

Emergence of Fairness in Repeated Group Interactions

S. Van Segbroeck,¹ J. M. Pacheco,^{2,3} T. Lenaerts,^{1,4} and F. C. Santos^{5,3}

¹MLG, Université Libre de Bruxelles, Brussels, Belgium

²Departamento de Matemática e Aplicações, Universidade do Minho, Braga, Portugal

³ATP-group, CMAF, Instituto para a Investigação Interdisciplinar, Lisboa, Portugal

⁴AI-lab, Vrije Universiteit Brussel, Brussels, Belgium

⁵DEI, & INESC-ID, Instituto Superior Técnico, TU Lisbon, Lisboa, Portugal

(Received 26 August 2011; published 10 April 2012)

Often groups need to meet repeatedly before a decision is reached. Hence, most individual decisions will be contingent on decisions taken previously by others. In particular, the decision to cooperate or not will depend on one's own assessment of what constitutes a fair group outcome. Making use of a repeated N -person prisoner's dilemma, we show that reciprocation towards groups opens a window of opportunity for cooperation to thrive, leading populations to engage in dynamics involving both coordination and coexistence, and characterized by cycles of cooperation and defection. Furthermore, we show that this process leads to the emergence of fairness, whose level will depend on the dilemma at stake.

DOI: 10.1103/PhysRevLett.108.158104

PACS numbers: 87.23.Kg, 89.75.Fb

Many problems of cooperation among humans boil down to the dilemma of helping others at a cost to ourselves or refraining from doing so while still profiting from the help provided by others [1–3]. Surprisingly often we take the first option, even though rational considerations encourage us not to [1,2]. This talent for cooperation forms one of the cornerstones of human society and is, as such, also largely responsible for the unprecedented success of our species [4]. But how did evolution succeed in shaping such cooperative beings, if the temptation to free ride on the benefits produced by others is always lurking? This paradox of cooperation [5] has been under intense scrutiny for decades and, fortunately, several mechanisms discourage us from actually giving in to this temptation [5–15]. Physicists have investigated some of these mechanisms (for an excellent review, see [8]), as human cooperation constitutes an excellent example of a complex system. Cooperation may, for instance, be worthwhile if your opponent has the chance to return you the favor later on. If he or she is not willing to do so, his or her cheating behavior can still be retaliated. This is Robert Trivers' *direct reciprocity* at work [16]. Theoretical and empirical studies show that individuals who pursue long-term relationships built on mutual cooperation are expected to prevail [17–21]. In this context, tit-for-tat players constitute the most famous example [17]: They always start by cooperating, subsequently repeating their opponent's last move.

Direct reciprocation may enhance cooperation for pairwise interactions, but when larger groups of actors are involved, decision-making becomes much more complex. Similar to the relation between 2-body and many-body interactions in Physics, also in human decisions there is a significant increase in complexity when going from pairwise cooperative game interactions to collective efforts in sizable groups. Technically, such an increase in complexity

is reflected in the number of possible behavioral equilibria, which scales linearly with the group size [22], even in the absence of reactive players. Moreover, it is far from clear under which conditions a cooperator (defector) should switch to defection (cooperation) when engaged in a repeated collective endeavor, wherein some may cooperate while others defect. To whom should one reciprocate [23]? One possibility is to reciprocate towards the entire group. As in previous studies of evolution and assessment of fair offers [24–27], reciprocating towards groups will depend on what is reckoned as a *fair* collective effort, as individuals may develop an aspiration level above which they cooperate, defecting otherwise. Such individuals constitute a N -person generalization of the 2-person reciprocators. Unsurprisingly, the spectrum of possible reciprocator strategies for group, N -person game interactions, is much larger than in the 2-person case. Some reciprocators may, for instance, be willing to cooperate only if the entire group did so in a previous encounter, whereas others may cooperate also in the presence of group members who defected.

Let us consider group decisions involving N individuals described in terms of the repeated N -person prisoner's dilemma (NPD) [28,29], in which all players have the opportunity to contribute a certain amount c ("cost") to the public good. The accumulated amount is multiplied by an investment factor F and subsequently shared equally among all group members, irrespective of their contribution. This entire process repeats itself with a probability w , resulting in an average number of $\langle r \rangle = (1 - w)^{-1}$ rounds per group [5,30]. The outcome of the game may differ from round to round, as individuals can base their decision to contribute on the result of the previous round. We distinguish N different aspiration levels, encoded in terms of the strategies R_M ($M \in \{1, \dots, N\}$). R_M players always contribute in the first round. Subsequently, they contribute

only if at least M players did contribute in the previous round. The threshold M can be regarded as their own perception of a *fair* number of contributions to the public good. In addition to these N different types of reciprocators, we include the strategy AD (always defect) to account for unconditional defectors.

Let us start by assuming an infinitely large population of individuals, where a fraction x of the population plays R_M —allowing one single value of M in the set of all reciprocators—while the remaining fraction plays AD. This will allow us to define the notation before addressing finite populations and an evolving M , while unveiling a dynamical scenario which differs strongly from the one obtained from (repeated) 2-person games. Behavioral dynamics often relies on individuals' propensity to be influenced by the actions and achievements of others. Such social learning or evolutionary dynamics can be described by the replicator equation [31] $\dot{x} = x(1-x)(f_{R_M} - f_{AD})$, where f_{R_M} (f_{AD}) stand for the fitness—or success—of R_M (AD) players, given by their payoff derived from the game group interactions (see Eq. 1 in [32]).

