Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 16(5): 3718–3733
DOI: 10.3934/mbe.2019184
Received: 17 December 2018
Accepted: 15 April 2019
Published: 27 April 2019

Research article

Quantifying fairness to overcome selfishness: A behavioural model to describe the evolution and stabilization of inter-group bias using the Ultimatum Game

Andrea Guazzini[1,3], Enrico Imbimbo[1,3], Federica Stefanelli[1,3], Franco Bagnoli[2,3] and Ezio Venturino[4,*]

[1] University of Florence, Department of Education, Languages, Intercultures, Literatures and Psychology, Via di San Salvi, 12, Building 26, Florence, Italy
[2] University of Florence Department of Physics and Astronomy. Also INFN, Sez. Firenze. Via G. Sansone 1, 50019 Sesto Fiorentino, Italy
[3] University of Florence, Center for Study of Complex Dynamics (CSDC), Florence, Italy
[4] Department of Mathematics "Giuseppe Peano", University of Turin, via Carlo Alberto 10, 10123 Torino, Italy; Member of the INdAM research group GNCS

* **Correspondence:** Email: ezio.venturino@unito.it; Tel: +390116702833; Fax:+390112362833.

Abstract: The ability to form groups to overcome problems has been crucial for the evolution of human beings. To favour the formation of cooperating groups, one of the mechanisms developed is the inter-group bias, namely the tendency of individuals to favour members of their group and hinder the external ones. It is the cognitive equivalent of the "green beard effect" in evolutionary biology, introduced by Hamilton and popularized by Dawkins, for which a group can profit of the altruistic behaviour of its members. Here, we use a behavioural model based on the Ultimatum Game, to shed light on how this behaviour cloud has been stabilized in the human population, estimating the magnitude of favouritism needed to overcome selfish individuals. Through both numerical simulations and analytic approaches, we study how a community of collectivist and individualist agents evolves. The key factor is the mechanism for the evolution of the population, i.e., the replacement of the poor-performing individuals. In the case of replacement by the reproduction of existing individuals, we observe a smooth phase transition and no coexistence. If the replacement is random, the transition smooths, and coexistence is possible. We developed analytical approaches for these two cases and performed numerical simulations. Although analytical calculations support the behaviour emerging from simulations, some differences ask for more refined treatments.

Keywords: Inter-group favouritism; fairness; selfishness; agent-based modeling; Ultimatum Game; evolution; mathematical models; dynamical systems

1. Introduction

In-group favouritism [1], also known as "inter-group-bias" [2] or "indirect reciprocity" [3], appears one of the most stable and diffused social mechanism promoted by natural selection in many different species [4]. It is the cognitive equivalent of the "green beard effect" in evolutionary biology, first suggested by Hamilton [5, 6] and popularized by Dawkins [7], for which a group can profit of the altruistic behaviour of its members if they are able to recognize and discriminate the members of their own group. The first interesting results on the inter-group bias date back to the 1970's. They show how people tend to behave more positively and favour the members of their own group (in-group) rather than the others (out-group). Thus, cooperation inside the group is preferred to the one outside it [8]. The first explanations of this phenomenon led back the preference group's internal sharing and aid to social identity theory [9, 10], based on the observation that this phenomenon could also be found in minimal groups whose interactions did not take place face to face [11].

More recently alternative explanations of this phenomenon have been proposed, for which the inter-group favouritism was caused by the naive representation that individuals would have about the behaviour of groups. A significant feature of these beliefs is the fact that the group seems to be perceived as a real generalized exchange system, within which an indirect reciprocity rule exists, for which, if a subject helps a member of the group, sooner or later in return he will receive help by another one [12]. This hypothesis has been supported by recent simulation experiments. They show how the tendency to favouritism improves if the strategy used by an agent is visible, even if only indirectly or partially, to other members of the group [13].

Overall, the evolution of favouritism and cooperation through indirect reciprocity has been well investigated in recent years, both through mathematical models of experimental subjects, and by the use of game-theoretic tools. The dynamics underlying decision making [14] and the altruistic behaviour [15] have been especially studied through the Ultimatum Game (UG) [16]. In this game, a player, called the proposer, must share a good with another player, the receiver. The latter can accept or reject the proposal; acceptance determines the success of the negotiation, by refusing instead no player receives anything [17]. In this simple situation the best strategy for a player when playing as a receiver is to maximize the minimum gain (i.e accept any offer) and to minimise losses when playing as proposer (i.e always offer the minimum possible amount), [18].

Not surprisingly [19, 20], human beings, when playing an UG, usually do not follow the utility function described above, but they tend to offer more than necessary, in an average range that goes from 30 to 40 percent of the good that has to be shared and refuse on average 16% of all offers bid, [21]. One explanation for such a behaviour could be found in the principle of fairness [22, 23, 24, 25, 26]: the people engaged in the UG usually play by dividing the common good equally, and expect to receive the same treatment from the other players. For this reason, they do not accept low offers that are judged selfish and unjust, and they could punish those who have proposed them [27, 28, 29]. Moreover, in UG it has been shown that decisions taken in group are closer to the optimal behaviour compared to those taken individually [30, 31].

Furthermore, some interesting findings indicate several features that arise in the UG game: belonging to the same group, the same nationality, the same origin, the same race, eliciting group membership, are characteristics that, when recognized, change the behavior of the offerer.

In one experiment, students belonging to two different institutes (the Washington and Lee University and the Virginia Military Institute) competed in a UG. When the proposer knew the receiver's provenance, his behaviour changed, increasing the offer for a receiver from his same institution [32]. Similar results have been also found through the use of the dictator game (i.e a variant of the UG where the receiver cannot refuse the offer) [33].

Another set of results shows how people participating in a UG behave differently if they know that they are playing against a player of their same nationality. An experiment involving Malaysian and British engaged in a cross-cultural UG shows how Malaysian offered significantly more to other Malaysians than to British people [34]. Similarly French players have shown a preference for their own countrymen with respect to players coming from India, by choosing more equitable amounts to be split, close to 50% for people from their country [35].

