

On Voluntariness of Nash Equilibrium

Paolo Caravani*

*Dept. of Electrical and Information Engineering
University of L'Aquila
Via Campo di Pile, 67100 L'Aquila, It
paolo.caravani@univaq.it*

Abstract The paper deals with pure strategy equilibria of bi-matrix games. It is argued that the set of Nash equilibria can contain voluntary as well as involuntary outcomes. Only the former are indicative of consistent expectations. In the context of repeated play with incomplete information, simulations show that involuntary equilibria tend to occur more frequently than voluntary equilibria. Consequences in econometric practice and philosophical implications are briefly hinted at.

Keywords: matrix games, expectations, pure equilibria, learning, incomplete information.

1. Introduction

In this note we discuss *voluntariness*, a psychological trait that despite its paramount importance in cognitive and behavioural sciences, seems to have been overlooked in Game Theory. It is argued that under certain conditions the very notion of Nash equilibrium together with its predictive and normative power are crucially related to the question: to what extent is the equilibrium outcome of a game what players wanted it to be? The exposition is initially framed in the simplest possible setting, one-shot bi-matrix games with known payoffs. It is assumed that rationality is exhaustively captured by players' optimizing behaviour in terms of best reply to an opponent's *expected* strategy. The mechanism by which expectations are formed is not assumed to be an ingredient of rationality, i.e. expectations may or may not be rational, consistent or inconsistent. It is shown that the set of outcomes arising from these assumptions may contain a Nash equilibrium even when expectations are inconsistent. We term *involuntary* such an equilibrium. We then remove the assumption of knowledge of the opponent's payoffs and show that the performance of a learning algorithm in presence of two pure-strategy equilibria, one voluntary one involuntary, exhibits - surprisingly - the prevalence of the latter. Our discussion dispels a common misconception present in the literature - the identification of Nash equilibrium to consistent alignment of expectations. Besides consequences in econometrics, specifically in the sort of statistical inference used to monitor agents' psychologies, the question involves deeper epistemic aspects. The possibility of involuntary equilibria is akin to whether Knowledge could be assimilated to 'Justified True Belief', a thesis denied in the classical paradox of (Gettier, 1963). Furthermore, involuntary equilibria provide substantive support to the compatibilist view in the debate over Free Will vs. Determinism (Dennet, 1984).

* HYCON support acknowledged. Work partially supported by European Commission under STREP project TREN/07/FP6AE/S07.71574/037180 IFLY.

2. Definitions

It is well known that a Nash equilibrium in two-person games is defined as a pair of strategies \hat{u}_1, \hat{u}_2 satisfying

$$J_1(\hat{u}_1, \hat{u}_2) \geq J_1(u_1, \hat{u}_2) \quad \forall u_1 \in U_1 \quad (1)$$

$$J_2(\hat{u}_1, \hat{u}_2) \geq J_2(\hat{u}_1, u_2) \quad \forall u_2 \in U_2 \quad (2)$$

where J_i, U_i define the objective function (maximand) and the strategy space of player i . Definition (1-2) is equivalent to

$$\hat{u}_1 = \operatorname{argmax}_{u_1} J_1(u_1, \hat{u}_2)$$

$$\hat{u}_2 = \operatorname{argmax}_{u_2} J_2(\hat{u}_1, u_2).$$

If the best-reply operator

$$B : U_1 \times U_2 \mapsto U_1 \times U_2 \quad (3)$$

$$B = \begin{bmatrix} B_1(u_2) \\ B_2(u_1) \end{bmatrix} \quad (4)$$

$$B_1(u_2) = \operatorname{argmax}_{u_1} J_1(u_1, u_2) \quad (5)$$

$$B_2(u_1) = \operatorname{argmax}_{u_2} J_2(u_1, u_2), \quad (6)$$

is single-valued, a second characterization of Nash equilibrium can be given in terms of the fixed point¹ of B

$$\begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \end{bmatrix} = \hat{u} \text{ is a Nash equilibrium} \iff \hat{u} = B(\hat{u}). \quad (7)$$

A third characterization, mainly focused on strategic aspects, can be given in terms of expectations. We suppose in the following B exists is continuous and single-valued over $U_1 \times U_2$.