A little algebra allows us to show that the (deterministic) replicator dynamics leads to scenarios in which cooperation may prevail, in connection with at most two internal fixed points, associated with unstable (coordination, x_L^*) and stable (coexistence, x_R^*) equilibria, which depend on the values of w , M , N , and F (for detailed derivations, see Section 1 in [32]). Intuitively, the simultaneous occurrence of these two equilibria, which happens when we face a repeated NPD ($F < N$), can be explained as follows. If the R_M frequency is smaller than x_L^* , there are only a few groups in which R_M players remain cooperative for the entire duration of the game. The benefits they receive from such interactions are insufficient to cover the cost for always being prepared to cooperate in the first round, making them disadvantageous with respect to ADs. Hence, R_M players will only endure as long as their prevalence remains above a minimum fraction x_L^* , representing an unstable fixed point (coordination). But even if they succeed, they will never take over the entire population, unless $M = N$. As long as $M < N$, R_M players will cooperate in partially cooperative groups, opening an escape hatch to the survival of a small fraction of ADs (a fraction $1 - x_R^*$), reflecting the stable coexistence between the two strategies.

Additional insight in the characterization of x_L^* and x_R^* is provided in Fig. 1 which shows that, for given M/N , there is a critical probability $\bar{w}$ above which the two equilibria emerge. $\bar{w}$ increases as we reduce M/N , meaning that more rounds are required to prevent AD from dominating the population. Naturally, the location of x_L^* and x_R^* follows the same trend, creating an interesting, but delicate, balance between the size of the basin of attraction of the coexistence state x_R^* and its actual value. Relaxing the criterium of fairness for reciprocators (lowering M) makes

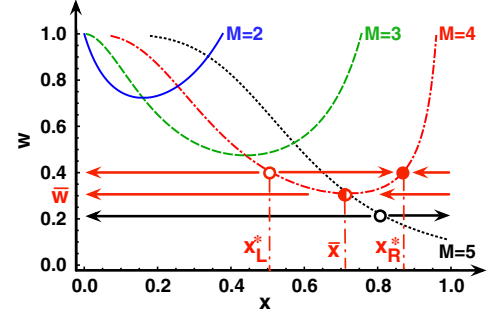


FIG. 1 (color online). Interior fixed points of the replicator equation as a function of w and F . (a) Each curve shows the position of the internal roots of the replicator equation as a function of w for a particular value of M . There are no roots if w is smaller than the critical value $\bar{w}$. A further increase in w leads either to two fixed points (if $M < N$), the left one being unstable and the other one stable, or to just one unstable fixed point (if $M = N$). The arrows indicate the direction of selection ($F = 3.5$, $N = 5$). Results for finite populations with evolving M are shown in Figs. 2 and 3.

the cooperative basin of attraction easier to reach (by reducing x_L^*), but less cooperative overall (reduction of x_R^*).

So far, we have investigated the competition between a single type of reciprocators and unconditional defectors. However, the assessment of what constitutes a fair level of cooperation in a group does not need to be unanimous in the population: The value of M itself may be under selective pressure, and in this case, the delicate competition just described becomes particularly important, mostly if we take into account that populations are finite [11] and selection is not free from errors of imitation [12,13,33] and behavioral mutations [14].

Let us then consider a population of finite size Z , and compute the average prevalence of each of the $N + 1$ available strategies—AD plus the N different R_M strategies—over time. We implement a stochastic, finite population analogue of the deterministic evolutionary dynamics defined before, in which strategies evolve according to a mutation-selection process defined in discrete time. At each time step, the strategy of one randomly selected individual A is updated. With probability μ , A suffers a mutation, adopting a strategy drawn randomly from the space of $N + 1$ available strategies; with probability $1 - \mu$, another randomly selected individual B acts as a role model for A : The probability that A adopts the strategy of B is given by the Fermi distribution $p = [1 + e^{\beta(f_A - f_B)}]^{-1}$ [8,12,33,34], where f_A (f_B) denotes the fitness of individual A (B) and $\beta \geq 0$ measures the strength of the fitness contribution to the update process, i.e., the so-called *intensity of selection* [12].

In the limit in which mutations are rare, we are able to compute analytically the relative prevalence of each of the different strategies [15,21,35] (for details, see Section 2 of [32]). This simplified limit turns out to be valid over a

much wider interval of mutation regimes, as we show below via numerical simulations. In this limit, the population will either end up wiping out the mutant or witness the fixation of the intruder long before the occurrence of a new mutation. Hence, there will be a maximum of two strategies present simultaneously in the population. The fixation probabilities of all possible mutants in all (otherwise) monomorphic populations can be readily computed analytically [12,15,32], defining a reduced (embedded) Markov chain, with which we compute the stationary distribution of the population, i.e., the average fraction of time the population spends in each of the $N + 1$ monomorphic configurations of the population [32,35].

The results are shown in Fig. 2(a), where we plot the stationary distribution for different values of the parameter F . The distribution of R_M players reveals some remarkable features: On one hand, there is a specific concept of fairness, associated with an aspiration level M^* whose corresponding strategy R_{M^*} is most favored by evolution, being the most prevalent among all R_M strategies. On the other hand, unanimity in the assessment of fairness does not occur, given that several values of M may coexist in the population. Finally, as the dilemma becomes harsher

(lower values of F), the higher the fraction of the population that adopts the most prevalent assessment of fairness— M^* —which is always much smaller than the group size N [see also Fig. 3(a)]. Naturally, the success of AD players increases with decreasing F [36].