Also, eliciting in some college students individual or group salience shows that those who have been subjected to a priming for the activation of a group identity have higher thresholds for both offers and acceptances (they offer more but also tend to demand more) to their in-group [36]. Racial affiliation (white or black) was evidenced by another experiment, where people tend to reject offers considered unfair more by people of their own race than those of the other [37].

Despite the huge amount of literature in the past decades devoted to in-group favouritism or inter-group-bias, still little is known about the "magnitude" that such effect should have had to bring human societies to the current "evolutionarily stable" state.

Modeling such a magnitude, as well as its dependencies from fundamental features of the system analyzed (e.g., dependence from the size of the population, sizes of the sub-populations, gap between the fitness advantages and costs related to be a "collectivist"), would allow to understand the inter-group favouritism complex dynamics, and support for instance many social engineering and psycho-social applications.

Evidence supports the claim that the human population consists of both collectivist people, willing to sacrifice themselves for the good of the group, and individualistic people, those more oriented to pursue their own interests at the expense of others. We are interested in understanding what is the magnitude of the in-group effect, i.e. the degree of favouritism among con-specifics, that allows the coexistence of these two types of people. Firstly, by setting up a numerical model we highlight the only possible magnitude of in-group effect allowing coexistence of these two opposite types of individuals. We simulate a very simple system using an Ultimatum Game where two different kinds of agents are implemented: individualists and collectivists. We are also interested in establishing a connection between a stochastic formulation of the system and its possible corresponding deterministic interpretation. Thus we translate the numerical model into a deterministic dynamical system. From its analysis, an evolutionary stable parameter ruling such a dynamics is obtained. The latter depends on the offers of the agents and their tendency to accept.

Our results reveal the existence of a smooth transition between the inter-group favouritism and the success of the group in the evolutionary stable state of the system. Such a transition appears to be independent to the size of the system. It seems related in a rather complex way to the initial density of "collectivists" in the population as well as to the fitness advantage provided by membership to the own group (i,e., individualists or collectivists).

2. The model

We model the interactions between agents involved in the Ultimatum Game, a non-cooperative game, in which two players interact. Their purpose is splitting an amount initially given to one of them by means of a repeated game. In each interaction, an agent called the offerer proposes to share a given amount of goods with a second agent, called the recipient. If the recipient accepts, the amount is actually partitioned among the two agents as proposed by the offerer; otherwise nobody gains anything.

Let N be the number of agents, partitioned into two sub-populations, the collectivists $C = cN$ and the individualists $I = (1 - c)N$. A game is made by a certain number T of tournaments, each one composed by M matches. In each match, each agent plays once as the offerer, in a random order, and may be randomly chosen as a recipient. Since a tournament is made by binary interactions, each agent plays as recipient once on average.

In the real UG game, offerers can dispose of a certain amount of money at each turn, part of which is offered to the recipient. So, the gain of each player for each turn is either positive or null. However, since the probability of accepting the offerer is fixed and does not change during the game, the total amount of money owned by plays increases at a fixed rate, and with respect to the average gain, the UG becomes a zero-sum game. Therefore, we can think that each player has a score, initially set to zero, that can become positive or negative with respect to the average gain. The amount of good (score) offered by individualists is q, an integer ranging from 1 to 10. The amount offered by collectivists depends on the type of the opponent: if it is a collectivist, the offerer is p, again an integer random from 1 to 10, otherwise it is p/h, where $1 \leq h \leq 4$ measures the degree by which collectivists try to oppose non-collectivists.

The probability that a collectivists accepts the offer is a, and that of individualists is b. In the following we always use $a = b = 1/2$.

At the end of a tournament, i.e., after M matches, a fraction m of agents with the lower score is replaced, either by random cloning the remaining individuals, or by immigration of a population composed, on average, by half collectivists and half individualists.

3. A simple mean-field analysis of the cloning replacement

Since the game is zero-sum, the average gain is zero. Let us try to compute the average gain for collectivists and individualists. The exchange between agents of the same class does not count.

When a collectivist plays the donor role against an individualist, it suffers a negative gain $G_C^{(CI)}$, given by the probability of having an individualist as a recipient ($I/(N-1)$) times the offer (p/h) times the acceptance probability (b). Therefore

$$G_C^{(CI)} = -\frac{I}{N-1}\frac{pb}{h} = -\frac{(1-c)N}{N-1}\frac{pb}{h}. \tag{3.1}$$

When a collectivist plays as a recipient against an individualists its gain G_{IC} is on average given by the probability of being chosen, $1/(N-1)$, times the number of individualists, I, times the offer, q, times the acceptance probability, a, i.e.,

$$G_C^{(IC)} = \frac{I}{N-1}qa = \frac{(1-c)N}{N-1}qa. \tag{3.2}$$

Similarly, $G_I^{(IC)}$, the negative gain of an individualist after its encounter with a collectivists is

$$G_I^{(IC)} = -\frac{C}{N-1}\frac{pb}{h} = -\frac{cN}{N-1}qa, \tag{3.3}$$

and the gain when playing as a recipient, $G_I^{(CI)}$, is

$$G_I^{(CI)} = -\frac{C}{N-1}qa = -\frac{cN}{N-1}\frac{pb}{h}. \tag{3.4}$$

So, finally, we have

$$\begin{aligned} G_C &= G_C^{(IC)} + G_C^{(CI)} = \frac{(1-c)N}{N-1}\left(qa - \frac{pb}{h}\right), \\ G_I &= G_I^{(CI)} + G_I^{(IC)} = -\frac{cN}{N-1}\left(qa - \frac{pb}{h}\right), \end{aligned} \tag{3.5}$$

It can be easily checked that the average gain, $cG_C + (1-c)G_I = 0$ is null.

