple change in group strategies and a deeper structural change is needed. In effect**,** this also means that there is a conflict between groups in the sense that, in the alternative equilibrium, total value is maximized but one group (formed by players 2 and 2') is partly better off at the expense of the other (formed by players 1 and 1').

Consequently, to the extent that the learning process mentioned in footnote 5, which works by analogy and is helped by the notion of recurrence, allow players to recognize the effect of inter-group relationships on intra-group relationships through the beliefs system, inverting that direction of causality or in other words passing from forward to backward induction, the implicit assumption of Bayesian updating –'that the dimensions of the internal mental models used to represent the external world are correct, in some sense' (Denzau and North, 1994, p. 17) – does no longer hold and players will use the minimax-maximin principle (Von Neumann and Morgestern, 1944). Indeed, given the positional nature of power and the already mentioned relational nature of the relevant knowledge, the strategic structure of the problem can be characterized as being zero-sum in types though not in pay-off.[11] In effect, while β is the probability that players are (*Upper and lower*), the probability that they are both (*same same but different*) is precisely (1-β). Accordingly, provided that if the resulting beliefs' system is to be shared it must be epistemologically objective, it will be determined by the intersection of the different probability assessments made by the two players which, to paraphrase the title of Aumann's 1976 paper, can be seen as the sub-set of commonly known posteriors on which players agree to disagree.

(4) $\beta_1 = \beta_1^1 \wedge \beta_1^2 = \beta^{\#}$, where

(5) $\beta_1^1 = \max_{\beta^1}[\min_{\beta^2} \Pi^1] = \min_{\beta^1}[\max_{\beta^2} \Pi^2] \geq b/a+b]$, and

[11] As an intuitive justification of the implausibility of Bayesian rationality in this context, note that Bayesian updating would amount to a max-max principle for row players and to a min-min principle for column players. With respect to the more sober mini-max criterion of the Wald principle which gave rise to the literature on non additive probabilities, they seem definitely too 'optimistic' or 'pessimistic'. Moreover, since stagnation in present context is equivalent to a crisis in a richer framework allowing for the role of finance in artificially creating booms and recessions, deviating from Bayesian updating may be simply seen as a consequence of its intrinsic 'conservative' nature. See Gilboa and Schmeidler (1989) for a different, but related, adaptation of expected utility theory to the case with more than one possible probability judgment, and Pagano (1989) for a detailed discussion of the positional nature of power

(6) $\beta_1^2 = \max_{\beta^2}[\min_{\beta^1} \Pi^2] = \min_{\beta^2}[\max_{\beta^1} \Pi^1] \leq b/b+c$.

If, instead, the equilibrium strategies profiles within the group maximize the value of the group as a function of individual strategies, naturally, there is not the strict separation between decision-makers and subordinates of the previous case, that is, there are no groups other than the original ones formed by players 1 and 2, and 1' and 2', so that the equilibrium beliefs system inherited from the past is accepted as epistemically objective by all the players, who therefore apply the Bayes law confirming what they already know.

(7) $\beta_1 \geq \beta_0 = \beta^*$.

These two case are represented in fig.3 below, where $r(\hat{\sigma}_i^0)$ indicates the probability that equilibrium strategy profile in γ_w at t=0 is (H,H).

