

## On the Evolutionary Fitness of Bounded Rationality: Heterogeneous Populations in Antagonistic Interactions

Tassos Patokos

UADPhilEcon, University of Athens, 14 Evripidou Street, Athens 10559, Greece

---

**Abstract:** Conventional game theory assumes hyper-rational players, while evolutionary game theory abandons the assumption. This paper studies what happens when agents of both profiles co-exist and get engaged in a series of antagonistic interactions (the Hawk-Dove game). It is shown that if rational agents are perfectly informed as to the type of their opponent, they find it optimal to always be aggressive (that is, always select “Hawk”) when paired with an irrational player. It is then shown that, generally, a similar result is also valid when rational agents fail to recognise the type of their opponent with certainty. Finally, a discussion on why it may be fruitful to consider populations heterogeneous as to the rationality of agents is provided.

**Key words:** Evolutionary game, antagonistic interactions, rational agents, heterogeneous populations

---

### INTRODUCTION

Conventional non-cooperative game theory assumes players who are endowed with admirable mental attributes. The word “rational”, usually employed in economics to describe mere utility-maximisers, is expanded in meaning and implies agents who are more than just “reasonable”; they are, to put it bluntly, overly *intelligent*. Such players know that they are facing equally sophisticated opponents and this common knowledge of rationality is somewhat reminiscent of the infinite reflection that happens when two mirrors are placed the one in front of the other. Not unusually, agents are also thought of as able to align their beliefs with those of their opponents and hold the same probabilistic expectations about the strategies to be chosen. If this seems as attaching too much faith in the intellectual capacities of real people, one partly satisfying response is to think that resolution of static games comes with logical-as opposed to historical-time. One period in logical time can be conveniently long, allowing players to think through their best reply to the opponent’s best reply and hence, it makes no difference if some agent is any “slow” to be rational. On the other hand, this is only partly a pleasing answer, because there is nothing to guarantee that people will actually choose what the theory predicts, no matter how much time they are given to think.

More often than not, this discussion is overshadowed by what indisputably is the primary trouble within game theoretic circles: Most games have multiple equilibria, which reduces the power of the theory for accurate predictions anyway, regardless of whether individuals are implausibly modelled as being “too rational”. The efforts

of trying to conjure up newer refinements may easily distract one from the fact that the rationality the theory imposes is generally too restrictive for the theory to successfully apply to real-life interactions. Empirical data frequently contradict even the simplest of games-like, for example, the prisoners’ dilemma, leading to questions as to what good is the theory for, if the rational agents it presupposes are not really living in the real world.

The argument hardly needs a formal treatment; it suffices to observe people all around to be convinced that the rationality in question is definitely not a trait of the average person. Choosing people at random and asking them to participate in some game, one would feel that several among them may even have trouble to understand the game in the first place, let alone reach a decision based on common knowledge of rationality and consistently aligned beliefs. This does not necessarily stand as a claim that people are not intelligent; the point here is that players are too diverse to be all covered with such a demanding assumption concerning their behaviour. Besides, knowledge of game theory itself can make a difference, for it is fair to assert that being conscious of a Nash equilibrium would probably make a player choose it with more confidence than an equally intelligent player who would be unaware of it.

All such issues concerning rationality seemingly resolve themselves with evolutionary game theory. The rationality assumption is not just relaxed, it is abandoned. Agents are thought of as somehow driven by animal instincts, only trying to get the most they can out of a game. Reason can no longer guide them into thinking what other players will do and the only criterion on how to play is to see how well they fared in

previous rounds and adjust their behaviour accordingly, by trial and error. Rational agents are therefore replaced by players who try to avoid unsuccessful choices and mimic rewarding strategies, in an attempt to end up being as well off as possible and secure their “evolutionary fitness”. Apparently, this newer framework no longer uses logical time, for an equilibrium unfolds in historical time, as agents’ current decisions are determined by what has happened in previous period interactions.

Given the contrast in the assumptions of the two theories with regards to how players behave, the fact that all evolutionary stable equilibria are Nash equilibria but not all Nash equilibria are evolutionary stable is rather startling. This implies that, even if one assumes players only motivated by imitation of successful behaviour and their instinct for survival, the evolutionary process makes them behave as if they were overly intelligent. The fact that rational behaviour is not necessary for reaching a sophisticated equilibrium could make some think that the assumed rationality of conventional game theory is not then too important in the first place; to quote Ken Binmore in his foreword to Weibull’s ‘Evolutionary Game Theory’<sup>[1]</sup> “...insects can hardly be said to think at all and so rationality cannot be crucial if game theory somehow manages to predict their behavior”. There is however no real paradox here; evolutionary theory may not require rationality, but it enforces a somewhat stricter behavioural code than the conventional theory (where there is nothing to guarantee the players’ “success”) and in addition to that, the introduction of historical time makes the informational setting a lot richer.

If evolutionary game theory addresses the criticism that the assumptions of the conventional theory on the rationality of humans are too demanding, it itself lies on an other extreme; not having faith that players can always be sophisticated is one thing, but considering them to behave like birds or ants is quite another. As Mailath<sup>[2]</sup> put it in an issue of Journal of Economic Theory devoted to evolutionary game theory in its entirety, “the [evolutionary] models rely on the players being implausibly stupid. Why are the players not able to figure out what the modeler can?”. The suggestion here is that it would be more realistic to consider an “in-between” scenario, that is, to assume that both types of agents can co-exist and see what the theory would have to say for interactions among them. After all, people clearly do not all possess the same mental and thinking abilities, which can point to the necessity for modelling agents as an heterogeneous population as to their rationality. Real-life games show that players actually condition their behaviour on who they think their opponent is; one would not play a game of, say, backgammon in the same way if

one played with a computer, a 6-year old kid or a skilled opponent. Introducing such “degrees of rationality” in the theory seems likely to produce more realistic conclusions about games.

**Rational players and automata playing Hawk-Dove under certainty:** Under study is the symmetric Hawk-Dove game for two players. It is an antagonistic game where the two opponents find themselves both contending the same good. Each player can either choose to be aggressive (and be a “hawk”) or retreat (and be a “dove”). If both players are aggressive, the good is eventually destroyed and fighting has both players suffer some loss. If both retreat, then they share the good and thus, they both enjoy some benefit. If the one fights and the other retreats, then the “hawk” gets all the reward, while the “dove” is left with nothing. Letting the payoff for the player who selected “Dove” in a “Hawk-Dove” outcome be zero makes the analytical task easier, without any sacrifice in generality (what matters is the relative ordering of the payoffs and hence, one among them may be chosen at random and the resulting payoff matrix will still reflect every possible Hawk-Dove interaction). Also, psychological considerations (such as feelings of fairness explored in Rabin<sup>[3]</sup>) that could potentially alter the payoffs of the game shall not be of concern here. In matrix notation, the game can, in its general form, be represented as follows:

	H	D
H	-L, -L	g,0
D	0,g	v,v

It is assumed that  $L > 0$ ,  $g > 0$ ,  $v > 0$ . Also, for this game to be a Hawk-Dove game, it is required that  $g > v$  (otherwise, D would always be a dominant strategy). The one-off version of this game has two Nash equilibria in pure strategies (H,D) and (D,H) and one Nash equilibrium in mixed strategies, where players choose H with probability  $p = (g-v)/(g-v+L)$  and D with probability  $1-p$ . In evolutionary game theory, when the population is homogeneous, there is a unique evolutionary stable equilibrium that coincides with the Nash equilibrium in mixed strategies. This means that either a fraction of the population equal to  $p$  always choose H and the remaining  $1-p$  always choose D, or that agents in that population choose H and D with probabilities  $p$  and  $1-p$  respectively. (Of these two interpretations, the former is favoured over the latter by researchers and is often referred to as the “purification view”, analysed, among others, in Harsanyi<sup>[4]</sup> and Fudenberg and Kreps<sup>[5]</sup>. The experimental findings of Friedman<sup>[6]</sup> are consistent with this view).