In the cloning replacement, the equilibrium condition is given by the Competitive Exclusion Principle [38], i.e., the average gain of coexisting population is the same. Imposing $G_C = G_I$, the result is independent of c, therefore no coexistence is possible. The transition point is given by the condition $G_C > G_I$, i.e., by assuming $a = b$,

$$h > \frac{p}{q}. \tag{3.6}$$

Thus the population is composed by all individualists for values of h less then p/q, and all collectivists otherwise.

4. Simulations and numerical results

The simulations use $T = 1000$, and $M = 1000$, varying the population N from 10 to 1000, and the parameter h from 1 to 4 in steps of 0.5. The replacement factor is $m = 0.2$, i.e., the 20% of the population is replaced after each tournament. For the cloning replacement, the numerical results reproduce the mean-field prediction, with sharp (first-order) phase transition from $c = 0$ to $c = 1$ at $h^* = p/q$. For $p = 7$ and $q = 3$, the transition is located at $h^* \simeq 2.33$.

Let us now study the case of random replacement, in which the fraction m of replaced population is composed, on average, by half-collectivists and half-individualist.

4.1. Critical in-group effect (h)

In Figure 1 the relation between the average final percentage c of collectivists at equilibrium versus the in-group effect (h) is shown for different population sizes. The parameters a, b, p and q are chosen in a conservative way and in agreement with literature [32, 33]. The dependence of c on h is a smooth transition, with respect to the cloning replacement case. The populations cannot vanish and are maintained at a level near 10%, namely the replacement rate for each of them. The transition is continuous, being located near $h^* = 2.33$.

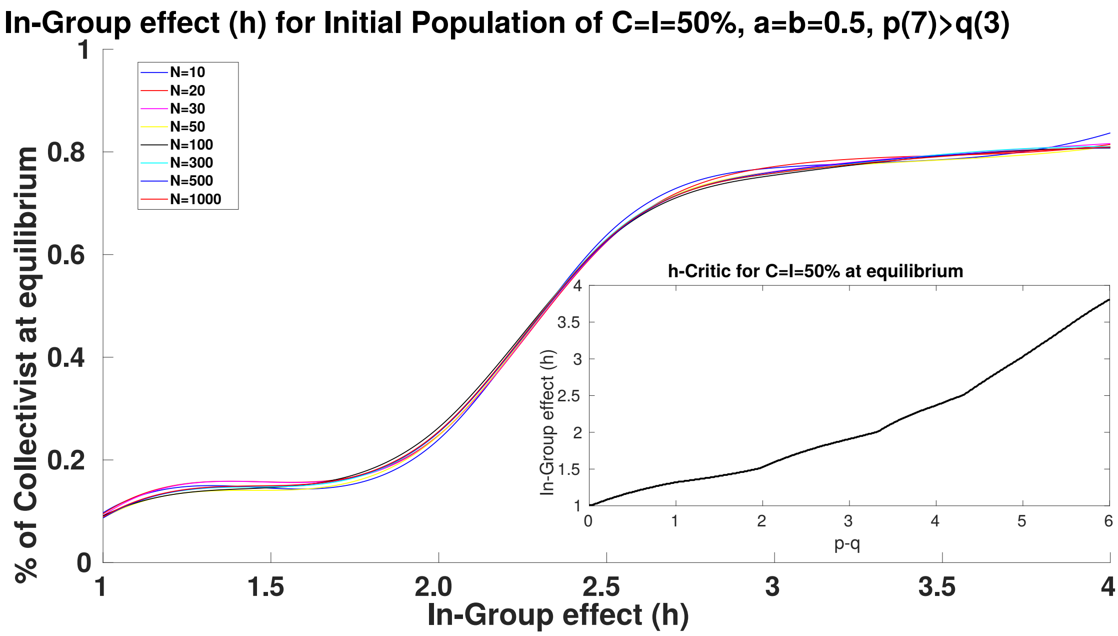

Figure 1. Smoothed phase transition. Main axes: percentage c of collectivists at equilibrium versus different values of h for an initial equal population of collectivists and individualists, for $a = b = 0.5$, $p = 7$ and $q = 3$. Subplot axes: critical values of h that lead to the final equilibrium of the population with a fifty-fifty distribution among C and I versus different values of the difference $p - q$, for $a = b = 0.5$.

The relationship emerging from Figure 1 suggests that, in the case of random replacement, there is always a coexistence between collectivists and individualists, with two zones characterized by a low-high asymptotic percentage of collectivists with respect to individualists. The results appear to be independent of N. We can divide the two zones fixing a separation threshold of 50%, from which one can get the separation value h^*, that depends on a, b, p and q.

The main frame of Figure 1 reports how different values of the in-group effect h influence different initial agent population sizes equally distributed between collectivist C and individualists I. The former have an acceptance probability $a = 0.5$ and an offer threshold $p = 7$; the latter have an acceptance probability $a = 0.5$ and an offer threshold $p = 3$. In these conditions, C agents need a factor $h > 2$ to effectively hinder the members of the selfish out-group and therefore to survive overcoming the I's. The effect of h on the final number of collectivists seems to be described by the same smooth transition regardless of the population size.

The subplot in Figure 1 shows the critical values of h^* for which it is possible to observe a fifty-fifty final distribution of C's and I's for various differences $p - q$ between offers of C (i.e., p) and I (i.e., q). Notably, from the subplot is it possible to appreciate that for values of $p - q = 4$ the in-group effect needed to reach the fifty-fifty condition at the equilibrium ($C = I$) is $2 < h < 2.5$. This result is consistent with the values reported in the main frame.