The intuition behind the emergence of an optimal level of fairness M can be understood with the help of Fig. 2(b), where we analyze the typical flow of probability between the different monomorphic states. Arrows represent transitions favored by natural selection, i.e., those whose fixation probability exceeds $1/Z$ (associated with the fixation probability of a mutant under neutral evolution). Suppose we start from a homogeneous population of ADs. Figure 2(b) shows that there are several R_M types with intermediate M who can invade AD (solid blue arrows). Clearly, such a modest assessment of what constitutes a fair group (intermediate M) combines the best of two worlds: avoiding continuous exploitation, but being sufficiently generous to maintain the level of cooperation in groups that are only partially cooperative. Once a given R_M takes over, neutral drift (grey dashed lines) can drive the population to any of the $N - 1$ other R_M states, which provides a foundation for the coexistence of different concepts of fairness in the population. Whenever the demands for fairness are too modest ($M < M^* = 3$), AD can take

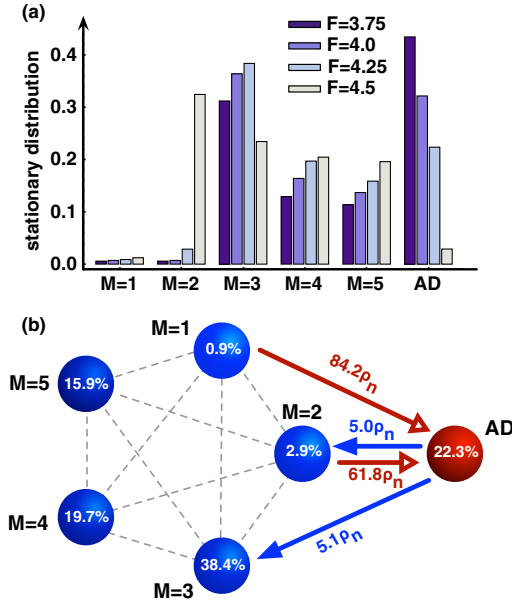


FIG. 2 (color online). Evolution of the threshold M in a finite population. (a) Stationary distribution for different values of F . Irrespective of the success of AD players, which decreases with increasing F , there is always an optimal threshold M^* , whose corresponding strategy R_{M^*} is the most prevalent ($w = 0.9$, $N = 5$, $Z = 100$, $\beta = 1.0$). (b) The percentages indicate the fraction of time the population spends in each composition of the population ($F = 4.25$). Arrows indicate transitions whose fixation probability is greater than $\rho_N = 1/Z$. One observes oscillations between cooperation and defection. The population moves from R_M with small M , over AD, back to R_M with moderate threshold. Neutral drift may bring us back to R_M with small M , as emphasized using grey dotted lines.

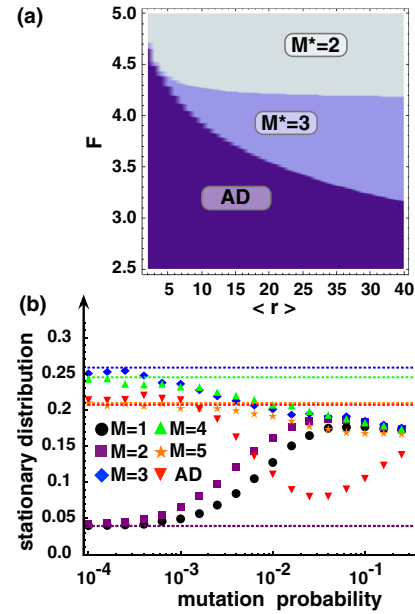


FIG. 3 (color online). Evolutionary dynamics for a) arbitrary number of rounds and b) mutation probabilities. (a) The optimal threshold M^* as a function of F and $\langle r \rangle$ in the limit of rare mutations ($w = 0.9$, $N = 5$, $Z = 100$, $\beta = 1.0$). (b) Dashed lines indicate the stationary distribution in the small-mutation limit. Each symbol indicates, for a given mutation probability, the fraction of the population that adopts the corresponding strategy, averaged over the simulation time (30 simulations, each lasting for 10^9 iterations; $w = 0.9$, $N = 5$, $Z = 100$, $\beta = 0.05$; our results are robust to changes in β [32]).

over the population again (open red arrows). Hence, the population oscillates continuously between cooperation and defection, resembling the cycles of war and peace similar in spirit to those identified in the context of repeated 2-person games of cooperation [21].

This scenario constitutes a general feature of the present model, and is not the result of a particular choice of the average number of rounds $\langle r \rangle$ (or w) or mutation probability μ , as demonstrated in Fig. 3. Figure 3(a) shows that, irrespective of the number of rounds $\langle r \rangle$, AD abounds when F is small, R_M with $M = 2$ when F is large (but still smaller than the group size N), and R_M with $M = 3$ for intermediate values of F , which corresponds exactly to the findings reported in Fig. 2(a). In other words, evolution shapes the population assessment of fairness, depending on the constraints imposed by the collective dilemma. In Fig. 3(b) we investigate the robustness of our results with respect to changes in μ . We abandon the limit of rare mutations, and determine the stationary distributions for arbitrary mutation rates via computer simulations. For $\mu < Z^{-2}$, the results match the limit of rare mutations. More importantly, the plot shows that our general conclusion remains valid for a wide range of mutation probabilities: R_M players with a moderately large aspiration are expected to prevail throughout a wide range of mutation values. For large mutation rates ($\mu > Z^{-1}$), all types of reciprocators become equally probable and dominant with respect to ADs. As a result, the overall outcome of cooperation is enhanced for high mutation rates. This is an important point, as one expects that, e.g., in human interactions, errors of decision making, well captured by the behavioral mutations introduced here, may be sizable [14], although at present a quantitative estimate is lacking. Needless to say, the results shown in Figs. 2 and 3 for $N = 5$, remain valid for other values of N , in the sense that the physical order parameter of the model remains the ratio M/N (see $\bar{x}$ in Section 1 of [32]).