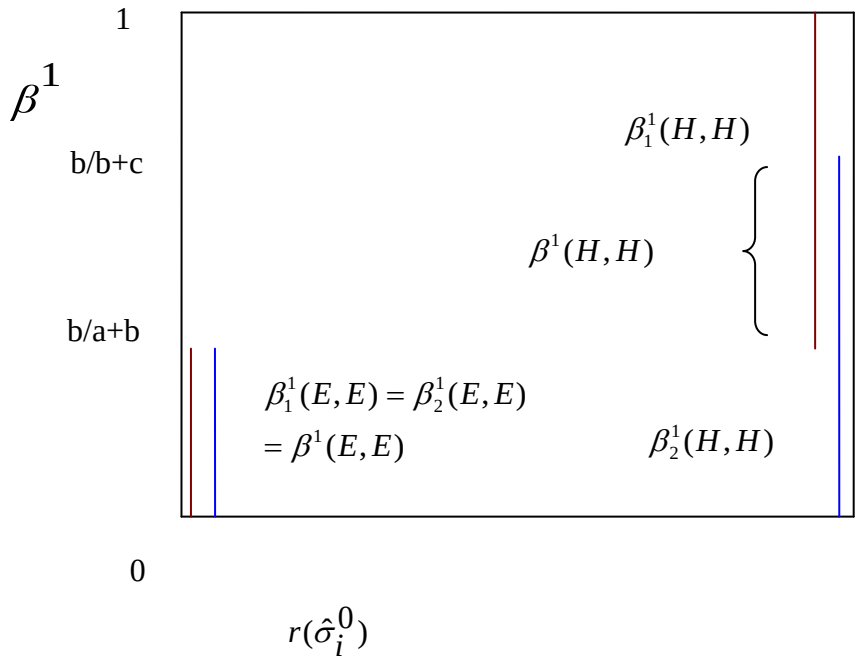

FIG. 3

The interesting feature of this proposition is that, when the resulting beliefs system is a disequilibrium beliefs system, precisely on the basis of such

a new beliefs system, different groups (or classes[12]) are formed from the original groups (or firms), which amounts to a change in the rules of the game in the game between groups and to a change in the original equilibrium. In effect, whilst having players 1 and 1' play the H-strategy and players 2 and 2' play the E-strategy would be quite irrational in the games within the groups taken in isolation, in the super-game this amounts to the formation of the group of the '*Uppers*' and to that of the '*lowers*'. Then, since $\hat{w} > X/2 > c$ because of the assumption about team production, the only equilibrium in the modified game between groups is the pay-off dominant R-equilibrium (fig. 4).

lowers / Uppers	UNREG.	REG.
UNREG.	2a, 2c	a+X/2, c+X/2
REG.	a+X/2, c+X/2	2b,2b

Fig.4

It is in this sense, therefore, that the pair of equilibria (U, U in γ_b, and H, H in γ_w) is not robust to a change in the rules of the game, or not perfect in the super-game because it is not an equilibrium of the modified game where players evaluate the pay-offs of the original games from a different point of view, leaving (R, R in γ_b, and E, E in γ_w) as the only multi-level equilibrium for the super-game.

More formally, therefore, we may state the following definition, where the requirement about absence of tendencies to change the rules turns out to be equivalent to the constancy of the beliefs system across generations:

DEFINITION 2:

[12] See Battistini (2013) for a more detailed discussion of the controversial concept of social classes. Here it may suffice to note that they are not given an ontological status and that class division is not taken as the exogenous driven of the analysis. To the contrary, it is the process of formation and exhaustion of class interests that is endogenously determined by the value-maximizing mechanisms implied by the multi-level operation of selection.

A super-game perfect multi-level equilibrium of $\Gamma = \{\gamma_b^t, \gamma_w^t\}$ at time *t* requires:

(i) $\sigma_g^t = \arg\max \Pi_g(\sigma_g^t, \sigma_h^t / \beta_t) \forall g, h \in g = 1,....,n/m$

(ii) $\sigma_i^t = \arg\max \Pi_i(\sigma_i^t, \sigma_j^t / \beta_t) \forall i, j \in i = 1,.....,n$

(iii) $\beta_t^* = \beta_{t-1}^*$

But now, since by Proposition 1 the maximization of total value is a necessary condition for a multi-level equilibrium and, on the other hand, by Definition 2 a multi-level equilibrium is a necessary condition for total value maximization, while because of wealth effects egalitarianism is always efficiency-enhancing, we may write the following proposition:

PROPOSITION 2

Assume $\hat{w} > X/2 > c$. Then at t=1 a super-game perfect equilibrium of $\Gamma = \{\gamma_b^1, \gamma_w^1\}$ exists if and only if it is efficient and egalitarian.

Basically, to conclude, this proposition can be understood as a consequence of a notion of 'structural interdependence' according to which individual (group) pay-offs depend not only on individual (group) strategies but also on group (individual) strategies through their effect on the rules of the game (as shown in Proposition 1). The requirement about the absence of tendencies to change the rules of the game can thus be reframed in terms of the absence of tendencies to change groups, which is the reason behind the egalitarian part of the result. It is with respect to such a notion, then, that the rationality of the updating rule above is relative.[13]

Analogously, since a multi-level equilibrium is consistent with the most general definition of a non-cooperative equilibrium as a self-enforcing

[13] Interestingly, such a notion of structural interdependence may be related to several ideas already present in the behavioral sciences literature (see Gintis, 2009, ch.12). In psychology, for example, there is a distinction between deliberative choices, which occur relatively infrequently in the course of a life, and routine choices, which only involve comparison among alternatives. Since as can be guessed by the price theory mentioned above, structural interdependence implies a choice *of* alternatives, it may be seen as an instance of the former, though such a choice is not made by any single individual or group but by society. Similarly, in the biological literature on niche construction, there is an analysis of species that not only adapt but also change the environment of selection, which in the present framework is clearly analogous to the change of the rules of games, though the role of intentionality is obviously more prominent.

agreement (Aumann, 1974), the same is true for the contribution of the paper, which lies in its attempt to model the process through which agreement, rather than being dogmatically[14] assumed from the outset, turns out to be socially constructed by interested parties.