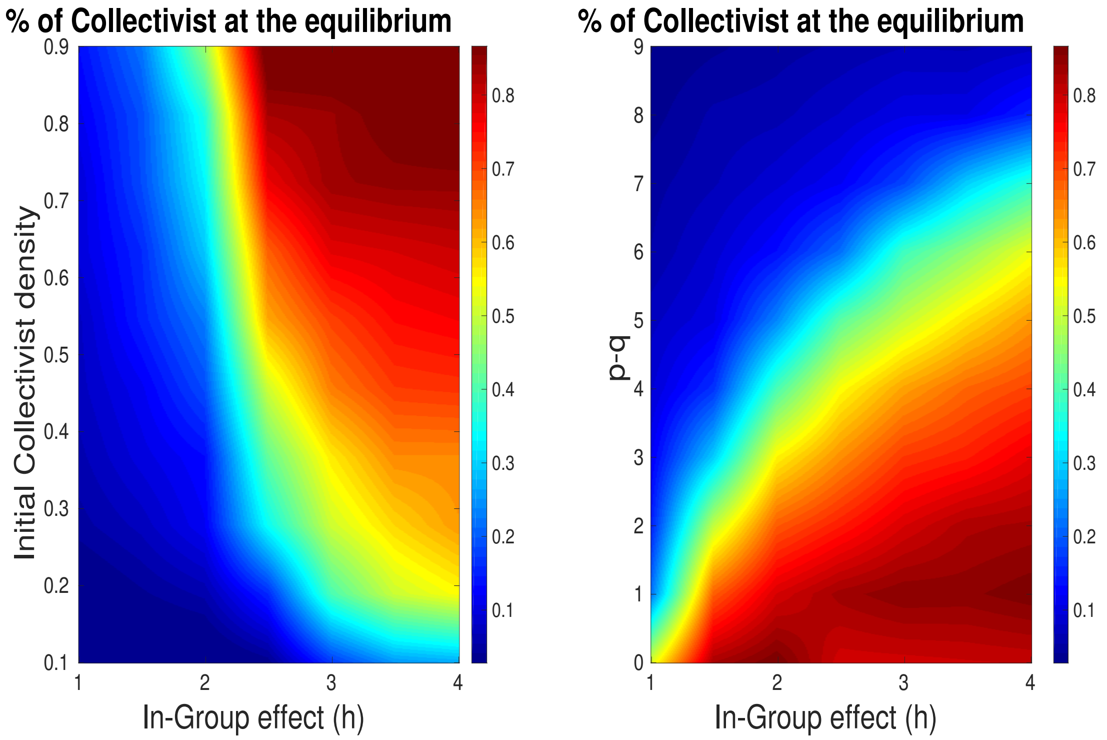

Figure 2. Percentage of collectivist at equilibrium for different initial density and different offers. Color range: red - large, green - medium, blue - small. Left panel: Collectivist percentage at equilibrium versus initial C density and values of h. Right panel: Collectivist percentage at equilibrium versus initial C density and values of the difference $p - q$.

4.2. Asymptotic density of collectivists

The effect of the initial C density and value of h on the C percentage at equilibrium has been accurately evaluated. The results are shown in the left panel of the Figure 2. For very small values of h individualist agents prevail at equilibrium, while increasing the initial number of C and for values of h larger than 2 collectivists start to increase their density at equilibrium. In the right panel of Figure 2 the results are presented on the combined effect of h and the difference between the offers of C and I agents. From this plot it appears that the larger the difference $p - q$ between the offers of C and I agents, the more the collectivists need to hinder the out-group members by increasing h, in order to survive.

To further examine the trends of collectivists density at the equilibrium, as function of the parameters investigated, we present in Figure 3 the effect of the amount offered by C agents and I agents for certain fixed values of a, b and h. We can observe that the larger the offer q of the individualists, the easier it is for collectivists to survive, regardless of their acceptance rate. Moreover, the more collectivists offer, the more difficult is their survival. They need to increase the hindering h inflicted to I agents in order not to become extinct. Furthermore, when both kinds of agents accept at the same rate, so that $a = b$, the higher the value of h is, the higher the surviving C population at equilibrium is for each parameters p and q taken into account. Instead, if collectivists accept with a higher probability

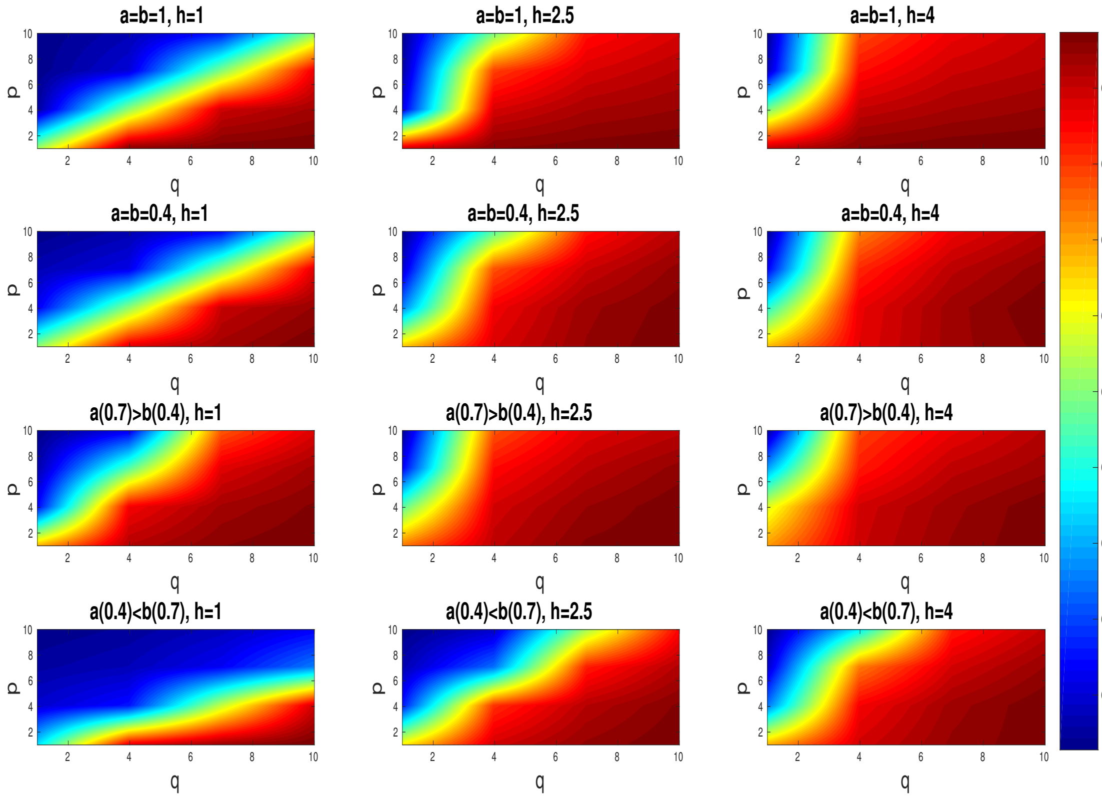

Figure 3. Density of collectivists varying p and q for different fixed values of a, b and h. Color range: red - large, green - medium, blue - small.