In summary, we have studied the evolutionary dynamics of repeated group interactions, in which individuals engage in an iterated NPD. Reciprocators are defined as individuals who may cooperate, contingent on their own individual assessment of what constitutes a fair group contribution. We found that evolution selects for a moderate, yet prevalent, concept of fairness in the population. This choice results from a detailed competition between the capacity to avoid continuous exploitation and the generosity of contributing in groups which are only partially cooperative. The prevalent concept of fairness that emerges in the population constitutes a compromise between too low aspiration levels, which lead reciprocators to extinction, and too high aspiration levels, associated with harsh coordination thresholds. Combined with the neutrality between different concepts of fairness, the emergent dynamics leads to cyclic behavior which, being ubiquitous in evolutionary games [8,33], also resembles the alternation

between cooperation and defection which seems to pervade throughout human history [37].

Financial support from FNRS Belgium (S.V.S., T.L.) and FCT-Portugal (F.C.S., J.M.P.) is gratefully acknowledged.

-
- [1] G. Hardin, *Science* **162**, 1243 (1968).
 - [2] P. Kollock, *Annu. Rev. Sociol.* **24**, 183 (1998).
 - [3] M. Olson, *The Logic Of Collective Action: Public Goods and the Theory Of Groups* (Harvard University Press, Cambridge, MA, 1971).
 - [4] J. Maynard Smith and E. Szathmáry, *The Major Transitions in Evolution* (Freeman, Oxford, 1995).
 - [5] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, NJ, 2010).
 - [6] F.C. Santos and J.M. Pacheco, *Phys. Rev. Lett.* **95**, 098104 (2005).
 - [7] J. Gómez-Gardenes, M. Campillo, L.M. Floría, and Y. Moreno, *Phys. Rev. Lett.* **98**, 108103 (2007).
 - [8] G. Szabó, G. Fáth, *Phys. Rep.* **446**, 97 (2007).
 - [9] M. Perc and A. Szolnoki, *Phys. Rev. E* **77**, 011904 (2008).
 - [10] G. Szabó and C. Hauert, *Phys. Rev. Lett.* **89**, 118101 (2002).
 - [11] A. Traulsen, J.C. Claussen, and C. Hauert, *Phys. Rev. Lett.* **95**, 238701 (2005).
 - [12] A. Traulsen, M. A. Nowak, and J. M. Pacheco, *Phys. Rev. E* **74**, 011909 (2006).
 - [13] A. Traulsen, J. M. Pacheco, and L. A. Imhof, *Phys. Rev. E* **74**, 021905 (2006).
 - [14] A. Traulsen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 709 (2009).
 - [15] S. Van Segbroeck, F.C. Santos, T. Lenaerts, and J.M. Pacheco, *Phys. Rev. Lett.* **102**, 058105 (2009).
 - [16] R. Trivers, *Q. Rev. Biol.* **46**, 35 (1971).
 - [17] R. Axelrod and W.D. Hamilton, *Science* **211**, 1390 (1981).
 - [18] M. Milinski, *Nature (London)* **325**, 433 (1987).
 - [19] D. Fudenberg and E. S. Maskin, *Am. Econ. Rev.* **80**, 274 (1990).
 - [20] M. A. Nowak and K. Sigmund, *Nature (London)* **355**, 250 (1992).
 - [21] L. A. Imhof, D. Fudenberg, and M. A. Nowak, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10797 (2005).
 - [22] C.S. Gokhale and A. Traulsen, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5500 (2010).
 - [23] K. Sigmund, *Trends Ecol. Evol.* **22**, 593 (2007).
 - [24] M. Nowak, K. Page, and K. Sigmund, *Science* **289**, 1773 (2000).
 - [25] J. Henrich, R. Boyd, S. Bowles, C. Camerer, H. Gintis, R. McElreath, and E. Fehr, *Am. Econ. Rev.* **91**, 73 (2001).
 - [26] K. Sigmund, E. Fehr, and M. Nowak, *Sci. Am.* **286**, 82 (2002).
 - [27] J. Henrich *et al.*, *Science* **327**, 1480 (2010).
 - [28] R. Boyd and P.J. Richerson, *J. Theor. Biol.* **132**, 337 (1988).
 - [29] S. Kurokawa and Y. Ihara, *Proc. Biol. Sci.* **276**, 1379 (2009).
 - [30] In this setting, w and $\langle r \rangle$ can be used interchangeably.

-
- [31] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, England, 1998).
- [32] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.108.158104> for details.
- [33] A. Szolnoki, M. Perc, and G. Szabó, *Phys. Rev. E* **80**, 056109 (2009).
- [34] G. Szabó and C. Toke, *Phys. Rev. E* **58**, 69 (1998).
- [35] D. Fudenberg and L. Imhof, *J. Econ. Theory* **131**, 251 (2006).
- [36] In the presence of execution errors a similar result is obtained, in which the emergent M^* depends on F , w , and the fraction of errors.
- [37] P. Turchin, *War and Peace and War: The Life Cycles of Imperial Nations* (Pi Press, New York, 2006).