An intuitive interpretation of the evolutionary stable equilibrium is that in a homogeneous population, there must exist enough doves for being a hawk to eventually pay off, which accordingly means that a population consisting of hawks (doves) in its entirety is threatened by invasion by doves (hawks), as the intruders shall initially outwin their peers, until the Nash equilibrium in mixed strategies is reached.

This equilibrium is valid when the population consists of agents whose strategies have little to do with rationality, but who are only motivated by a will to secure “evolutionary fitness”, that is, perform well in the game, having the outcome of previous rounds as a kind of guide. The setting changes when this assumption is dropped and the population is taken to consist of rational agents-in the conventional game theory sense-too. The irrational players are still choosing their strategies by adaptation and imitation of successful behaviours, while the rational players have full knowledge of how their opponents form their strategies and can adjust their best replies accordingly.

More precisely, the population is taken to consist of a fraction  $a$  of rational players and a fraction  $1-a$  of “automata”, that is, agents who fulfill the non-rationality assumption of evolutionary game theory. The former will often be referred to as players of “type A”, while the latter will be referred to as “type B”. Type A players are fully rational and they are aware of the fact that if the population consisted of players of type B only, the evolutionary process would result in the Nash equilibrium in mixed strategies. It is also assumed that players of type A can recognise what type of player they are interacting with. This last assumption is not as implausible as it may at first seem: It is common for people to know who their co-player is and how strong an opponent they make. This assumption shall be dropped later in the text, where the rational players will be assumed to not know with certainty what type of opponent they are interacting with. Players of type B are thought of as totally unaware of the heterogeneity, i.e., of the existence of type a players in the population.

In this newer setting, there are three possible kinds of interactions: Interactions between one player of type A and one player of type B, interactions between players of type B and interactions between players of type A. Among the aforementioned kinds of meetings, only the latter is undetermined; because of the Folk Theorem, there can be no prediction as to an equilibrium when rational agents are engaged in a repeated game that has no definite ending. Neither is it justified to assume they choose the Nash equilibrium in mixed strategies, because they simply have no reason to: Once one player believes their opponent will choose the

Nash equilibrium in mixed strategies, then any choice of strategy makes them indifferent and thus, the choice of the Nash equilibrium in mixed strategies is as good as any other strategy. The bottom line is that any behaviour of agents of type A in meetings between themselves can be rationalised and therefore, what happens in these interactions has to be treated as a “black box”.

In contrast, when players of type B are called to play the game, their behaviour is known. These agents are motivated by evolutionary success and therefore, their choice of strategy is expected to approach the evolutionary stable point, that is, the Nash equilibrium in mixed strategies. The complication here is that agents of type A know this and they are able to use this knowledge in their favour. More precisely, rational players know that the more aggressive players of type A are against players of type B, then the less “hawks” there will exist in the population of players of type B. In more technical terms, at meetings between one player of type A and one player of type B, if players of type A choose H with probability  $r$ , then players of type B will choose H with probability  $q = q(r)$  and  $dq/dr < 0$ . Because  $q$  is known by players of type A, they can choose the optimal  $r = r^*$  such that it generates the response  $q^*$  that will maximise type A players’ expected returns.

Calculation of  $q(r)$  is straightforward: given that a fraction  $r$  of type A players choose H and that a fraction  $q$  of type B players choose H, then the expected return for a player of type B from choosing H is:

$$ER_{B,H} = a \cdot (-Lr + g \cdot (1-r)) + (1-a) \cdot (-Lq + g \cdot (1-q))$$

The first term of the above equation is relevant to meetings of type B players with type A players and the second term refers to meetings between agents belonging to type B. Each term has obviously been multiplied by its relevant population weight, for it also stands for the probability for each meeting to happen. The implicit assumption here is that the population (denoted  $N$ ) is sufficiently large so that  $aN \approx aN \pm 1$ . This assumption also ensures that players meet new opponents in each round and therefore, no previous history between specific agents can possibly affect their choices. As will be seen below, the populations’ sizes do not affect the decisions of rational players and thus, this assumption does not cause any analytical problems.

In similar fashion, the expected return for a player of type B by choosing D is:

$$ER_{B,D} = av \cdot (1-r) + (1-a) \cdot v \cdot (1-q)$$

Equating  $ER_{B,H}$  with  $ER_{B,D}$  gives the Nash equilibrium in mixed strategies, which is what players of type B shall opt for, given the existence of a type A “hawks” in the population. By doing the calculus, the resulting formula is:

$$q(r) = \frac{g - v - (g - v + L) \cdot ar}{(g - v + L) \cdot (1 - a)} \quad (1)$$

As was expected,  $q$  is a strictly decreasing function of  $r$  (for  $g-v+L$  is always positive). Also, if  $a = 0$  (that is, if there are only type B players in the population), then  $q$  expectedly collapses to the formula for  $p$  given above, which is the evolutionary stable strategy for the case of homogeneous populations (and, of course, the Nash equilibrium in mixed strategies).

With  $q(r)$  being known as the response of type B players in equilibrium, type A players can choose  $r$  so as to maximise their own expected returns. Because, like it was discussed above, meetings between rational players have indeterminate resolutions, players of type A can only maximise expected returns from their interactions with players of type B. The maximisation problem can be written as:

$$\max_r \{r \cdot (1 - a) \cdot [-Lq(r) + g \cdot (1 - q(r))] + (1 - r) \cdot (1 - a) \cdot v \cdot (1 - q(r))\}$$

The first term of the objective function is the expected return for a rational player when they play H against a type B player, multiplied with the probability  $r$  that they shall play H. Likewise, the second term refers to type A players' expected return when they choose D against a type B player. Because the non-negative term  $1-a$  appears in both these terms, it can be omitted from the calculus. Thus, the maximisation problem gets written Eq. (2):

$$\max_r \{r \cdot [-Lq(r) + g \cdot (1 - q(r))] + (1 - r)v \cdot (1 - q(r))\} \quad (2)$$

By substituting the above equation (1) for  $q(r)$  into the objective function and doing the calculus, the maximisation problem gets equivalently written as Eq. (3):

$$\max_r \{(L + g - v) \cdot r^2 + (2v - g) \cdot r\} \quad (3)$$

It is interesting to note that the population factor  $a$  (which appears in the  $q(r)$  formula) is not present in the objective function of maximisation problem (3) and thus, the solution will not depend on how large the population of rational agents is.

It is easy to see that the solution of problem (3) is  $r^*=1$ , which means that agents of type A maximise their expected returns in games with players of type B when they always choose H, for any configuration of payoffs  $L$ ,  $g$  and  $v$  and regardless of what their population  $a$  is (a simple proof is provided in Appendix A). Type B players' reaction then is:

$$q^* = q(r^*) = \frac{g - v - (g - v + L) \cdot a}{(g - v + L) \cdot (1 - a)} \quad (4)$$

**Implications:** With  $r^*$  being equal to 1, type B agents always have to face rational agents who choose H. How aggressive will type B agents be shall naturally depend on the size of type A players' population. It is easy to see that the derivative of  $q$  with regards to  $a$  is always negative, which confirms the intuitively obvious claim that the larger the population of rational agents, the less aggressive type B players will be. As was said above, when  $a = 0$  then  $q^* = p$  (the homogeneous case); this value of  $q^*$  is the largest possible to be attained in equilibrium. Also, because  $q^* \rightarrow -\infty$  as  $a \rightarrow 1$  and  $q^*$  cannot take negative values, there will be a value of  $a = \bar{a}$  to minimise  $q^*$ , that is, to make it zero. Solving the equation  $q^*=0$  for  $a$ , it can easily be found that  $\bar{a} = p$ . Evidently, for any  $a \geq \bar{a}$ ,  $q^* = 0$ . Thus, for populations of rational agents that are larger than the Nash equilibrium in mixed strategies, type B agents learn to always play D, even in meetings where both players are of type B (which is not surprising, since type B players do not understand the heterogeneity).