than individualists, so that $a > b$, they are able to survive and overcome individualists more frequently also when they offer more than the I players. Conversely, when the I agents accept with a higher probability than C, so that $a < b$, collectivists survive at a lower level and, in order not to disappear, they need to increase h as p and q change.

To examine the hindering role of C toward I, we show in Figure 4 the effect of h on the equilibrium density of collectivists for different values of agents offers, p and q, when the acceptance probabilities a and b are kept at fixed values. First of all, for all the reported values of p, q, a and b, the system attains a stable point for high values of h where collectivists dominate. This outcome is reached faster when $a > b$, namely when C agents accept more frequently than I agents, also for values of $h < 1.5$. Conversely, C's outperform I's in a slower way when their acceptance probability is smaller than that of the out-group members, i.e., $a < b$. In particular, in this situation, collectivists need values of $h > 3$ to overtake the number of individualists when $p > q$, namely when the C players offer more than the I ones.

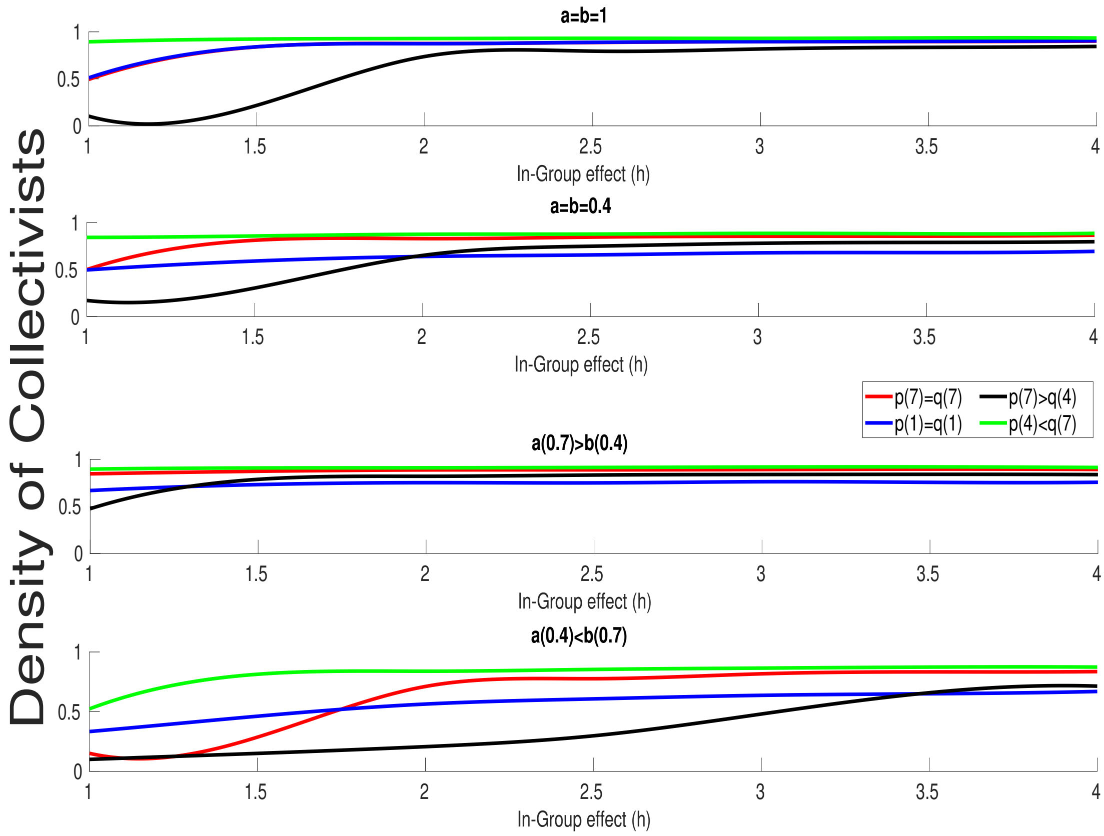

Figure 4. Effect of the in-group effect (h) on density of collectivists for fixed values of a and b.

5. A more refined dynamic approximation for the random replacement case

We now reformulate the problem with an equivalent deterministic approach based on dynamical systems. We consider the collectivists C and individualists I, with total constant population

$$C + I = N. \tag{5.1}$$

There is a constant influx per unit time of new individuals replacing an equal amount of individuals that are removed from the system, at rate m. New members enter instantaneously into either one of the two sets, with the same probability 0.5. The new individuals represent a fraction mN of the total population. Individuals are eliminated from each population following a rule on the reward system, $E_i(x)$, $i \in \{C, I\}$, where x represents the outcome of the rule, to be described below. Thus

$$\frac{dC}{dt} = \frac{1}{2}mN - mCE_C(x), \quad \frac{dI}{dt} = \frac{1}{2}mN - mIE_I(x). \tag{5.2}$$

The fundamental assumption is that the total population is constant, so that we must have $(C + I)' = 0$ and then by adding the equations in (5.2) we obtain

$$N = CE_C(x) + IE_I(x). \tag{5.3}$$

From (5.3) we thus conclude that

$$E_I(x) = \frac{1}{I}[N - CE_C(x)]. \tag{5.4}$$

Thus substituting (5.4) into the second equation of (5.2) the system becomes

$$\frac{dC}{dt} = \frac{1}{2}mN - mCE_C(x), \quad \frac{dI}{dt} = \frac{1}{2}mN - m[N - CE_C(x)]. \tag{5.5}$$