Supplemental Material (SM)

Emergence of Fairness in Repeated Group Interactions

Sven Van Segbroeck¹, Jorge M. Pacheco^{2,3}, Tom Lenaerts^{1,4} and Francisco C. Santos^{5,3}

¹MLG, Université Libre de Bruxelles, Boulevard du Triomphe - CP 212, 1050 Brussels, Belgium

²Departamento de Matemática e Aplicações, Universidade do Minho, Campus de Gualtar, 4710 - 057 Braga, Portugal

³ATP-Group, CMAF, Complexo Interdisciplinar, P-1649-003 Lisboa Codex, Portugal

⁴AI-lab, Department of Computer Science, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

⁵Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal

Summary

1. Evolutionary dynamics in infinite populations
 - 1.1 Overview
 - 1.2 Detailed analysis of $Q(x)$
 - a. Analysis for $1 < M < N$
 - Lemma a.1
 - Proposition a.1
 - b. Analysis for $M = N$
 - Proposition b.1
 - c. Analysis for $M = 1$
 - Proposition c.1
2. Evolutionary dynamics in finite populations
 - 2.1. Overview
 - 2.2. Computer simulations
3. References

1. Evolutionary dynamics in infinite populations

1.1. Overview

Given that most mathematical terminology of population dynamics has been proposed for the case of infinite populations [1, 2], we shall ground our analysis of evolutionary dynamics resulting from repeated group interactions on the dynamics obtained from infinitely large populations. In section 2, this discussion will be extended to arbitrary population sizes.

Let us assume an infinitely large, well-mixed population. Groups of N individuals, sampled randomly from the population, interact in repeated **NPD**'s. There are two game strategies: AD and R_M . R_M players always contribute in the first round. Subsequently, they contribute only if at least M players did contribute in the previous round. AD s always opt for defection. We use x to denote the fraction of the population playing R_M . The expected average payoff associated with each strategy is given by [3-5]

$$\begin{aligned} f_{R_M}(x) &= \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \Pi_{R_M}(k+1) \\ f_{AD}(x) &= \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \Pi_{AD}(k) \end{aligned} \quad (1)$$

$\Pi_{R_M}(k)$ ($\Pi_{AD}(k)$) denotes the expected net payoff a R_M (AD) player acquires in a group with k R_M players and $N-k$ AD players. These payoff values are defined as follows

$$\begin{aligned} \Pi_{AD}(k) &= \frac{Fkc}{N} \left[1 + \theta(k-M)(\langle r \rangle - 1) \right] \\ \Pi_{R_M}(k) &= \Pi_{AD}(k) - c \left[1 + \theta(k-M)(\langle r \rangle - 1) \right] \end{aligned} \quad (2)$$

where $\langle r \rangle$ is the average number of rounds and $\theta(x)$ the Heaviside step function ($\theta(x < 0) = 0$ and $\theta(x \geq 0) = 1$).

The payoff of an individual measures the success of his/her strategy in the population. Successful strategies spread; the rest disappears. The replicator equation [1]

$$\dot{x} = x(1-x)(f_{R_M} - f_{AD}) \quad (3)$$

describes this dynamical process. The roots of the fitness difference $Q(x) \equiv f_{R_M}(x) - f_{AD}(x)$ determine the non-trivial *equilibria* of the replicator dynamics.

In the following section we detail the analysis of $Q(x)$ showing that it can be written as

$$Q(x) = c(\lambda - 1) + c(\langle r \rangle - 1)R(x), \quad (4)$$

with $\lambda = F/N$ and

$$R(x) = (\lambda M - 1) \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} + (\lambda - 1) \sum_{k=M}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1}. \quad (5)$$

We also prove several properties of the polynomial $R(x)$ that allow us to come up with a detailed analysis of the roots of $Q(x)$. This analysis consists of a general part, valid for $I < M < N$, and the degenerate cases $M=I$ and $M=N$.

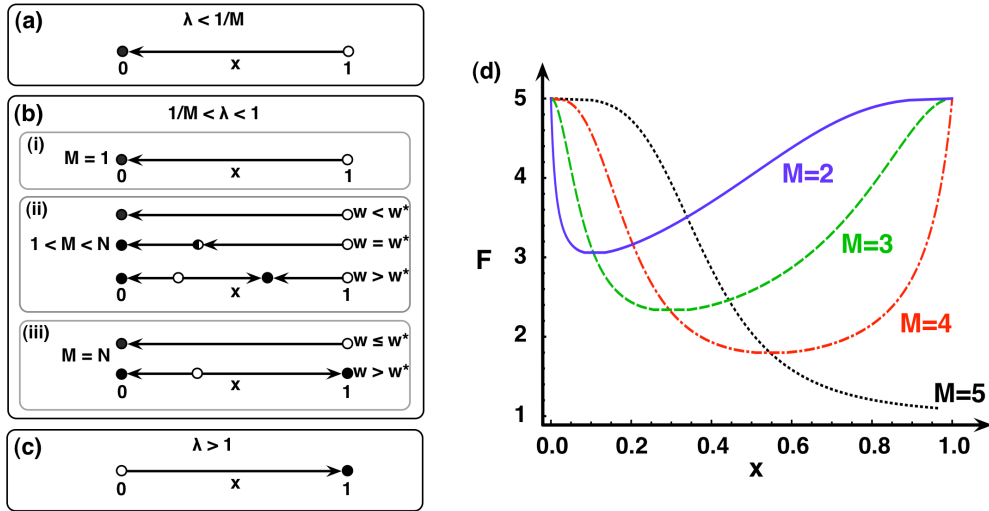


Fig. S1. Classification of all possible dynamical scenarios when evolving an infinitely large population of R_M and AD as a function of M , N , w and $\lambda = F/N$. A fraction x of an infinitely large population adopts the strategy R_M ; the remaining fraction $1-x$ adopts AD . The replicator equation describes the evolution of x over time. Solid (open) circles represent stable (unstable) *equilibria* of the evolutionary dynamics; arrows indicate the direction of selection. **(a)** Defection dominates if $\lambda < 1/M$, irrespective of the other parameters. **(b)** If $1/M < \lambda < 1$, there can be either zero, one or two interior equilibria, depending on the value of w , M and N . **(c)** Cooperation is dominant if $\lambda > 1$. **(d)** Each curve shows the position of the roots of the fitness difference $Q(x)$ as a function of F for a particular value of M . It illustrates the dynamical scenarios pictures on the left panels, and is qualitatively similar to those shown in Fig. 1 of the main text ($w = 0.8, N = 5$).