Naturally, although in this example such a construction is spontaneous in the sense that, in principle, neither an external intervention is required, nor anyone is forced to do anything, as already mentioned, it would be rather ingenuous not to consider that, in reality, the mechanisms through which social knowledge is formed may be quite different. In effect, the issue itself of this paper in some way is a proof of this. Nevertheless, the fact that the group behaviors analysed in this paper do not require the solution of any collective action problem also means that the so-called 'n-problem' is no longer a problem and instead becomes an asset.

References

Alchian, A., Demsetz, H. (1972), Production, Information Costs, and Economic Organization, *American Economic Review*, 62, pp. 777-795.

Aoki, M. (2007), Endogenizing Institutions and Institutional Change, *Journal of Institutional Economics*, 3:1, pp. 1-31.

Aumann, R.J. (1974), Subjectivity and Correlation in Randomizing Strategies, *Journal of Mathematical Economics*, 67-96.

Aumann, R.J. (1976), Agreeing to Disagree, *Annals of Statistics*, 4:6, 1236-1239.

Aumann, R.J. (1987), Correlated Equilibrium as an Expression of Bayesian Rationality, *Econometrica*, 55, 1-18.

Aumann, R.J., Brandeburger, A. (1995), Epistemic Conditions for Nash Equilibrium, *Econometrica*, 65:5, 1161.1180.

Battistini, A. (2013), A Theory of Profit and Competition, *Evolutionary and Institutional Economics Review*, forthcoming.

Bowles, S. (2004), *Microeconomics: Behavior, Institutions, and Evolution*, Princeton University Press.

Coase, R.H. (1992), The Institutional Structure of Production, *American Economic Review*, 82, 713-719.

Denzau, A.T. North, D.C. (1994), Shared Mental Models: Ideologies and Institutions, *Kyklos*, 47:1, pp. 3-31.

Gilboa, I, Schmeidler, D. (1989), Maxmin Expected Utility with Non-Unique Priors, *Journal of Mathematical Economics*, 18, 141-153.

Gintis, H. (2009), *The Bounds of Reason: Game Theory and the Unification of Social Sciences*, Princeton University Press.

[14] I borrow the term from Kreps (1990, ch. 3) See also Denzau and North (1994).

Greif, A. (2005), *Institutions and the Path to Modern Economy: Lessons from Medieval Trade,* Cambridge University Press.
Harsanyi, J. (1967-1968), Games of Incomplete Information Played by Bayesian Players, Parts I-III, *Management Science*, 159-182; 320-334; 486-502.
Henrich, J. (2004), Cultural Group Selection, Coevolutionary Processes and Large-Scale Cooperation, *Journal of Economic Behavior and Organization*, 53-1, 3-35.
Kreps, D.M. (1990), *A Course on Microeconomic Theory*, Harvester-Wheatsheaf.
Maddison, A. (1982), *Phases of Capitalistic Development,* Cambridge University Press.
Maynard-Smith, J. (1982), *Evolution and the Theory of Games*, Cambridge University Press.
Mokyr, J. (2002), *The Gifts of Athena. Historical Origins of the Knowledge Economy*. Princeton University Press.
Pagano, U. (1999), Is Power an Economic Good? Notes on Social Scarcity and the Economics of Positional Goods, in S. Bowles, M. Franzini and U. Pagano, eds., *The Politics and the Economics of Power*, London: Routledge, 63-85.
Perez, C. (2002), *Technological Revolutions and Financial Capital: The Dynamics of Bubble and Golden Ages*, Oxford University Press.
Searle, J. (2005), What Is an Institution?, *Journal of Institutional Economics*, 1:1, pp. 1-21.
Sober, E., Wilson, D.S. (1999), *Unto Others. The Evolution and Psychology of Unselfish Behavior,* Harvard University Press.