The "mortality" in each population has opposite effects, it is large for one and small for the other one depending on the sign of x. To model this situation we further assume the specific form for the elimination function E_C, which has to be non-negative, so that it is maximal for C when x is large and positive and minimal when x is large and negative:

$$E_C(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}. \tag{5.6}$$

Now we describe the rule for the reward that each population obtains. For this, the encounters between individuals are relevant. Since each individual can encounter C collectivist individuals and I individualist individuals, the total number of encounters that are possible are as follows. There are C^2 encounters between a collectivist and another collectivist, CI encounters between a collectivist and an individualist and as many between an individualist and a collectivist, and finally I^2 encounters among individualists. Let R_C and R_I denote the rewards obtained by each set of individuals during the encounters. To assess the reward 1, which instantaneously introduced into the system, we need to distinguish the two cases of the offering and the receiving individuals. Assume C is the offerer: let $0 \le p \le 1$ be the fraction of the amount offered to another collectivist and ph^{-1}, $h \ge 1$ the amount offered to an individualist. In such cases the offerer gets the amount that remains after the offer is accepted, respectively then $1 - p$ and $1 - ph^{-1}$. Let a be the rate at which a collectivist accept an offer, no matter from which individual, either C or I, is being offered. Let b the acceptance rate of an individualist, again independent of the offerer being C or I. The offer made by an individualist is always the same independently of the other individual encountered, being the fraction $0 \le q \le 1$ of the money instantaneously generated by the system. In case of refusal, no reward is given to anybody in any case. So we need only to concentrate on just the offers being accepted. Note that for the whole of collectivists, the total reward comes from the encounters of one of them with another one or with an individualist, the offer being accepted, or from the acceptance of an offer from an individualist:

$$R_C = [ap + (1-p)a]C^2 + b\left(1 - \frac{p}{h}\right)CI + aqIC = aC^2 + \left[aq + b\left(1 - \frac{p}{h}\right)\right]CI. \tag{5.7}$$

In the first part of this formula we have emphasized the offerer, writing it as the first variable in the interactions. With the same notation, for the individualist set we similarly find

$$R_I = (1-q)aIC + [(1-q)b + qb]I^2 + \frac{p}{h}bCI = bI^2 + \left[(1-q)a + \frac{p}{h}b\right]IC. \tag{5.8}$$

We can now calculate the average reward that each agent receives:

$$\frac{R_C}{C} = aC + \left[aq + b\left(1 - \frac{p}{h}\right)\right]I, \quad \frac{R_I}{I} = bI + \left[(1-q)a + \frac{p}{h}b\right]C. \tag{5.9}$$

The rule by which individuals, in particular C's, are instantaneously eliminated, is related to the discrepancy between the above two quantities. We therefore set

$$x = \frac{R_I}{I} - \frac{R_C}{C} = \left[\frac{p}{h}b - aq\right](C + I) = \rho(C + I) = \rho N$$

We can now use this expression in (5.6), and substitution into (5.5) leads to

$$\frac{dC}{dt} = \frac{1}{2}mN - mCE_C(\rho N) = m\left[\frac{1}{2}(N - C) - \frac{C}{\pi}\arctan(\rho N)\right] \tag{5.10}$$

Now, the dynamics of the system (5.5) occurs on the line $C + I = N$, and, what is the same, for the reduced equivalent equation (5.10) the equilibrium $E^* = (C^*, I^*)$, $I^* = N - C^*$, is obtained by annihilating the derivative, from which C^* is found

$$C^* = \frac{\pi N}{\pi + 2\arctan(\rho N)} \tag{5.11}$$

Note that also in this approach, as well as in the numerical simulations, the ratio C^*N^{-1} never attains the values 0 and 1. This equilibrium point is feasible if $N \geq C^* \geq 0$. Now observe that $\pi + 2\arctan(\rho N) > 0$ unconditionally. Thus we need to concentrate only on the upper bound. It turns out that it is satisfied if and only if

$$bp \geq ahq, \tag{5.12}$$

and whenever E^* is unfeasible, then

$$C^* > N. \tag{5.13}$$

Now, $2\pi > \pi + 2\arctan(\rho N)$, from which using this estimate in (5.11), we find:

$$C^* > \frac{1}{2}N. \tag{5.14}$$

Further $C' \geq 0$ if and only if

$$\frac{1}{2}(N - C) \geq \frac{C}{\pi}\arctan(\rho N)$$

giving

$$C \leq C^*, \tag{5.15}$$

from which the stability of E^* is unconditional, whenever it is feasible. Because of the constraints $0 \leq C \leq N$, we must take into account also the boundary points $\hat{E} = (\hat{C}, \hat{I}) = (N, 0)$ and $\widetilde{E} = (\widetilde{C}, \widetilde{I}) = (0, N)$ as possible equilibrium, when E^* is feasible, i.e. for (5.12) or, explicitly, $\rho > 0$. Recalling that C grows whenever (5.15) is satisfied, the equilibrium $\widetilde{E}$ is unconditionally unstable. On the other hand, the converse result holds for $\hat{E}$, for which the derivative is negative, see (5.10), implying instability as well. Indeed,

$$\frac{dC}{dt}|_{C=0} = \frac{1}{2}mN > 0, \quad \frac{dC}{dt}|_{C=N} = -\frac{1}{\pi}mN\arctan(\rho N) < 0 \tag{5.16}$$

The latter inequality however reverses for $\rho < 0$, i.e. in case (5.13) holds, in which case E^* coalesces with the boundary equilibrium $\widehat{E}$. In view of the fact that C grows whenever (5.15) holds, $\hat{E}$ is ultimately approached by the trajectories and therefore it results to be stable when E^* becomes unfeasible, i.e. explicitly for

$$ahq > bp. \tag{5.17}$$

The game then always ends up at equilibrium E^* for (5.12), with $C^* \geq I^*$ in view of (5.14). In case instead of (5.13), the system settles to equilibrium $\hat{E}$, in which the whole population becomes collectivist, $C = N$. In conclusion, the sub-population C always wins the game. These results are summarized in Table 1.