Fig. S1 illustrates the resulting classification of the different dynamical scenarios. If $1 < M < N$, $R(x)$ attains a maximum at $\bar{x} = \frac{\lambda M - 1}{\lambda N - 1}$. Furthermore, $R(x)$ increases monotonically between 0 and $\bar{x}$, and decreases monotonically between $\bar{x}$ and 1. Then: (i) for $\lambda > 1$, we have $Q(x) > 0$, for all $x \in [0, 1]$; (ii) for $\lambda < 1/M$, we have $Q(x) < 0$, for all $x \in [0, 1]$; (iii) for $1/M < \lambda < 1$ and $\langle r \rangle < 1 + \frac{1 - \lambda}{R(\bar{x})} \equiv \bar{r}$, we have $Q(x) < 0$, for all $x \in [0, 1]$; (iv) for $1/M < \lambda < 1$ and $\langle r \rangle = \bar{r}$, $Q(x)$ has a double root at $x = \bar{x}$; (v) for $1/M < \lambda < 1$ and $\langle r \rangle > \bar{r}$, $Q(x)$ has two simple roots $\{x_L^*, x_R^*\}$, $x_L^* \in]0, \bar{x}[$ and $x_R^* \in]\bar{x}, 1[$. x_L^* is unstable because $Q'(x_L^*) > 0$, x_R^* is stable because $Q'(x_R^*) < 0$.

If $M=N$, our analysis consists of five cases again. The first three are exactly the same as for $1 < M < N$, discussed in the previous paragraph. The final two read as follows: (iv) for $1/M < \lambda < 1$ and $\langle r \rangle = \bar{r}$, $Q(1) = 0$ and $Q(x) < 0$ for all $x \in [0, 1[$; (v) for $1/M < \lambda < 1$ and $\langle r \rangle > \bar{r}$, $Q(x)$ has one simple root in $x^* \in]0, 1[$. $Q(x) < 0$ for $x \in [0, x^*[$ and $Q(x) > 0$ for $x \in]x^*, 1]$.

If $M=1$, R_M is essentially the same as unconditional cooperation, making the analysis independent of $\langle r \rangle$. There are only two cases: (i) for $\lambda > 1$, we have $Q(x) > 0$, for all $x \in [0, 1]$; (ii) for $\lambda < 1$, we have $Q(x) < 0$, for all $x \in [0, 1]$.

1.2. Detailed analysis of $Q(x)$

In section 1.1 we sketch the evolutionary dynamics of an infinitely large, well-mixed population of individuals playing the repeated N -person Prisoner's Dilemma (**NPD**). Here, we study in detail the direction of evolution by analyzing the fitness difference

$Q(x) \equiv f_{R_M}(x) - f_{AD}(x)$. The population is in equilibrium when $Q(x) = 0$. Evolution favors $R_M(AD)$ players if $Q(x) > 0$ ($Q(x) < 0$). Following Equation (1), $Q(x)$ equals

$$Q(x) = \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} (\Pi_{R_M}(k+1) - \Pi_{AD}(k)). \quad (6)$$

The payoff differences in this equation are given by (see Equation (2))

$$\Pi_{R_M}(k+1) - \Pi_{AD}(k) = \Pi_{AD}(k+1) - \Pi_{AD}(k) - c(1 + (r-1)\theta(k+1-M)). \quad (7)$$

Note that we use r as a shorter notation for the average number of rounds $\langle r \rangle$. The payoff difference at the right-hand side of Equation (7) is given by

$$\Pi_{AD}(k+1) - \Pi_{AD}(k) = \lambda c + \lambda c(r-1)[(k+1)\theta(k+1-M) - k\theta(k-M)], \quad (8)$$

so that Equation (7) reduces to

$$\Pi_{R_M}(k+1) - \Pi_{AD}(k) = \begin{cases} cr(\lambda-1) & \text{if } k > M-1 \\ \lambda c(1 + M(r-1)) - cr & \text{if } k = M-1. \\ c(\lambda-1) & \text{if } k < M-1 \end{cases} \quad (9)$$

This allows us to rewrite Equation (6) as follows

$$\begin{aligned} Q(x) = & c(\lambda-1) \sum_{k=0}^{M-2} \binom{N-1}{k} x^k (1-x)^{N-k-1} \\ & + c[\lambda(1 + M(r-1)) - r] \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} \\ & + cr(\lambda-1) \sum_{k=M}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \end{aligned} \quad (10)$$

Since

$$\begin{aligned} 1 &= 1^{N-1} \\ &= (x + 1 - x)^{N-1} \\ &= \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \\ &= \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} + \sum_{k=0}^{M-2} \binom{N-1}{k} x^k (1-x)^{N-k-1} + \sum_{k=M}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \end{aligned} \quad (11)$$

we have that

$$Q(x) = c(\lambda - 1) + c(r - 1) \left[(\lambda M - 1) \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} + (\lambda - 1) \sum_{k=M}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \right]. \quad (12)$$

By introducing the polynomial

$$R(x) \equiv (\lambda M - 1) \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} + (\lambda - 1) \sum_{k=M}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1}, \quad (13)$$

$Q(x)$ can be written as follows

$$Q(x) = c(\lambda - 1) + c(r - 1)R(x). \quad (14)$$

In the following section, we analyze the shape of $Q(x)$, assuming $1 < M < N$, later on followed by the analysis for the degenerate cases $M = N$ and $M = 1$. Together, these results prove the classification of all possible dynamical scenarios shown in Fig. S1.