3. Expectations

Denote by the symbol \mathcal{E} the *subjective* expectation that each player has over the action of his opponent. In the case, for example, of Bayesian statistics \mathcal{E} may coincide with the mathematical expectation E or with the conditional expectation $E(\cdot|\cdot)$ if posterior information is available. This of course would presuppose that players' expectations are formed on the basis of specific (and known) probability distributions over the opponent's strategy sets, perhaps in connection with externally observed events. Since any such specification would introduce a degree of arbitrariness at this stage, we stay clear of any distributional assumption. By $\mathcal{E}_j U_i$ we mean that

¹ This of course requires the operator B to be defined on a domain $\mathcal{D} \subset U_1 \times U_2$ containing the point \hat{u} . Unfortunately, it is frequent the case that a Nash equilibrium exists at a point where B is discontinuous, as in mixed strategy equilibria of matrix games. In other cases B may turn out to be nowhere defined *except* at the point \hat{u} , hence \mathcal{D} may be an *a-priori* unknown domain. Even when B is defined and continuous on some domain it may not be single-valued. In game theory these are well known sources of computational difficulty.

element of U_i with respect to which player j evaluates - and implements - his own best reply, regardless of the reason why such a conviction came to materialize in his mind. We will say that player j 's expectation is *fulfilled* if, after i 's move is taken, it turns out $u_i = \mathcal{E}_j U_i$. Players' expectations will be said *consistent* (in two-person games) if they are both fulfilled.

Expectation Principle: if expectations are consistent, then a Nash equilibrium exists and it coincides with the expected outcome.

Proof. Being at a best-reply v_1, v_2 means

$$v_1 = \operatorname{argmax}_{u_1} J_1(u_1, \mathcal{E}_1 U_2) \tag{8}$$

$$v_2 = \operatorname{argmax}_{u_2} J_2(\mathcal{E}_2 U_1, u_2) \tag{9}$$

and if $v_1 = \mathcal{E}_2 U_1$ $v_2 = \mathcal{E}_1 U_2$ it follows

$$J_1(\mathcal{E}_2 U_1, \mathcal{E}_1 U_2) \geq J_1(u_1, \mathcal{E}_1 U_2) \quad \forall u_1 \in U_1 \tag{10}$$

$$J_2(\mathcal{E}_2 U_1, \mathcal{E}_1 U_2) \geq J_2(\mathcal{E}_2 U_1, u_2) \quad \forall u_2 \in U_2 \tag{11}$$

hence v_1, v_2 are equilibrium strategies and coincide with the expected outcome. \square

The condition is only sufficient because

$$\operatorname{argmax}_{u_1} J_1(u_1, \mathcal{E}_1 U_2) = \operatorname{argmax}_{u_1} J_1(u_1, v_2)$$

is implied by – but it does not imply – $v_2 = \mathcal{E}_1 U_2$ (*idem* for $\mathcal{E}_2 U_1$). In other words, u_1 may constitute a best reply to more than just one pure strategy of the opponent. If this occurs when u_1, u_2 is an equilibrium, such an equilibrium does not require player 1's expectation to be fulfilled. It may well happen that inconsistent expectations lead to a Nash equilibrium. In this case we should speak of an *involuntary* equilibrium.

To say that a given equilibrium is involuntary is better suited to the judgment of a psychologist than to the speculation of a game theorist. What *can* be established within our discipline though is whether or not the hypothesis of involuntariness can be *logically* held. Involuntariness can be ruled out in cases where *any* pair of inconsistent expectations is in dis-equilibrium. To illustrate, consider²

	<i>L</i>	<i>C</i>	<i>R</i>
<i>t</i>	(2, 3)	(4, 1)	(0, 1)
<i>m</i>	(1, 0)	(0, 4)	(5, 5)
<i>b</i>	(1, 3)	(2, 1)	(3, 0).