Rational players both favour and dislike their population getting larger. On the one side, a larger value for  $a$  means that type B agents are less aggressive. On the other side, with a bigger value for  $a$ , it becomes less probable for a rational player to interact with a player of type B and thus, the aforementioned advantage progressively weakens. This suggests that there is a value for  $a$ , denoted  $a^*$ , which maximises type A players' expected returns in meetings with type B opponents, given that the former always play H. To find  $a^*$ , it is easy to solve the maximisation problem again, but with regards to  $a$  and by setting  $r = r^* = 1$ . When  $a \geq \bar{a}$ ,  $q^*$  will be zero, which means that the problem simply becomes:

$$\max_a \{(1 - a) \cdot g\} \quad (5)$$

Obviously, in this case,  $a^* = \min\{a\} = \bar{a}$ ; this means that once the rational players' population becomes larger than  $\bar{a} = p$ , any increase in type A players' population results in a decrease of expected returns. This result confirms the intuition that any increase of  $a$  when  $a \geq \bar{a}$  shall only generate loss for type A players, for type B players cannot get any less aggressive ( $q^*$  is kept equal to zero) and at the same time, they become less likely to be encountered.

If  $a \leq \bar{a}$ ,  $q^* = q(r^*)$  is given by Eq. (4) derived above. With  $r = r^* = 1$ , the maximisation problem is:

$$\max_a \{(1 - a) \cdot [-Lq(r^*) + g \cdot (1 - q(r^*))]\} \quad (6)$$

By substituting formula (4) for  $q^*$  and doing some calculus, the problem comes down to maximising a linearly increasing function of  $a$ , which means that  $a^* = \max\{a\} = \bar{a} = p$ . Because the two cases give identical results, it follows that  $a^* = p$  for any initial level of  $a$ . When the population of rational agents is less than  $p$ , then type A agents favour their population getting larger, but once it hits  $\bar{a}$ , the rational players want it to stay there. When the former applies, type A players may affect the size of their population by converting a fraction of type B agents to type A (for instance, by making them aware of the heterogeneity and teaching them how to be rational)-provided, of course, that mobility across populations is possible. If, however, the population of type A players surpasses the Nash equilibrium in mixed strategies threshold, type A players have an incentive to keep their advantage a secret, because if more agents adhere to their population, the expected returns shall be lower. After that point, rational agents get engaged in a situation which recalls, in some sense, the prisoners' dilemma: for each one of them it is individually optimal to keep always playing H when interacting with a type B opponent, although the payoff everyone gets becomes progressively smaller as  $a$  increases. The phenomenon does not disappear but in the limit, that is, when  $a = 1$ , which means that the heterogeneity is no longer and that the whole population consists of rational agents only. Naturally, with no type B players, rational players cannot enjoy any advantage and how the interactions will be carried out is now totally left to the Folk Theorem and as a consequence, is anyone's guess.

For the above analysis, it was assumed that rational agents have perfect information; also, all that was said referred to equilibrium behaviour. Before studying what happens when there is uncertainty, a short qualitative discussion on how the initial conditions might affect the equilibrium follows.

**A note about out of equilibrium behaviour:**

Everything that was said above is relevant with what happens in equilibrium. For the analysis to be valid, it is implicitly assumed that type B agents will respond to type A players' strategy automatically (through the  $q$  function); nothing, however, has been said about the speed of this process. Type A agents are rationally expecting that type B players will adopt the predicted behaviour, but this may take a long while before it happens (that is, the evolutionary process might bring about the predicted  $q$  in the very long run).

In the model studied above, it is immediately clear that if  $q > p$ , then the rational players are better off if

they play D rather than H, because (as can also be seen from inspection of maximisation problem (2) above), their expected returns from playing H against a type B agent are less than the expected returns from playing D. If, therefore, type A players are not persistent enough and willing to suffer some loss now in order to gain more in the future, this shall lead to a decrease of  $r$ . Because  $q > r$ ,  $q$  will tend to decrease, but because  $r$  decreases too, there shall also be a tendency for  $q$  to increase. In other words, if type A players do not see to the problem intertemporally, the evolutionary process that would naturally have  $q$  decrease until it is equal to  $p$  will slow down. In the end, the equilibrium may be exactly the inverse of what was studied previously: rational players never play H when they interact with a type B player. There are three remarks to be made here:

- The equilibrium described above is unstable; for when ultimately  $q = p$ , then the rational agents will be indifferent as to whether they shall play H or D against a type B opponent. Therefore, even if a tiny fraction of them start play H,  $r$  shall become positive (from zero that it previously was) and  $q$  will decrease, which means that henceforth, it will be optimal for *all* rational agents to play H and  $r$  shall eventually become equal to 1
- Even when  $q > p$ , type A players may still rationally choose to play H, if they expect that the intertemporal gains when  $q(r)$  stabilises to  $q(r^*) = q(1)$  shall be larger than the losses in the current and near-future periods. Naturally, it comes down to the speed of the evolutionary process: if this process is slow enough, then type A agents shall not find it worthwhile to suffer the loss now (let alone if a discount factor is used). Nevertheless, as was said in the previous remark, once  $q = p$ , rational players will take over
- The interesting observation here is that type A players generally have no way of knowing the initial value of  $q$ . If this kind of uncertainty applies, then rational agents' expectations about current value of  $q$  determines their behaviour. If  $E(q) > p$ , then type A players may or may not want to take some expected loss now; but once  $E(q)$  equals  $p$  (or becomes lower than  $p$ ), then this will prompt  $r$  to rise. How can  $E(q)$  be formulated? The safest way is by observing how type B agents actually behave in the games. Thus, the hyper-rational type A agents may actually have to resort to a pattern of behaviour that type B players typically adopt: adjusting their behaviour based on what they observe to have happened in previous interactions.

**Introducing uncertainty:**

**Rational players and automata playing Hawk-Dove under uncertainty:** Uncertainty can have numerous sources; in this particular setting, it only affects type A agents, for type B players are supposed to act on instinct and heuristics anyway. Rational players might not fully know the payoffs of the game, they may not know the size of the population  $a$ , or, they may not be certain as to what type of player they are interacting with. More radically, they may not know the dimension of the heterogeneity, but more complex cases that involve more types of agents than just two are out of the scope of this text (a brief comment on more complicated scenarios appears later).

Uncertainty as to  $L$ ,  $g$  and  $v$  is of no concern here. It was previously seen that, as long as the interaction is a symmetric Hawk-Dove game (which means that  $L > 0$ ,  $g > v > 0$ ), type A players' optimal strategy ( $r = r^*$ ) remains the same, no matter the specific values of  $L$ ,  $g$  and  $v$ . In cases where rational players are uncertain as to the type of their opponent, it will be shown that uncertainty as to the payoffs is of small importance. (Naturally, if the game is not symmetric, then the whole setting changes, but the non-symmetric Hawk-Dove game will not be under study here). Similarly, it was shown that, in the perfect information scenario, knowledge of the size of population  $a$  does not affect rational players' decision. Even if type A agents were capable of affecting  $a$ , this would by no means alter the conclusion that the optimal strategy for type A players is to always play H against an irrational opponent. A similar result is true even for games when rational agents cannot tell with certainty if they are interacting with a rational player or not.

When rational players are aware of the heterogeneity but cannot fully recognise what type of player their opponent is, the interaction needs to be modelled from scratch; it can be assumed that rational agents recognise rational agents with probability  $x$  and irrational players with probability  $y$ . If, in any possible game, a rational player thinks they are interacting with a rational player with the same probability (regardless of what type the opponent really is), then  $x = 1 - y$ , which is a special case. In general, for each game, exactly one of the following four distinct cases shall apply:

- A rational player successfully recognises their rational opponent
- A rational player thinks they are playing with an irrational opponent, while, in reality, they are facing a rational player
- A rational player successfully recognises their irrational opponent, or

- A rational player thinks they are playing with a rational opponent, while, in reality, they are facing an irrational player

Because the population of rational agents is  $a$ , the aforementioned cases happen with probabilities  $ax$ ,  $a(1-x)$ ,  $(1-a) \cdot y$  and  $(1-a) \cdot (1-y)$  respectively.