1.2.1. Analysis for $1 < M < N$

The following lemma facilitates proving the main result of this section, comprised in Proposition a.1.

Lemma a.1 *Let $1 < M < N$. The polynomial $R(x)$ satisfies following properties*

- 1) $R(0) = 0$
- 2) $R(1) = \lambda - 1$
- 3) $R(x)$ attains a maximum at $\bar{x} = \frac{\lambda M - 1}{\lambda N - 1}$. Moreover, $R'(x) > 0$ for all $x \in]0, \bar{x}[$ and $R'(x) < 0$ for all $x \in]\bar{x}, 1[$.

Proof: The first two properties follow immediately from the definition of $R(x)$. We now prove the third property. First we rewrite $R(x)$ as follows

$$\begin{aligned}
 R(x) &= \lambda(M-1) \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} + (\lambda-1) \sum_{k=M-1}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} \\
 &= x^{M-1} \left[\lambda(M-1) \binom{N-1}{M-1} (1-x)^{N-M} + (\lambda-1) \sum_{k=M-1}^{N-1} \binom{N-1}{k} x^{k-M+1} (1-x)^{N-k-1} \right]. \quad (15)
 \end{aligned}$$

By substituting $N-k-1$ by k' , we obtain

$$\begin{aligned}
 R(x) &= x^{M-1} \left[\lambda(M-1) \binom{N-1}{M-1} (1-x)^{N-M} + (\lambda-1) \sum_{k'=0}^{N-M} \binom{N-1}{k'} x^{N-M-k'} (1-x)^{k'} \right] \\
 &= x^{N-1} \left[\lambda(M-1) \binom{N-1}{N-M} \left(\frac{1-x}{x} \right)^{N-M} + (\lambda-1) \sum_{k'=0}^{N-M} \binom{N-1}{k'} \left(\frac{1-x}{x} \right)^{k'} \right], \quad (16) \\
 &\equiv x^{N-1} p(z)
 \end{aligned}$$

where $z = \frac{1-x}{x}$. The polynomial $p(z)$ is of the form $\sum_{i=0}^{N-M} a_i z^i$, with

$$\begin{aligned}
 a_i &= (\lambda-1) \binom{N-1}{i}, \quad 0 \leq i \leq N-M-1 \\
 a_{N-M} &= (\lambda M-1) \binom{N-1}{N-M}
 \end{aligned} \quad (17)$$

Noting that $z' = -\frac{1}{x^2} = -\frac{z+1}{x}$, we obtain the following expression for $R'(x)$

$$\begin{aligned}
 R'(x) &= (N-1)x^{N-2}p(z) - x^{N-2}p'(z)(z+1) \\
 &= x^{N-2} \left[(N-1)p(z) - p'(z)(z+1) \right] \\
 &= x^{N-2} \left[(N-1) \sum_{i=0}^{N-M} a_i z^i - \sum_{i=1}^{N-M} i a_i z^i - \sum_{i=1}^{N-M} i a_i z^{i-1} \right] \\
 &= x^{N-2} \left[(N-1)a_0 - a_1 + (N-1) \sum_{i=1}^{N-M} a_i z^i - \sum_{i=1}^{N-M} i a_i z^i - \sum_{i=2}^{N-M} i a_i z^{i-1} \right] \\
 &= x^{N-2} \left[(N-1)a_0 - a_1 + (N-1) \sum_{i=1}^{N-M} a_i z^i - \sum_{i=1}^{N-M} i a_i z^i - \sum_{i=1}^{N-M-1} (i+1) a_{i+1} z^i \right] \\
 &= x^{N-2} \left[(N-1) \sum_{i=1}^{N-M} a_i z^i - \sum_{i=1}^{N-M} i a_i z^i - \sum_{i=1}^{N-M-1} (i+1) a_{i+1} z^i \right] \\
 &\equiv x^{N-2} S(z)
 \end{aligned} \quad (18)$$

The polynomial $S(z)$ can be further simplified

$$\begin{aligned}
 S(z) &= (M-1)a_{N-M} z^{N-M} + [M a_{N-M-1} - (N-M)a_{N-M}] z^{N-M-1} \\
 &\quad + \sum_{i=1}^{N-M-2} [(N-i-1)a_i - (i+1)a_{i+1}] z^i. \quad (19)
 \end{aligned}$$

For $1 \leq i \leq N - M - 2$, the coefficients a_i satisfy the recurrence relation

$$\begin{aligned} a_{i+1} &= (\lambda - 1) \binom{N-1}{i+1} \\ &= (\lambda - 1) \frac{N-1-i}{i+1} \binom{N-1}{i}, \\ &= \frac{N-1-i}{i+1} a_i \end{aligned} \quad (20)$$

so that

$$\sum_{i=1}^{N-M-2} [(N-i-1)a_i - (i+1)a_{i+1}] z^i = 0 \quad (21)$$

and

$$\begin{aligned} S(z) &= (M-1)a_{N-M} z^{N-M} + [Ma_{N-M-1} - (N-M)a_{N-M}] z^{N-M-1} \\ &= \left[(\lambda M - 1) \binom{N-1}{N-M} z - \lambda M \binom{N-1}{N-M-1} \right] (M-1) z^{N-M-1}. \end{aligned} \quad (22)$$

Therefore, we can write the derivative of $R(x)$ as follows

$$\begin{aligned} R'(x) &= x^{N-2} \left[(\lambda M - 1) \binom{N-1}{N-M} z - \lambda M \binom{N-1}{N-M-1} \right] (M-1) z^{N-M-1} \\ &= M(M-1) \binom{N-1}{N-M-1} x^{M-1} (1-x)^{N-M-1} \left[\frac{\lambda M - 1}{N-M} z - \lambda \right]. \end{aligned} \quad (23)$$