Table 1. The possible system outcomes. In the first four columns the summary of the analytic findings, the last column contains instead the corresponding results of the numerical simulations.

Equilibrium	Analytic Results			Numerical Results
	feasibility	stability	remarks	metastability
$\widetilde{E} = (\widetilde{C}, \widetilde{I}) = (0, N)$	—	unstable	-	$bp > ahq$
$\hat{E} = (\hat{C}, \hat{I}) = (N, 0)$	—	$bp > ahq$: unstable $ahq > bp$: stable	$C^* = \hat{C} = N$	$ahq > bp$
$E^* = (C^*, N - C^*)$	$bp \geq ahq$	stable	$C^* \geq I^*$	$pb = ahq$

6. Discussion

This study analyses the emergence and stabilization of inter-group bias as a stable strategy among humans. Both numerical simulations and a theoretical analysis simulating a social dilemma game played by two clusters of agents, collectivists and individualists are implemented. We aim at bridging the stochastic interpretation of an Agent-Based Model with the corresponding deterministic one, thereby describing the same phenomenon in two alternatives but complementary ways. Our model attempts at replicating in a simple way the conditions in which the first groups of human beings found themselves fighting each other for survival and where the first forms of selective altruism behaviors appeared [5, 6]. The model, through the use of the Ultimatum Game, simulates the emergence of aggregates in which the tendency of individuals to favour members of their own groups (i.e. the inter-group bias) arises, at the same time hindering the external members. The Ultimatum Game is a barraging game in which two agents have to decide how to split an amount initially given to one of them. In each interaction, an agent, the offerer, has to decide how to partition the amount with a second agent, the recipient. If the recipient accepts, the amount is actually shared among the two agents; if the recipient declines, nobody receives anything. Collectivists are the agents who decide in each encounter to hinder players of the other group with an offer dependent on a parameter, the in-group effect. Instead, the individualists agents are those players who, regardless of the other agent encountered, offer always the same amount.

A game is composed by several tournaments. After each of them, a fraction of the population with lower score is replaced. We examined two cases: one in which the replacements are given by the random cloning of the remaining population, as in biological systems, and the other in which the replacement is given by a random choice.

In the first case the coexistence is impossible, and we have a first-order phase transition for the asymptotic fraction of collectivists, for a specific value of the in-group effect, $h^* = p/q$.

For the random replacement case, we ran the numerical simulations on 8 different agent population sizes, i.e. $N = 10, 20, 30, 50, 100, 300, 500, 1000$. We varied the value of the parameter h from 1 to 4 and changed the agent probabilities of accepting the offers, i.e., a and b, as well, in dependence also of the amount donated, i.e., p and q. The results show that a smooth transition exists. It involves the effect of in-group effect, h, on the density of collectivist players at equilibrium regardless of the size of the initial population. This transition arises for values of h between 2 and 2.5 allowing C agents to survive and become the dominant group (Figure 1). Further, from the numerical results the effects of the amount donated by agents and the initial density of collectivists on the final equilibrium population (Figure 2) is noticeable. The higher the difference between the amount offered by C and I agents, the more difficult for collectivists is to survive. Thus they need to increase the hinder, h, toward individualist players. Moreover, alongside h, the initial number of C also plays a significant role in establishing the equilibrium density of C. For instance even for a large initial number of C a value $h \approx 2$ is needed for them to thrive and to attain the equilibrium. The simulations shed light also on the hindering level h that C agents have to exert on I's in order to survive (Figure3 and Figure 4). Increasing the value of h, collectivists finally manage to survive and overcome individualists for all the different amounts offered and for all values of the acceptance probability.

We formulated a dynamical deterministic model for the random replacement case as well. The theoretical analysis shows one unstable equilibrium, another one unconditionally stable and one that is conditionally stable (Table 1). The equilibrium where all the final population is composed only by individualist players ($N = I$) appears to be unstable. Further, the equilibrium where collectivists dominate completely the game ($N = C$) and individualists vanish is also unstable if the inequality $ahq > bp$ is satisfied. It is instead stable for the opposite condition. Finally, the equilibrium in which both groups thrive, but always with a higher value of collectivists, is attained exactly when the former boundary equilibrium ($N = C$) is unstable, i.e. for the inequality $bp > ahq$ being satisfied.

From Table 1 we can compare the theoretical and numerical results. Indeed, the stable states emerging from the simulations deviate from the ones obtained analytically in all cases except for coexistence, and when in the latter case both groups have the same size $C^* = I^*$, a fact that occurs in the very particular case $pb = ahq$. Thus, the states found by means of the numerical simulation, in light of the theoretical analysis, seem to represent metastable equilibrium states maintained by the thermalization induced by the random replacement of new agents. The thermalization could be secondary, for $C > 90\%$ at equilibrium, and primary, for $I \geq 50\%$.

This study represents only a first step in this direction. Future work to improve it could address the existence of a possible relationship between the amount offered by the agents and their probability to accept the offers.