Rational players choose their strategies according to who they think they are playing with. However, they are now aware that, in some games (namely, case b above), the behaviour they intend for irrational opponents shall actually be received by rational opponents and that, in some other games (case d above), the behaviour they intend for rational opponents shall eventually be directed to irrational opponents. It is assumed that rational players choose to be aggressive with probability  $d$  when they think they are interacting with rational opponents (that is, in cases a and d above) and that they choose to be aggressive with probability  $r$  when they think they are meeting irrational opponents (cases b and c). At least two fundamental differences from the certainty case can be seen here: for one thing, it is not necessarily optimal for  $r$  to be equal to 1, because this increases the probability of conflict (a "Hawk-Hawk" outcome) between two rational players. Also, the behaviour that rational players intend for rational co-players shall unavoidably be affecting the irrational population too, which means that both  $r$  and  $d$  are now needed as control variables (and not just  $r$ , like was the case under certainty).

To find the reaction function of irrational players, the same procedure as above will be followed; it is assumed that irrational players choose H with probability  $q$ . The expected return for an irrational player from choosing H can be split into three terms, for the cases (i) they are meeting a rational player who fails to recognise them as irrational, (ii) they are meeting a rational player who correctly recognises them as irrational and (iii) they are meeting with an irrational co-player. Given the notation explained above, it is:

$$ER_{B,H} = ay \cdot (-Lr + g \cdot (1-r)) + a \cdot (1-y) \cdot (-Ld + g \cdot (1-d)) + (1-a) \cdot (-Lq + g \cdot (1-q))$$

Similarly, the expected return for a player of type B by choosing D is:

$$ER_{B,D} = ayv \cdot (1-r) + a \cdot (1-y) \cdot v \cdot (1-d) + (1-a) \cdot v \cdot (1-q)$$

Equating  $ER_{B,H}$  with  $ER_{B,D}$ , as before, gives the equilibrium response of type B players to type A players' behaviours  $d$  and  $r$ :

$$q(d,r) = \frac{g-v-a \cdot (g-v+L) \cdot [yr+(1-y) \cdot d]}{(g-v+L) \cdot (1-a)} \quad (7)$$

As it was expected, when  $d = r$  or  $y = 1$  (or both),  $q(d,r)$  collapses to the formula for  $q(r)$  (equation (1)) found in the certainty case (because the uncertainty collapses and, as a result, the control variables vector is no longer in two dimensions).

Switching to rational players, their expected return from choosing H can again be split into three terms, to capture the situations when a rational player selects H (i) against a rational opponent who correctly thinks they are playing with a rational co-player, (ii) against a rational opponent who mistook them for an irrational co-player and (iii) against an irrational opponent. Thus:

$$ER_{A,H} = ax \cdot (-Ld + g \cdot (1-d)) + a \cdot (1-x) \cdot (-Lr + g \cdot (1-r)) + (1-a) \cdot (-Lq + g \cdot (1-q))$$

Similarly, the expected return for a player of type A by choosing D is:

$$ER_{A,D} = axv \cdot (1-d) + a \cdot (1-x) \cdot v \cdot (1-r) + (1-a) \cdot v \cdot (1-q)$$

Rational players choose H with probability  $d$  whenever they think they are facing a rational opponent, that is, with probability  $ax + (1-a) \cdot (1-y)$ . They also choose H with probability  $r$  whenever they think they are facing a type B player, that is, with probability  $a \cdot (1-x) + (1-a) \cdot y$ . Thus, H is played with probability  $f = d \cdot [ax + (1-a) \cdot (1-y)] + r \cdot [a \cdot (1-x) + (1-a) \cdot y]$  and D is played with probability  $1-f$ . Thus, the maximisation problem is:

$$\max_{d,r} \{ f \cdot ER_{A,H} + (1-f) \cdot ER_{A,D} \} \quad (8)$$

After the calculus and after substituting Eq. (7) found above for  $q = q(d,r)$ , the maximisation problem comes down to:

$$\max_{d,r} \{ Ar^2 + Bd^2 + Cdr + Dr + Ed + F \} \quad (9)$$

Where:

$$\begin{aligned} A &= (1-x-y) \cdot [-a \cdot (1-x-y) - y] \cdot a \cdot (L+g-v), \\ B &= (1-x-y) \cdot [-a \cdot (1-x-y) - y + 1] \cdot a \cdot (L+g-v), \\ C &= (1-x-y) \cdot [2a \cdot (1-x-y) + 2y - 1] \cdot a \cdot (L+g-v), \\ D &= -av \cdot (1-x-y), \\ E &= av \cdot (1-x-y) \text{ and} \\ F &= vL / (L+g-v). \end{aligned}$$

The full solution of the maximisation problem is given in Appendix B. The conclusion is reproduced here; to save on notation, let  $M = -a \cdot (1-x-y) \cdot y$  and  $\Delta = L+g-v > 0$ :

a. When  $1-x-y = 0$ , the objective function is a constant.

b. When  $1-x-y > 0$ , then:

b<sub>1</sub>. If  $M \geq (-v-\Delta)/2\Delta$ , then  $d^* = 1$  and  $r^* = 0$ .

b<sub>2</sub>. If  $M < (-v-\Delta)/2\Delta$ , then  $d^* = 1$  and  $r^* = 1+(v+\Delta)/2M\Delta$ .

c. When  $1-x-y < 0$ , then:

c<sub>1</sub>. If  $M > (v-\Delta)/2\Delta$ , then

$d^* = (\Delta-v+2M\Delta)/2 \cdot (M+1) \cdot \Delta$  and  $r^* = 1$ .

c<sub>2</sub>. If  $M \leq (v-\Delta)/2\Delta$ , then  $d^* = 0$  and  $r^* = 1$ .

It can be seen that the decisions of rational players are primarily influenced by  $x$  and  $y$ . If  $1-x-y$  is positive (negative), then a rational agent finds it optimal to always select "Hawk" when they think they are meeting a rational (an irrational) opponent and play "Dove" or a mixed strategy when they think they are meeting an irrational (rational) opponent; what of the latter two shall happen in each case depends on the configuration of the payoffs, the population  $a$  and  $x$  and  $y$ . The result confirms that the payoffs of the game and the population are not crucial in determining the best action of a rational player; they do not determine whether the player will fall into category a, b or c above, but only if they must choose between  $b_1$  or  $b_2$  in the case of b and  $c_1$  or  $c_2$  in the case of c.

The next part of this section analyses and discusses some implications of the above result.

**Implications of uncertainty:** Because  $x$  and  $y$  are the probabilities of successful recognition of rational and irrational agents respectively, it is evident that type A agents are more able to recognise their opponent correctly, as the expression  $1-x-y$  becomes lower. If indeed  $1-x-y$  becomes negative, it is optimal for type A agents to always select "Hawk" against an opponent whom they think is irrational, which is the same policy encountered in the certainty case. On the other hand, if successful recognitions are not frequent enough and  $1-x-y > 0$ , then the optimal thing for type A agents to do is to always select "Hawk" when they think they are playing with a rational co-player.

It is interesting to focus a bit on the case  $1-x-y = 0$ , which reflects a situation where the rational players recognise other rational players with the same probability in every possible meeting, no matter what the opponents truly are. That is, there is no additional information available-like, for example, something in the appearance of the agents signalling different odds for being either of type A or of type B-to make rational players recognise one of the two types more or less successfully than the other; if this is the case, then it shall be that  $x = 1-y$  and thus,  $1-x-y = 0$ . When this happens, the expected return is always equal to  $vL/(L+g-v) > 0$ , whatever the chosen strategies. Intuitively, this means that opposite effects cancel out (for example,

unexpected-or to put it better, unintended-“Hawk-Hawk” outcomes are cancelled out by unexpected “Hawk-Dove” outcomes between rational players).