This game has exactly two equilibria in pure strategies. These are (t, L) yielding utilities (2, 3); and (m, R) yielding utilities (5, 5). Suppose both players expect to be facing a *very cautious* opponent. For example, Row expects Column to play *C* (this is the only strategy ruling out zero utility for Column, e.g. a max-min strategy). Consequently, Row will play *t*, his best reply to *C*. Symmetrically, if Column expects Row to play *b* (Row's max-min strategy) her best reply is *L*. So the outcome associated to these expectations is (t, L) . It is easy to check that this is indeed an equilibrium outcome although the expectations leading to it were *not*

² left numbers in brackets are utilities of Row player, right numbers of Col player.

fulfilled. They predicted (b, C) , i.e. they were *inconsistent*. Hence we cannot rule out that when the equilibrium (t, L) is reached, it is reached involuntarily. To be sure, outcome (t, L) is a Nash equilibrium, but it does not necessarily mirror a system of self-fulfilling expectations.

Notice that this conclusion does not apply to the other equilibrium of the game. The outcome (m, R) cannot be considered an involuntary equilibrium as m is the best reply to R and to no other strategy of Column; and R is the best reply to m and to no other strategy of Row.

On the other hand consider the game dubbed *Chicken* in game theory folklore. Two drivers arrive at the same time at an intersection. Each can keep going (g) or stop (s). Both prefer (g) to (s) but only if the other stops, otherwise they prefer (s) and can be represented by the pay-off matrix

$$\begin{array}{cc} & g & s \\ g & (0, 0) & (3, 1) \\ s & (1, 3) & (2, 2). \end{array}$$

The game has two pure strategy equilibria, (g, s) and (s, g) . There are 4 possibilities for the expectations

1. *Row expects Column to stop and Column expects Row to go.* Best-reply prescribes Row to go and Column to stop, expectations are consistent, the outcome is equilibrium.
2. *Row expects Column to go and Column expects Row to stop.* Best-reply prescribes Column to go and Row to stop, expectations are consistent, the outcome is equilibrium.
3. *Both expect that the other goes.* Best-reply prescribes both to stop, expectations are inconsistent, the outcome is not equilibrium.
4. *Row expects Column to stop and Column expects Row to stop.* Best-reply prescribes both to go, expectations are inconsistent, the (disastrous) outcome is not equilibrium.

In this case involuntariness is to be ruled out on purely *algorithmic* ground: if an equilibrium is reached, it is reached voluntarily - no psychology required.

Motivated by the above, we give the following

Definition 1. A pure-strategy equilibrium of bi-matrix game is *Involuntary* if it contains strategies that are best-replies to more than one strategy of the opponent; otherwise is *Voluntary*.

4. Best-Reply Graph

A useful tool to examine the existence and the nature of pure-strategy equilibria is the Best-Reply Graph (BRG). This is a bi-partite graph

$$G(\mathbb{I}_n, \mathbb{I}_m, E_{nm}, E_{mn})$$

whose nodes are strategies and edges are best-replies. An edge $e_{ji} \in E_{nm}$ connects node $j \in \mathbb{I}_m$ to node $i \in \mathbb{I}_n$ if pure strategy i is a best-reply to pure strategy j ,

e.g. $i = \mathcal{B}_1(j)$, and an edge $e_{ij} \in E_{mn}$ connects node $i \in \mathbb{I}_n$ to node $j \in \mathbb{I}_n$ if pure strategy j is a best-reply to pure strategy i , e.g. $j = \mathcal{B}_2(i)$. For example the game

	L	C	R
t	(6, 7)	(8, 4)	(0, 3)
m	(4, 5)	(3, 4)	(7, 6)
b	(5, 5)	(4, 0)	(5, 4)

has a BRG

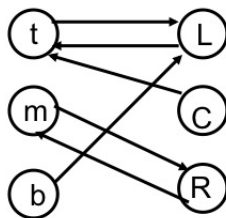


Figure1: Best-Reply Graph

A node with in-degree= m is a *Dominant* strategy of player 1 (similarly for player 2 with in-degree= n); nodes with in-degree= 0 are called *Never-Best-Reply* (NBR) strategies; a cycle of the BRG of length > 2 is an *Improvement Cycle*. In Fig. 1 there are no dominant strategies, b and C are NBR, and there are no improvement cycles. Moreover,