Contribution

A.G. designed the model and carried out the simulations. E.V. performed the analytic calculations. F.B contributed to the interpretation of the results. E.I. and F.S. wrote the manuscript in consultation with A.G. and E.V. All the authors revised the manuscript.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. D. Balliet, J. Wu and C. K. De Dreu, Ingroup favoritism in cooperation: a meta-analysis, *Psychol. Bullet.*, **140** (2014), 1556.
2. M. Hewstone, M. Rubin and H. Willis, Intergroup bias, *Annu. Rev. Psychol.*, **53** (2002), 575–604.
3. A. Romano, D. Balliet and J. Wu, Unbounded indirect reciprocity: Is reputation-based cooperation bounded by group membership?, *J. Exp. Soc. Psychol.*, **71** (2017), 59–67.
4. M. A. Nowak, Five rules for the evolution of cooperation, *Science*, **314** (2006), 1560–1563.
5. W. D. Hamilton, The genetical evolution of social behaviour. I, *J. Theor. Biol.*, **7** (1964), 1–16.
6. W. D. Hamilton, The genetical evolution of social behaviour. II, *J. Theor. Biol.*, **7** (1964), 17–52.
7. R. Dawkins, *The Selfish Gene*. Oxford University Press, Oxford, 1976.
8. T. Yamagishi and N. Mifune, Does shared group membership promote altruism? fear, greed, and reputation, *Rationality and Society*, **20** (2008), 5–30.
9. H. Tajfel and J. C. Turner, The social identity theory of intergroup behavior, in *Key readings in social psychology. Political psychology: Key readings* , (eds. J. T. Jost & J. Sidanius), New York, NY, USA: Psychology Press, (2004), 276–293.
10. M. A. Hogg, *The social psychology of group cohesiveness: From attraction to social identity*. New York University Press, New York, 1992.
11. H. Tajfel, M. G. Billig, R. P. Bundy, et al., Social categorization and intergroup behaviour, *Europ. J. Soc. Psychol.*, **1** (1971), 149–178.
12. T. Yamagishi and T. Kiyonari, The group as the container of generalized reciprocity, *Soc. Psychol. Quart.*, **63** (2000), 116–132.
13. N. Masuda and H. Ohtsuki, Tag-based indirect reciprocity by incomplete social information, *Proceed. Royal Soc. London B Biol. Sci.*, **274** (2007), 689–695.
14. A. G. Sanfey, J. K. Rilling, J. A. Aronson, et al., The neural basis of economic decision-making in the ultimatum game, *Science*, **300** (2003), 1755–1758.
15. E. Fehr and U. Fischbacher, The nature of human altruism, *Nature*, **425** (2003), 785–791.
16. W. Güth and M. G. Kocher, More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature, *J. Econ. Behav. Organ.*, **108** (2014), 396–409.
17. W. Güth, R. Schmittberger and B. Schwarze, An experimental analysis of ultimatum bargaining, *J. Econ. Behav. Organ.*, **3** (1982), 367–388.
18. M. A. Nowak, K. M. Page and K. Sigmund, Fairness versus reason in the ultimatum game, *Science*, **289** (2000), 1773–1775.
19. V. Denes-Raj and S. Epstein, Conflict between intuitive and rational processing: when people behave against their better judgment., *J. Person. Soc. Psychol.*, **66** (1994), 819–829.

20. D. Kahneman, A perspective on judgment and choice: mapping bounded rationality., *Am. Psychol.*, **58** (2003),697–720.

21. H. Oosterbeek, R. Sloof and G. Van De Kuilen, Cultural differences in ultimatum game experiments: Evidence from a meta-analysis, *Exp. Econ.*, **7** (2004), 171–188.

22. R. H. Thaler, Anomalies: The ultimatum game, *J. Econ. Perspect.*, **2** (1988), 195–206.

23. J. H. Kagel, C. Kim and D. Moser, Fairness in ultimatum games with asymmetric information and asymmetric payoffs, *Games Econ. Behav.*, **13** (1996), 100–110.

24. D. G. Rand, C. E. Tarnita, H. Ohtsuki, et al., Evolution of fairness in the one-shot anonymous ultimatum game, *Proceed. Nat. Aca. Sci.*, **110** (2013), 2581–2586.

25. C. Feng, Y. J. Luo and F. Krueger, Neural signatures of fairness-related normative decision making in the ultimatum game: A coordinate-based meta-analysis, *Human Brain Map.*, **36** (2015), 591–602.

26. S. Debove, N. Baumard and J.-B. André, Models of the evolution of fairness in the ultimatum game: a review and classification, *Evol. Human Behav.*, **37** (2016), 245–254.

27. R. Slonim and A. E. Roth, Learning in high stakes ultimatum games: An experiment in the Slovak republic, *Econometrica*, **66** (1998), 569–596.

28. L. A. Cameron, Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, *Econ. Inq.*, **37** (1999), 47–59.

29. J. Henrich, R. Boyd, S. Bowles, et al., In search of Homo Economicus: behavioral experiments in 15 small-scale societies, *Am. Econ. Rev.*, **91** (2001), 73–78.

30. G. Bornstein and I. Yaniv, Individual and group behavior in the ultimatum game: are groups more rational players?, *Exp. Econ.*, **1** (1998), 101–108.

31. T. Kugler, E. E. Kausel and M. G. Kocher, Are groups more rational than individuals? a review of interactive decision making in groups, *Wiley Int. Rev. Cogn. Sci.*, **3** (2012), 471–482.

32. J. H. Kim, The role of identity in intra-and inter-group bargaining in the ultimatum game, *Undergra. Econ. Rev.*, **4** (2008), Article 6.

33. A. Ockenfels and P. Werner, Beliefs and ingroup favoritism, *J. Econ. Behav. Organ.*, **108** (2014), 453–462.

34. S.-H. Chuah, R. Hoffmann, M. Jones, et al., Do cultures clash? Evidence from cross-national ultimatum game experiments, *J. Econ. Behav. Organ.*, **64** (2007), 35–48.

35. R. Boarini, J.-F. Laslier and S. Robin, Interpersonal comparisons of utility in bargaining: evidence from a transcontinental ultimatum game, *Theory Decis.*, **67** (2009), 341–373.

36. K. N. McLeish and R. J. Oxoby, Social interactions and the salience of social identity, *J. Econ. Psychol.*, **32** (2011), 172–178.

37. S. A. Mendoza, S. P. Lane and D. M. Amodio, For members only: ingroup punishment of fairness norm violations in the ultimatum game, *Soc. Psychol. Personal. Sci.*, **5** (2014), 662–670.

38. H. Garrett, The competitive exclusion principle, *Science*, **131** (1960), 1292–1297.