With total uncertainty about types and no asymmetries between them, if the population size of type A players  $a$  is known, then rational players can use it as a probability of successful recognition. More precisely,  $x = a$  and  $y = 1-a$ . In this case,  $1-x-y$  shall again be zero and rational players will be indifferent as to their choice of strategies. Even with population size being unknown, rational players can use expectations and the same result shall apply:  $x = E(a)$  and  $y = 1-E(a)$ , which again gives  $1-x-y = 0$ .  $E(a)$  need not be the same across type A players; even if they formulate different expectations, the conclusion that they remain indifferent as to the choice of strategies is safe.

For  $1-x-y$  to be non-zero, there must be some trait to distinguish type A and type B agents, known to rational players, such that successful recognition of the former happens less or more frequently than successful recognition of the latter. In other words, rational players must be able to formulate  $x$  and  $y$  based on some information that distinguishes the two possible types of their opponents; for example, if  $a = 1/2$  and rational players know that half of the irrational players are black-haired, then when rational players encounter a black-haired opponent, they recognise the opponent as rational with probability  $1/2$  and as irrational with probability  $1/4$ . In this example,  $1-x-y = 1/4$  and Case b of the conclusions presented in the previous section applies. (This situation presents no “information gap” for type A players, as at first might seem. It could be argued that, in the particular example, rational players are called to play  $r^*$  in the  $1/4$  of the interactions and  $d^*$  in the  $1/2$  of the interactions, but what happens in the remaining  $1/4$ ? If the randomisation starts anew, then in that 25% probability of allegedly having an information gap, the rational agents could again think the opponent is rational with probability  $1/2$  and irrational with probability  $1/4$ . But this would elevate the probability of thinking the opponent is rational to  $5/8$  ( $1/2+1/8$ ) and the probability of thinking the opponent is irrational to  $5/16$  ( $1/4+1/16$ ) and there would still be an information gap in the  $1/16$  of the interactions. If the iterations continue likewise, somehow reminding of the structure of a fractal, then eventually the information gap would vanish and  $1-x-y$  would become zero. This could be a tricky argument claiming that it would never make sense for  $1-x-y$  to be anything else than zero; the mistake with this logic overlooks the fact that type A players shall actually play  $d^*$  not in  $1/2$  of the interactions, but in  $ax+(1-a)\cdot(1-y)$  of the interactions and play  $r^*$  in  $a\cdot(1-x)+(1-a)\cdot y$  of the interactions (in the

specific example, these probabilities are  $5/8$  and  $3/8$  respectively). For  $ax+(1-a)\cdot(1-y)+a\cdot(1-x)+(1-a)\cdot y = 1$ , there is ultimately no such “information gap”).

It is important to underline here that  $x$  and  $y$  are objective probabilities-as opposed to subjective beliefs. Whether a rational player believes, for whatever idiosyncratic reason, that they are meeting with a rational agent with some probability  $x$  or  $y$  is irrelevant; by construction, the maximisation problem involves objective probabilities for  $x$  and  $y$  (for if they are just based on beliefs and these beliefs are not confirmed, then the maximisation shall fail). Another issue with probabilities  $x$  and  $y$  is that they need not be the same for all rational agents; type A players might calculate  $x$  and  $y$  based on different information they have about the population. This means that, unlike the certainty case where the whole population of rational agents had a unique best action, strategies here may differ across rational agents. This makes type A agents’ behaviour harder to predict, for some of them shall fall within case b of the solution presented above, while some others will fall within case c, which is case b’s stark opposite (not to mention that anything goes if  $1-x-y = 0$ ). However, except for the special (but nonetheless very reasonable) case where  $1-x-y = 0$ , any strategic choice of rational players is in fact an attempt to make advantage of type B players’ adaptive behaviour. That the rational players may choose to play always “Hawk” when they think they are playing with a rational opponent is in fact an act of aggression towards the irrational players, for this happens when the rational players are not relatively successful in recognising the types of their opponents and as a consequence, these “Hawk” strategies shall eventually be directed to irrational players more frequently than not.

The probabilities  $x$  and  $y$  have so far been treated as exogenous. Their values are crucial in determining the optimal choices of  $d$  and  $r$  in the maximisation problem (9) above. However, it might not be altogether true that type A agents cannot change  $x$  and  $y$  at all. After all,  $x$  and  $y$  are formulated by the information available to rational agents and as such, they can be altered with the (conscious) acquisition of more information relevant to the identification of the opponent’s type.

What makes this discussion interesting in this particular setting is that a quick inspection of (9) reveals that different configurations of  $x$  and  $y$  yield different expected returns. It can easily be seen that when  $1-x-y = 0$ , the expected return for rational agents (equal to  $vL/\Delta$ ) is the minimum they can get, if compared to the other cases where  $1-x-y$  is non-zero. In other words, type A players can always do better if  $x$  and  $y$  are such that  $1-x-y$  is not zero. This means that the case where rational agents recognise rational opponents regardless of what the opponents really are



(that is, when  $x = 1-y$ ) is always inferior to a case where these recognition probabilities are different. Thus, rational players have an incentive to avoid  $x = 1-y$ ; whether this will actually entail more success in recognising the type of the opponent correctly does not matter, for anything other than  $x = 1-y$  shall be preferred to  $x = 1-y$ . As a consequence, type A agents shall try to condition their recognition performance (by acquiring information) in such a way that  $1-x-y$  is either positive or negative and regardless of whether this will make them better at actually recognising the type of an opponent successfully. To put it in a more concrete way, the colour of an agent's hair may mean nothing about how rational they are, but it pays for rational agents to know that, for example, more irrational than rational agents have black hair, because this makes  $x$  different from  $1-y$ .

And if, for instance, it happens (by mere accident) that there are more irrational agents with black hair than rational agents with black hair, agents with black hair receive different treatment than others. This happens because type A players enjoy greater expected returns when they base their recognition performance on any piece of information that may distinguish the two populations and for doing that, any arbitrary trait that shall stand as yet another heterogeneity will do. In other words, except for the heterogeneity in rationality, rational agents favour additional dimensions of heterogeneity within the population, because they help them fare better in the interactions. And if such additional dimensions of heterogeneity do not really exist, they might as well be "artificially" sought up by the rational agents themselves, for they are devices that help type A agents increase their expected payoffs: there need not be any correlation between the colour of one's hair and if one is actually rational; all that suffices is an observation (which may as well be an accident at that) that instructs rational players on how to recognise rational and irrational players differently. This perhaps can explain why people are eager to discriminate on one another and are quick to deduce general rules that are based on one's characteristics-and then adjust their behaviour accordingly. And if these beliefs are also communicated between agents (that is, if the reasons why  $x$  and  $y$  take some particular values are common knowledge), it is then no wonder that social groups usually come with labels attached.

**More heterogeneity:** Mixing rational with irrational agents makes sense only partially, for both profiles are extremes and in-between situations are also likely to exist. A player may not be hyper-rational and able to arrive at the above maximisation conclusions, but this does not necessarily mean that they will have the behavioural profile of an ant or a bee. A more realistic scenario would therefore involve more "levels of

rationality" and thus, greater heterogeneity within the population as to the rationality of the agents.

Modelling such a scenario would inevitably be analytically hard, but a possible way out would be to implement it with a computer simulation. Such a task would, of course, involve making explicit assumptions as to the speed of learning and the probabilities of "mutations" (if positive at all); also, the dimension of the heterogeneity would have to be predetermined.