- i. each node of a BRG has out-degree > 0
- ii. nodes of Nash equilibria belong to cycles of length= 2
- iii. nodes of *strict* equilibria have out-degree= 1
- iv. an equilibrium is involuntary if at least one of its nodes has in-degree > 1
- v. every potential game has at least one pure-strategy equilibrium
- vi. potential games have no improvement cycles.

i. follows from existence of B ; ii., iii. and iv. hold by definition; v. and vi. are well known (e.g. Cor 2.2 and 2.3 in (Monderer and Shapley, 1996)). The game in Fig.1 has one voluntary equilibrium (m, R) and one involuntary equilibrium (t, L) .

Remark 1. Notice that involuntary equilibria are all but rare in matrix games. For instance, all equilibria in dominant strategies are involuntary since by definition a dominant strategy is a best-reply to more than one (indeed to any) of the opponent's strategies. Perhaps the epitomy of equilibria in dominant strategies is the Prisoner's Dilemma³

	c	d
c	(1, 1)	(3, 0)
d	(0, 3)	(2, 2)

³ pay-offs are years of prisonment.

where Deny (d) dominates Confess (c) for both players. If both players expect (c) from the opponent, their best replies lead to (d, d) with inconsistent expectations, i.e. (d, d) is an involuntary equilibrium.

Furthermore

Proposition 1. *Every potential game with $n \neq m$ has at least one Involuntary Equilibrium*

Proof. If Row-strategies are $r_1 \dots r_n$ and Col-strategies are $c_1 \dots c_m$ (take $n < m$ wlog) re-number nodes such that the equilibria of the potential game are $(r_i, c_i), i \in \mathcal{P}$ and assume they are all voluntary. If $|\mathcal{P}| = n$, then there is an outgoing edge

$$c_h \mapsto r_k \quad \text{with } h \notin \mathcal{P}, k \in \mathcal{P}$$

showing that r_k has indegree > 1 hence (r_k, c_k) is involuntary. If $|\mathcal{P}| < n$, use outgoing edges to re-order nodes with index $k \notin \mathcal{P}$

$$r_{k_1} \mapsto c_{k_1} \mapsto r_{k_2} \mapsto c_{k_2} \dots \mapsto r_n \mapsto c_n$$

Since c_n must have one outgoing edge

$$c_n \mapsto r_i \quad \text{for some } i \in \{1 \dots n\}$$

But $i = n$ for otherwise there would be an improvement cycle and the game would not be potential. Thus (r_n, c_n) is a 2-cycle containing node r_n with in-degree > 1 , i.e. an Involuntary Equilibrium. \square

This result confirms the set of cases in which matrix games exhibit involuntary equilibria is of non-negligible importance.

5. Incomplete information

Notice that *involuntariness* is entirely based on *best-reply* and not at all on *expectations* nor on the mechanism by which they are formed. In the first example of Sec. 3 we assumed, for the sake of the argument, each player expected the opponent to be very cautious. Given the structure of the payoff matrices, that seemed a plausible argument: individual rationality may well motivate max-min strategies, for example when common knowledge cannot be invoked. However involuntary equilibria may arise with completely irrational expectations. For example, in the game of Fig. 1 max-min rationality would suggest the row-player to play b rather than t . In fact, expectations not even require to be formed on the basis of *known* opponents payoffs. Assume players only know their *own* payoffs - not their opponent's. In the frame of repeated games with simultaneous moves, let each player know only *own* current and *own* past performance. Assume players adopt, among the many proposed, a so called linear-reward learning scheme (Thathachar and Sastry, 2004). It works like this. A pure strategy s of player 1 is sampled randomly at round t , according to a probability vector $p(t) \in S^n$ (unit simplex of \mathbb{R}^n) which updates as

$$p_i(t+1) = \begin{cases} p_i(t) + \eta P(t)(1 - p_i(t)) & \text{if } i = s \\ p_i(t) - \eta P(t)p_i(t) & \text{if } i \neq s \end{cases} \quad (12)$$

where $P(t)$ is the outcome payoff of player 1 normalized to the unit interval, and η a learning rate chosen in the same interval. Similarly, player 2 selects pure strategy r sampled with probability $q(t) \in S^m$ with

$$q_i(t+1) = \begin{cases} q_i(t) + \eta Q(t)(1 - q_i(t)) & \text{if } i = r \\ q_i(t) - \eta Q(t)q_i(t) & \text{if } i \neq r. \end{cases} \quad (13)$$

It is known (Sastry et al., 1994) that, for small η , each pure-strategy strict Nash equilibrium is a locally asymptotically stable point of (12,13) and any point which is not a Nash equilibrium is unstable.