A potentially promising way to implement this would be to assume an *ordered* range of populations, each one of them "more rational" than the ones before them. The population at the top of the rationality scale would behave like type A agents of this study, while at the other extreme would lie type B agents. A population  $i$  between these two extremes would behave like type A agents with probability  $p_i$  and like type B agents with probability  $1-p_i$  and therefore, a population  $k$  would be considered as "more rational" than population  $m$  if  $p_k > p_m$ . In the most general case, this rationality space could be continuous, with an infinite number of different populations. Each population  $i$  would acknowledge (with certainty or with probabilities of successful recognition like) the existence of all other populations less rational than  $i$ , but would be "blind" as to the existence of the more rational. A mutation here would be a (random or otherwise) accession of an agent belonging to population  $i$  to a population higher than  $i$  in the rationality scale. Such mutations being possible, a convergence to a homogeneous population where all agents are of type A is expected. To revisit the certainty case studied in the first part of this text, if type B agents became aware of type A agents' strategic advantage, they would start acting like them and the heterogeneity would disappear (but at the same time, so would the strategic advantage).

A related work in this direction is Camerer *et al.*<sup>[7]</sup>. The authors distinguish agents as to the degree of the common knowledge of rationality they can perceive and offer an empirically driven and dynamic learning model. After having parameterised bounded rationality, one of their concerns is to fit the model with data and pin down the frequencies of players with different rationality bounds within populations. Apparently, their results are valuable for identifying plausible initial conditions for an evolutionary analysis, should one want to work with a heterogeneous population as in the model presented in this text.

However, according to what meaning one would want to attach to rationality (which obviously needn't involve any kind of maximisation), the task of ordering different rationality profiles does not seem an easy one. Generally, the above discussion makes obvious that there are several decisions to be made while modelling a setting like this and that such theoretical

pronouncements are bound to have a big impact on the according conclusions, if social conjectures are being sought after. Clearly, an evolutionary setting like the one studied here cannot solely rely on standard biology and the need for substantial input from social sciences, such as sociology, anthropology or even history, is more than evident. Sugden<sup>[8]</sup> provides a comprehensive discussion on why it is inadequate for evolutionary game theory to remain attached to biology-and the relevant mathematics). Can the modeller allow that type B agents can learn from the behaviour of type A agents and, if yes, how exactly this learning process can be assumed to work, given that type A agents do not always have an incentive to support it? What may change if type B agents do not become any more rational but learn to tell if their opponent is type A or type B? If type B agents are not just automata but real human beings, would it not be more realistic to assume that mutations in one direction are less or more likely than others? Questions like these come with no short answers, for it might be appropriate to address them in one way for some social context and then in another should this context change. It seems that an ad-hoc approach of games is inevitable, a need that was already clear in game-theoretic circles, at least since Schelling's<sup>[9]</sup> classic "*The strategy of conflict*".

## CONCLUSION

The introduction of this text reproduced a comment by Mailath<sup>[2]</sup> which expresses the view that portraying individuals as unsophisticated is not too realistic. This is not considered to be really problematic; in fact, Mailath<sup>[10]</sup> argues that concerns about EvGT's realism are "*misplaced*". He writes that "*the role of models is to improve our intuition and to deepen our understanding of how particular economic or strategic forces interact. [...] The games are intended as examples, experiments and allegories. Modelers do not make assumptions of bounded rationality because they believe players are stupid, but rather that players are not as sophisticated as our models generally assume*". Of course, the modeller is all the way justified to dismiss the hyper-rational agent of conventional game theory, but it is quite a whole different thing to replace this profile with that of an agent who is clueless about their surroundings and can learn or imitate only in the specific ways suggested by the selection dynamics under use. The point here is that, on the way to substitute a lot less demanding profile for the hyper-rational individual, the theory came up with a set of presumptions that can be challenged just as easily. Mailath asserts that this is inevitable if the modeller wants "*simple and tractable games [...] that can be solved*". To be sure, nobody can deny that simple and

simplified models can yield interesting and insightful conclusions, but on the other hand, it is not really fruitful to rest on these models' ideal settings and not explore more complicated and realistic scenarios only because the latter may cause analytical problems.

The discomfort with both the neoclassical paradigm and the evolutionary models lies not so much in that their assumptions are implausible on their own right; the problem is rather that the theories assume a certain type of economic agents and then consider any deviation as a special and uninteresting case. Thus, while there certainly can be individuals who, for example, act in accordance with rational choice theory, or agents who fall into the category of bounded rationality (whatever its form), none of these profiles can exclusively be used as an assumption for a generic enough framework. After all, biologists themselves never claimed that their evolutionary models apply to all species; in a recent retrospective article, Smith<sup>[11]</sup> seems almost hard-pressed to actually name species that behave according to the predictions of the evolutionary models, concluding that these models are good for pronouncing some qualitative conjectures, which can *sometimes* (as opposed to "in most cases") get confirmed.

The central suggestion here has been that a promising way for putting more realism in evolutionary models is to let different rationality profiles co-exist and model the individuals under study as an heterogeneous population, consisting of individuals that differ in rationality. In the model presented above, the heterogeneity altered the predictions of the conventional analysis and, to the eyes of type A players, it made the antagonistic interaction seem more like a prisoners' dilemma kind of game. Naturally, this is not to say that the particular model is realistic to any satisfactory degree, but it knowingly seems as a step towards acknowledging the need for not taking one specific rationality profile for granted. Its greatest merit is that it began by a more accurate view of the population's rationality (albeit the arbitrariness in deciding upon the initial state) and offered a few findings that would not occur otherwise. In addition to this, such a venture can obviously trigger fruitful discussions on what rationality really is, in how many different ways one can be rational or how this rationality may evolve, let alone lead to positive conclusions relevant with emerging conventions and discriminatory phenomena.

From a methodological perspective, the primary motives for suggesting heterogeneity in rationality are flexibility and generality. Since behaviour of individuals is unarguably context specific, it is not hard to find reasons for disagreeing with the predictions of most models that presuppose a specific behavioural profile and to come up with counterexamples or empirical paradoxes. On the contrary, allowing for parameterisation of rational behaviour leads to models

with broader applicability, that can easily be customised for many games. Of course, there is a trade-off: enabling this kind of interweaving between EvGT and behavioural game theory is bound to increase the theory's level of abstraction and make the problem of indeterminacy more acute. This is a logical compromise for models that have more explanatory power and stay in line with the findings of experimental studies.

As its analogies with mechanics have it, economics sees individuals as lifeless particles of matter, their actions being directed by preference orderings and utility functions. The parallel with biology gives economic agents more credit and endows them with a conscience and instincts, albeit animal-like. This is certainly better, but not good enough: fitness for animals cannot easily be translated in terms of fitness for humans, whose motives, expectations, mental abilities, moral codes and levels of sophistication are so variable and ambiguous. It then follows that the models of biologists are generally in need of substantial and radical amendments before claiming to apply to the world of humans. Tractable and simplifying models may have their own theoretical value and mathematical elegance, but for an efficient study of human interactions and evolution, there can be no easier way other than trying to instill to these models some of the complexity of human nature.

**Appendix A:** Problem (3) is an inequality constrained maximisation problem, the constraints being that  $r$  must lie within  $[0,1]$ . The problem can be easily solved without resorting to the Lagrangian or writing down the Karush-Kuhn-Tucker conditions: Because the second derivative of the objective function is equal to  $2(L+0-v)>0$ , the first-order condition shall give a minimum rather than a maximum. By writing down the first-order condition and temporarily forgetting about the constraints, this minimum, denoted  $\bar{r}$ , is given by the formula:

$$\bar{r} = \frac{g - 2v}{2 \cdot (L + g - v)}$$

Obviously, if  $r > \bar{r}$  then the objective function is increasing with  $r$ , while, if  $r < \bar{r}$  then the objective function is decreasing with  $r$ . Because  $r \in [0,1]$ , one of the following three cases shall apply:

- $\bar{r} > 1$ : This means that  $f$  is decreasing everywhere in  $[0,1]$ , which means that  $r^* = \min\{r\} = 0$ . But  $\bar{r} > 1$  means that  $g - 2v > 2(L + g - v)$  or that  $g > 2L$  which cannot hold, since both  $g$  and  $L$  are positive. Thus, this case is not valid.
- $\bar{r} < 0$ : This means that  $f$  is increasing everywhere in  $[0,1]$ , which means that  $r^* = \max\{r\} = 1$ . This happens when  $g - 2v < 0$ .