Given the results of the previous section it is interesting to use the above learning scheme to assess the relative frequency of occurrence of an involuntary equilibrium in the class of potential games possessing one voluntary and one involuntary equilibrium. The first example of Sec. 3 belongs to this class and the above learning scheme typically produces results of the kind shown below. Initial probabilities p, q are all set to $1/3$ and $\eta = 0.1$. Fig. 2 shows convergence of p to $[0\ 1\ 0]$ (top) and of q to $[0\ 0\ 1]$ (bottom). This is the voluntary equilibrium (m, R) yielding payoffs $(5, 5)$.

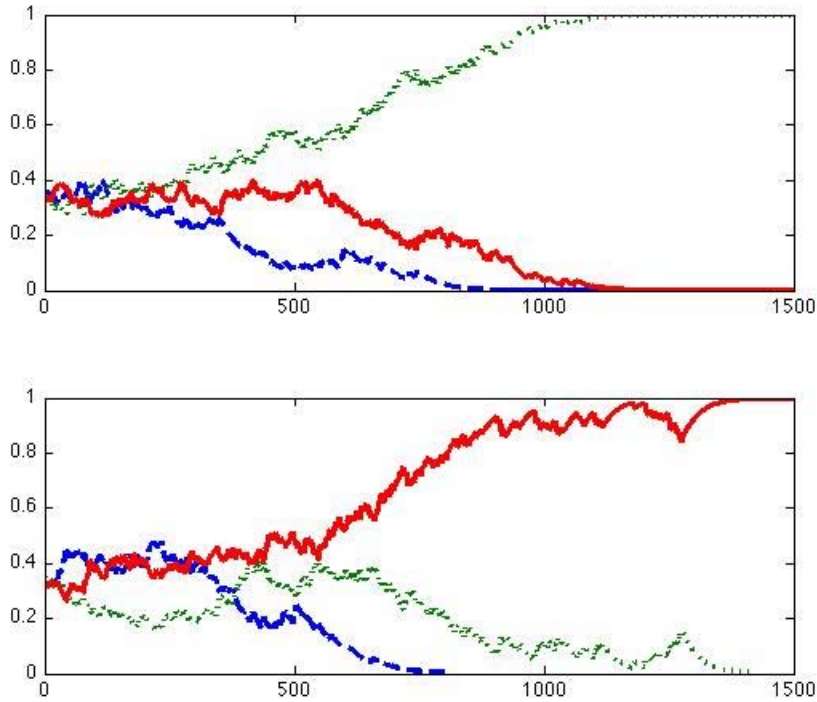


Figure2: Convergence of p (top) and q (bottom) to voluntary equilibrium

Fig. 3 shows convergence of p to $[1\ 0\ 0]$ (top) and of q to $[1\ 0\ 0]$ (bottom). This is the involuntary equilibrium (t, L) yielding payoffs $(2, 3)$. Convergence to one equilibrium or the other is regulated by chance, given the random sampling that takes

place at each round. The question of interest is long-run behaviour: which of the two outcomes prevails in a reiterated simulation cycle.

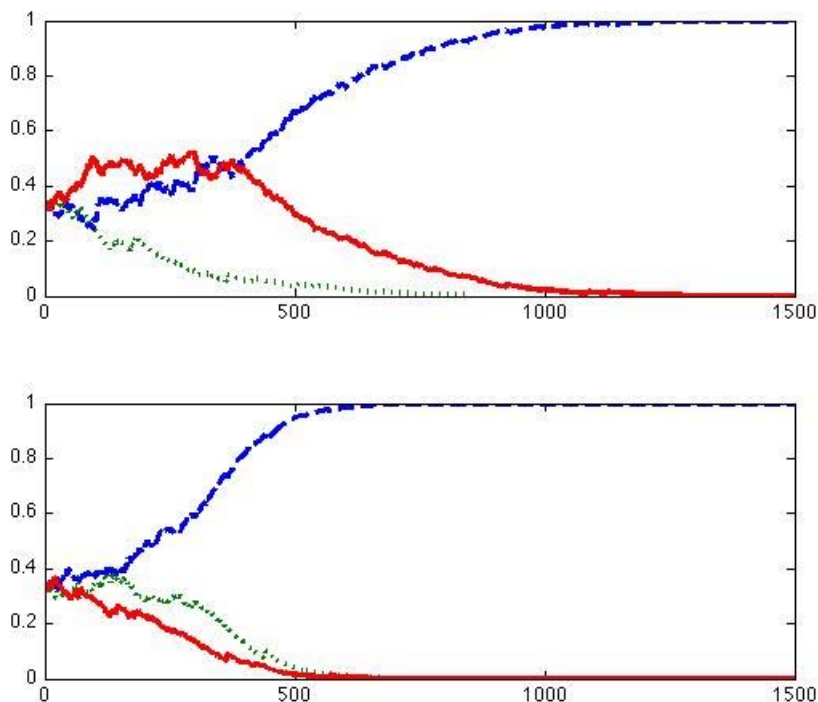


Figure3: *Convergence of p, q to involuntary equilibrium*

In Fig. 4 (bottom) the simulations shown in Figs. 1,2 have been replicated 2000 times. The dashed line represents the number of times an involuntary equilibrium has been reached, the solid line the number of times a voluntary equilibrium has been reached. The top graph of Fig. 4 (top) is a zoom on the first 20 rounds. It is apparent that in the long-run the involuntary equilibrium prevails.

6. Discussion and conclusion

The idea of relating equilibrium to self-fulfilling expectations pervades Economics before and after the onset of the rational expectations paradigm. Expectations determine behaviour. The standard problem is to deduce equilibrium price and quantities from consumer and producer's preferences. However inverse questions can be – and usually are – posed in Econometrics, including for example guessing consumers preferences from market-clearing prices; assessing investors risk-preferences from bond yields; dubbing stockmarket attitudes bullish or bearish, etc. Drawing positive conclusions on the nature of expectations from revealed behaviour means reversing the expectations-behaviour link somewhat arbitrarily through statistical inference. That self-fulfilling expectations lead to equilibrium is a truism following

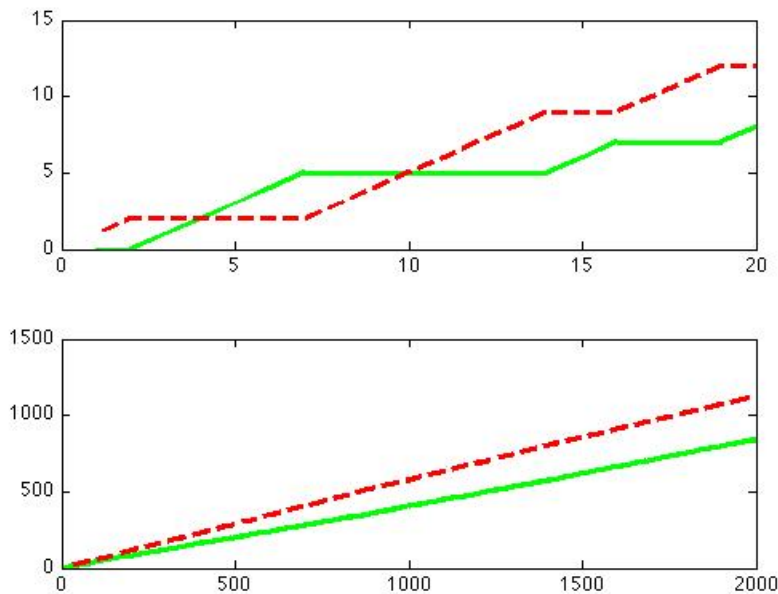


Figure4: Solid: number of times VOL is reached. Dash: number of times INVOL is reached.

straight from the definition of Nash (1,2). Revealing agents' expectation from observed equilibrium is a far more ambitious and definitely non-trivial task. Indeed, this task is not always viable and often doomed to fail. This is so, it is argued in this note, when equilibria are involuntary.