- $0 \leq \bar{r} \leq 1$ : This means that the minimum lies somewhere in the interval  $[0,1]$  and hence, the objective function is first decreasing and then increasing within this interval. Thus, the objective function is maximised either when  $r = \max\{r\} = 1$  or when  $r = \min\{r\} = 0$ . But when  $r = 1$  the objective function gives  $L + g - v + 2v - g = L + v > 0$ , while when  $r = 0$  the objective function becomes zero. Therefore,  $r^* = 1$ .
- Case a is invalidated, while both cases b and c give the solution that  $r^* = 1$ .

**Appendix B:** This appendix presents the solution for the following maximisation problem:

$$\max_{d,r} \{Ar^2 + Bd^2 + Cdr + Dr + Ed + F\}$$

Where:

$$\begin{aligned} A &= (1-x-y) \cdot [-a \cdot (1-x-y) - y] \cdot a \cdot (L+g-v), \\ B &= (1-x-y) \cdot [-a \cdot (1-x-y) - y + 1] \cdot a \cdot (L+g-v), \\ C &= (1-x-y) \cdot [2a \cdot (1-x-y) + 2y - 1] \cdot a \cdot (L+g-v), \\ D &= -av \cdot (1-x-y), \\ E &= av \cdot (1-x-y) \text{ and} \\ F &= vL / (L+g-v). \end{aligned}$$

The coefficient  $F$  can be dropped as a constant. Also, all other coefficients  $A, B, C, D, E$  include the term  $1-x-y$  and thus this term can be dropped off all coefficients, as long as it is not zero. Interestingly, if  $1-x-y = 0$ , then the objective function is a constant, which means that type A players are indifferent as to their choice of  $d$  and  $r$ , for their expected returns are always the same. Thus, when  $1-x-y = 0$ , or  $x+y = 1$ , then there is indeterminacy and any strategic profile can be an equilibrium.

To simplify notation,  $M$  is used to denote the expression  $-a \cdot (1-x-y) - y$  and  $\Delta$  is used to denote  $L+g-v > 0$ . By simplifying the term  $a$  from all coefficients and by assuming  $1-x-y > 0$ , the maximisation problem becomes:

$$\max_{d,r} \{M\Delta r^2 + (M+1) \cdot \Delta d^2 - (2M+1) \cdot \Delta dr - vr + vd\} \quad (A1)$$

There are four constraints to this problem, of which at most two can be binding:

$$r \geq 0, r \leq 1, d \geq 0, d \leq 1 \quad (A2)$$

The associated Lagrangian is:

$$L = M\Delta r^2 + (M+1) \cdot \Delta d^2 - (2M+1) \cdot \Delta dr - vr + vd + \lambda_1 r - \lambda_2 \cdot (r-1) + \lambda_3 d - \lambda_4 \cdot (d-1) \quad (A3)$$

The first-order conditions are:

$$\frac{\partial L}{\partial r} = 2M\Delta r - (2M+1) \cdot \Delta d - v + \lambda_1 - \lambda_2 = 0 \quad (A4)$$

$$\frac{\partial L}{\partial d} = 2 \cdot (M+1) \cdot \Delta d - (2M+1) \cdot \Delta r + v + \lambda_3 - \lambda_4 = 0 \quad (A5)$$

$$\lambda_1 \cdot r = 0 \quad (A6)$$

$$\lambda_2 \cdot (r-1) = 0 \quad (A7)$$

$$\lambda_3 \cdot d = 0 \quad (A8)$$

$$\lambda_4 \cdot (d-1) = 0 \quad (A9)$$

$$r \geq 0, r \leq 1, d \geq 0, d \leq 1 \quad (A2)$$

$$\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0 \quad (A10)$$

Because at most two of the constraints (A2) can be binding, the following cases can be examined separately for potentially giving a solution:

- $r = 0, d = 0$
- $r = 1, d = 0$
- $r = 0, d = 1$
- $r = 1, d = 1$
- $r = 0, 0 < d < 1$
- $r = 1, 0 < d < 1$
- $0 < r < 1, d = 0$
- $0 < r < 1, d = 1$
- $0 < r < 1, 0 < d < 1$  (the unconstrained problem)

Before proceeding to each of the above cases, it is useful to notice that  $M = -a \cdot (1-x-y) \cdot y$  can alternatively be written as  $M = -a \cdot (1-x) \cdot (1-a) \cdot y$ , which obviously means that:

$$-1 \leq M \leq 0, \text{ or } 0 \leq M+1 \leq 1 \quad (A11)$$

**Case 1:** If  $r = 0$  and  $d = 0$ , then (A7) and (A9) mean that  $\lambda_2 = \lambda_4 = 0$ . Then, (A5) gives  $\lambda_3 = -v < 0$ , which contradicts (A10). Thus,  $r = 0$  and  $d = 0$  is not a solution for (A1).

**Case 2:** If  $r = 1$  and  $d = 0$ , then (A6) and (A9) mean that  $\lambda_1 = \lambda_4 = 0$ . Then, (A4) gives  $2M\Delta - v - \lambda_2 = 0$ , which cannot hold, because of (A10) and (A11). Thus,  $r = 1$  and  $d = 0$  is not a solution for (A1).

**Case 3:** If  $r = 0$  and  $d = 1$ , then (A7) and (A8) mean that  $\lambda_2 = \lambda_3 = 0$ . Then:

$$(A4) \Rightarrow -(2M+1) \cdot \Delta - v + \lambda_1 = 0 \text{ and} \\ (A5) \Rightarrow 2 \cdot (M+1) \cdot \Delta + v - \lambda_4 = 0$$

Solving for  $\lambda_1, \lambda_4$ , these two last equations yield  $\lambda_1 = (2M+1) \cdot \Delta + v$  and  $\lambda_4 = 2 \cdot (M+1) \cdot \Delta + v$ . Because of (A11), the latter of these expressions satisfies (A10). For the former expression to be valid, it is required that  $(2M+1) \cdot \Delta + v \geq 0$ ,

or  $M \geq (-v-\Delta)/2\Delta$ . If this condition applies, then  $r = 0$  and  $d = 1$  is a candidate solution for being a maximum.

To check the second-order condition, the bordered Hessian is:

$$A = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 2 \cdot (M+1) \cdot \Delta & -(2M+1) \\ -1 & 0 & -(2M+1) & 2M\Delta \end{bmatrix}$$

whose determinant is  $|A| = 1 > 0$ , which means that the solution is a maximum indeed.

**Case 4:** If  $r = 1$  and  $d = 1$ , then (A6) and (A8) mean that  $\lambda_1 = \lambda_3 = 0$ . Then, (A4) gives:

$$2M\Delta - (2M+1) \cdot \Delta - v - \lambda_2 = 0, \text{ or } \lambda_2 = -v - \Delta < 0, \text{ which} \\ \text{contradicts (A10). Thus, } r = 1 \text{ and } d = 1 \text{ is not a} \\ \text{solution for (A1)}$$

**Case 5:** If  $r = 0$  and  $0 < d < 1$ , then (A7), (A8) and (A9) mean that  $\lambda_2 = \lambda_3 = \lambda_4 = 0$ . Then:

$$(A4) \Rightarrow -(2M+1) \cdot \Delta d - v + \lambda_1 = 0 \text{ and} \\ (A5) \Rightarrow 2 \cdot (M+1) \cdot \Delta d + v = 0$$

Adding these two last equations by parts yields  $\Delta d + \lambda_1 = 0$ , which can only hold if  $d = 0$  and  $\lambda_1 = 0$ , which contradicts the starting assumption. Thus,  $r = 0$  and  $0 < d < 1$  is not a solution for (A1).