Consistent alignment of beliefs is not a new idea. The term *rationalizable* has been used to describe strategies a player can defend (i.e. rationalise) on the basis of beliefs about the beliefs of the opponent that are not inconsistent with the game's data (Bernheim, 1984), (Pearce, 1984). More precisely, rationalizable equilibria are outcomes of a game in which dominated strategies are iteratively eliminated (see also the notion of sophisticated equilibria in (Moulin, 1986)). Although rationalizable outcomes need not be Nash equilibria, every Nash equilibrium is a rationalizable outcome. But this is not to say that Nash equilibrium is sustained by consistent beliefs, as we have just shown. Despite this, some confusion seems to be present in the literature. For example, in (Hargreaves Heap and Varoufakis, 1995Sec 1.2.1) we read

Nash strategies are the only rationalizable ones which, if implemented, confirm the expectations on which they were based. This is why they are often referred to as self-confirming strategies or why it can be said that this equilibrium concept *requires* that players' beliefs are consistently aligned.

The implicit root of our question lies deep in epistemology, namely in the analysis of Knowledge and its controversial identification with Justified Belief (Dennet, 1989). Involuntary equilibria are akin to the Gettier case (Gettier, 1963) rephrased here as

Smith has applied for a job but, from rumors, has a justified belief that "Jones will get the job". Also he knows "Jones has 10 coins in his pocket", as he counted the coins in Jones's pocket ten minutes ago. Smith therefore (justifiably) concludes that "the man who will get the job has 10 coins in his pocket". In fact, Jones does not get the job. Instead, Smith does. However, as it happens, Smith (unknowingly and by sheer chance) also had 10 coins in his pocket. So his belief that "the man who will get the job has 10 coins in his pocket" was justified and true. But it does not appear to be knowledge.

The similarity is only partial. Whereas justified beliefs are necessary but not sufficient to Knowledge in the case of Gettier, they are sufficient but not necessary to reach an involuntary equilibrium. Indeed they not even need to be *justified* in this case.

Now if free-will originates from beliefs and beliefs are irrelevant to outcomes, then free-will is irrelevant to outcomes. Thus involuntary equilibria are an instance of determinism. Game theory, to the extent it can exhibit both voluntary and involuntary equilibria, offers a powerful paradigm in favour of compatibilism in the debate (Dennet, 1984) over Free-will vs. Determinism.

References

- Bernheim, D. (1984). *Rationalizable Strategic Behavior*. *Econometrica*, **52**, 1007–1028.
- Dennet, D. C. (1984). *Elbow Room*, Oxford Univ. Press, UK.
- Dennet, D. C. (1989). *The Intentional Stance*. MIT Press, Cambridge, Ma, US
- Gettier, E. (1963). *Is Justified True Belief Knowledge?* *Analysis*, **23**, 121–123.
- Hargreaves Heap S.P. and Y. Varoufakis (1995). *Game Theory: a critical introduction*. Routledge, New York.
- Monderer, D. and L. S. Shapley (1996). *Potential Games*. *Games and Economic Behaviour* 14-0044: 124143
- Moulin, H. (1986). *Game Theory for the Social Science*. New York Univ. Press, Washington Sq. New York.
- Pearce, D. (1984). *Rationalizable Strategic Behavior and the Problem of Perfection*. *Econometrica*, **52**, 1029–1050.
- Sastry, P., V. Phansalkar and M. A. L. Thathachar (1994). *Decentralized Learning of Nash Equilibria in Multi-person Stochastic Games with Incomplete Information*. *IEEE Trans. Sys., Man, Cybern.*, **24-5**, 769–777.
- Thathachar, M. A. L. and P. S. Sastry (2004). *Networks of Learning Automata*. Kluwer Academic Publ., Norwell, MA, USA.