**Case 6:** If  $r = 1$  and  $0 < d < 1$ , then (A6), (A8) and (A9) mean that  $\lambda_1 = \lambda_3 = \lambda_4 = 0$ . Then:

$$(A4) \Rightarrow 2M\Delta - (2M+1) \cdot \Delta d - v - \lambda_2 = 0 \text{ and} \\ (A5) \Rightarrow 2 \cdot (M+1) \cdot \Delta d - (2M+1) \cdot \Delta + v = 0$$

Adding these two last equations by parts yields  $d = 1 + \lambda_2/\Delta$ , which can only satisfy (A2) if  $\lambda_2 = 0$ . This would however mean that  $d = 1$ , which contradicts the starting assumption. Thus,  $r = 1$  and  $0 < d < 1$  is not a solution for (A1).

**Case 7:** If  $0 < r < 1$  and  $d = 0$ , then (A6), (A7) and (A9) mean that  $\lambda_1 = \lambda_2 = \lambda_4 = 0$ . Then:

$$(A4) \Rightarrow 2M\Delta r - v = 0, \text{ which contradicts (A11). Thus, } d \\ = 0 \text{ and } 0 < r < 1 \text{ is not a solution for (A1)}$$

**Case 8:** If  $0 < r < 1$  and  $d = 1$ , then (A6), (A7) and (A8) mean that  $\lambda_1 = \lambda_2 = \lambda_3 = 0$ . Then:

$$(A4) \Rightarrow 2M\Delta r - (2M+1) \cdot \Delta - v = 0 \text{ and} \\ (A5) \Rightarrow 2 \cdot (M+1) \cdot \Delta - (2M+1) \cdot \Delta r + v - \lambda_4 = 0$$

First of all, (A4) implies that  $M$  is non-zero. Solving (A4) for  $r$  gives that  $r = 1 + (v+\Delta)/2M\Delta$ . For this

result to be valid,  $(v+\Delta)/2M\Delta$  must be in the interval  $(-1,0)$ . The condition that must apply for this to happen is that  $M < (-v-\Delta)/2\Delta$ .

To check the second-order conditions, the bordered Hessian is the same as shown in Case 3, bar the first row and the first column (since there is no binding constraint on  $r$  here). Then, the determinant is  $|A| = -2M\Delta$ , which is positive because of (A11) and because  $M$  is non-zero. Thus, again, the solution is a maximum.

**Case 9:** If  $0 < r < 1$  and  $0 < d < 1$ , then (A6), (A7), (A8) and (A9) mean that  $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ . Then:

$$(A4) = > 2M\Delta r - (2M+1)\Delta d - v = 0$$

$$(A5) = > 2\cdot(M+1)\Delta d - (2M+1)\Delta r + v = 0$$

Summing these last two equations by parts yields  $d = r$ . By substituting  $d = r$  into any of these equations gives that  $\Delta r = -v$ , which cannot hold. Thus,  $0 < r < 1$  and  $0 < d < 1$  is not a solution for (A1).

If  $1-x-y < 0$ , the problem must be resolved anew, with the signs of (A1) reversed. If the same procedure is used, it is easy to find, that there is, again, a unique solution, featuring (not unexpectedly) a kind of symmetry if compared with the solution found above: More specifically, in the case where  $1-x-y < 0$ , if  $M > (v-\Delta)/2\Delta$ , then  $r^* = 1$  and  $d^* = (\Delta - v + 2M\Delta)/2\cdot(M+1)\Delta$ , while if  $M \leq (v-\Delta)/2\Delta$ , then  $r^* = 1$  and  $d^* = 0$ .

Thus, to sum it all up:

- a. When  $1-x-y = 0$ , the objective function is a constant.
- b. When  $1-x-y > 0$ , then:
  - b<sub>1</sub>. If  $M \geq (-v-\Delta)/2\Delta$ , then  $d^* = 1$  and  $r^* = 0$ .
  - b<sub>2</sub>. If  $M < (-v-\Delta)/2\Delta$ , then  $d^* = 1$  and  $r^* = 1 + (v+\Delta)/2M\Delta$ .
- c. When  $1-x-y < 0$ , then:
  - c<sub>1</sub>. If  $M > (v-\Delta)/2\Delta$ , then  $d^* = (\Delta - v + 2M\Delta)/2\cdot(M+1)\Delta$  and  $r^* = 1$ .
  - c<sub>2</sub>. If  $M \leq (v-\Delta)/2\Delta$ , then  $d^* = 0$  and  $r^* = 1$ .

**Appendix C:** What follows is a list of the symbols used in the text, for quick reference.

- L: the payoff for player  $i$  when  $i$  selects H and  $j$  selects H
- g: the payoff for player  $i$  when  $i$  selects H and  $j$  selects D
- v: the payoff for player  $i$  when  $i$  selects D and  $j$  selects D
- p: the Nash equilibrium in mixed strategies
- q: the probability with which an irrational player chooses H
- r: the probability with which a rational player chooses H when they think they are playing with an irrational player (or when, in fact, they are

actually playing with an irrational player, when there is no uncertainty)

- d: the probability with which a rational player chooses H when they think they are playing with a rational player
- x: the probability with which a rational player successfully recognises a rational opponent
- y: the probability with which a rational player successfully recognises an irrational opponent
- a: the fraction of rational agents within the whole population
- $\Delta$ : introduced for notation purposes; it is defined as  $\Delta = g - v + L$
- M: introduced for notation purposes; it is defined as  $M = -a \cdot (1-x-y) - y$

## REFERENCES

1. Weibull, J.W., 1997. *Evolutionary Game Theory*. 1st Edn., MIT Press, Cambridge, ISBN-10: 0262731215, pp: 265.
2. Mailath, G.J., 1992. Introduction: Symposium on evolutionary game theory. *J. Econom. Theory*, 57: 343-362. DOI: 10.1016/0022-0531(92)90036-H
3. Rabin, M., 1993. Incorporating fairness into game theory and economics. *Am. Econom. Rev.*, 83: 1281-302. <http://www.behaviouraldesignlab.org/latestpapers/incorporating-fairness/>
4. Harsanyi, J.C., 1973. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *Intl. J. Game Theory*, 2: 1-23. DOI: 10.1007/BF01737554
5. Fudenberg, D. and D. Kreps, 1993. Learning mixed equilibria. *Games Econom Behavior*, 5: 320-367. DOI: 10.1006/game.1993.1021
6. Friedman, D., 1996. Equilibrium in evolutionary games: Some experimental results. *Econom J.*, 106: 1-25. <http://ideas.repec.org/p/cla/levarc/393.html>
7. Camerer, C.F., T.H. Ho and J.K. Chong, 2001. Behavioral game theory: Thinking, learning and teaching. *J. Risk Uncertainty*, 19: 7-42. DOI: 10.2139/ssrn.295585
8. Sugden, R., 2001. The evolutionary turn in game theory. *J. Econom. Method.*, 8: 113-130. <http://classes.maxwell.syr.edu/ecn611/Sugden.pdf>
9. Schelling, T.C., 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge, pp: 309.
10. Mailath, G.J., 1998. Do people play Nash equilibrium? Lessons from evolutionary game theory. *J. Econom. Literature*, 36: 1347-1374. <http://ideas.repec.org/a/aea/jecolit/v36y1998i3p1347-1374.html>
11. Smith, J.M., 2003. *The Mathematics of Evolution. In: It Must be Beautiful: Great Expectations of Modern Science*, Farmelo, G., (Ed.), Granta Books, ISBN-10: 1862075557, pp: 284-284.