It is clear that $R'(x)$ vanishes at

$$\begin{aligned} \bar{z} &= \frac{\lambda(N-M)}{\lambda M - 1} \\ &= \frac{\lambda N - 1}{\lambda M - 1} - 1. \end{aligned} \quad (24)$$

Since $z = \frac{1-x}{x} = \frac{1}{x} - 1$, $\bar{z}$ corresponds to

$$\bar{x} = \frac{\lambda M - 1}{\lambda N - 1}. \quad (25)$$

Following Equation 23, $R'(x) > 0$ for $z > \bar{z}$ and $R'(x) < 0$ for $0 < z < \bar{z}$. The function

$z = \frac{1-x}{x}$ decreases monotonously and maps $]0,1[$ on $]0,\infty[$. Hence, the interval

$0 < z < \bar{z}$ corresponds to $\bar{x} < x < 1$. The region $z > \bar{z}$ corresponds to $0 < x < \bar{x}$. This proves the third property of Lemma a.1. $\square$

Proposition a.1 *Let $1 < M < N$. $Q(x)$ satisfies the following properties:*

1) *If $1/M < \lambda < 1$, then there is a critical number of rounds $\bar{r} \equiv 1 + \frac{1-\lambda}{R(\bar{x})}$ which*

determines the behavior of $Q(x)$:

i) If $r < \bar{r}$, then $Q(x) < 0$ for all $x \in [0,1]$.

ii) If $r = \bar{r}$, then $Q(x)$ has a double root at $x = \bar{x}$.

iii) If $r > \bar{r}$, then $Q(x)$ has two simple roots $\{x_L^, x_R^*\}$,*

with $x_L^ \in]0, \bar{x}[$ and $x_R^* \in]\bar{x}, 1[$.*

2) *If $\lambda > 1$, then $Q(x) > 0$ for all $x \in [0,1]$.*

3) *If $\lambda < 1/M$, then $Q(x) < 0$ for all $x \in [0,1]$.*

Proof:

1.i) Let $r < \bar{r}$. We obtain the following inequality by applying Lemma a.1

$$\begin{aligned} Q(x) &= c(\lambda - 1) + c(r - 1)R(x) \\ &< c(\lambda - 1) + c(\bar{r} - 1)R(\bar{x}) \\ &= c(\lambda - 1) + c(\bar{r} - 1)\frac{1 - \lambda}{\bar{r} - 1}, \\ &= 0 \end{aligned} \tag{26}$$

for all $x \in]0, 1[$. At the boundaries of the interval $[0, 1]$, we have $Q(0) = c(\lambda - 1) < 0$ and

$Q(1) = cr(\lambda - 1) < 0$. Hence, $Q(x) < 0$ for all $x \in [0, 1]$.

1.ii) For $r = \bar{r}$, we have

$$\begin{aligned} Q(\bar{x}) &= c(\lambda - 1) + c(r - 1)R(\bar{x}) \\ &= c(\lambda - 1) + c\frac{1 - \lambda}{R(\bar{x})}R(\bar{x}). \\ &= 0 \end{aligned} \tag{27}$$

To prove that $\bar{x}$ is a double root, we show that $Q'(\bar{x}) = 0$ and $Q''(\bar{x}) \neq 0$. The existence of a root $Q'(\bar{x})$ in $\bar{x}$ follows directly from Equation 9 and Lemma a.1:

$Q'(\bar{x}) = c(r-1)R'(\bar{x}) = 0$. To find the value of $Q''(\bar{x})$, we first derive an expression for $R''(x)$. We know from Equation 18 that $R'(x) = x^{N-2}S(z)$. The second derivative of $R(x)$ is therefore given by

$$R''(x) = (N-2)x^{N-3}S(z) - x^{N-3}S'(z)(z+1). \quad (28)$$

Calculating the derivative of $S(z)$ gives us

$$S'(z) = M(M-1) \binom{N-1}{N-M-1} z^{N-M-2} [(\lambda M - 1)z - \lambda(N-M-1)]. \quad (29)$$

Since $S'(\bar{z}) \neq 0$, it follows that

$$\begin{aligned} R''(\bar{x}) &= -\bar{x}^{N-3}S'(\bar{z})(\bar{z}+1) \\ &\neq 0 \end{aligned} \quad (30)$$

which proves that $\bar{x}$ is a double root.

1.iii) For $r > \bar{r}$, we have

$$\begin{aligned} Q(\bar{x}) &= c(\lambda - 1) + c(r - 1)R(\bar{x}) \\ &> c(\lambda - 1) + c(\bar{r} - 1)R(\bar{x}). \\ &= 0 \end{aligned} \quad (31)$$

Since $Q(0) = c(\lambda - 1) < 0$ and $Q(1) = cr(\lambda - 1) < 0$, the Intermediate Value Theorem predicts that $Q(x)$ will have at least two roots: one root x_L^* in $]0, \bar{x}[$ and another root x_R^* in $]\bar{x}, 1[$. Since $R'(x)$ has only one root (see Lemma a.1), at $\bar{x}$, $Q(x)$ cannot have more than two roots. Hence, $Q(x)$ increases monotonically in $]0, \bar{x}[$ and decreases monotonically in $]\bar{x}, 1[$.

2) For $\lambda > 1$, $Q(x)$ is positive in 0, 1, and $\bar{x}$:

$$\begin{aligned} Q(0) &= c(\lambda - 1) > 0 \\ Q(1) &= cr(\lambda - 1) > 0 \\ Q(\bar{x}) &= c(\lambda - 1) + c(r - 1)R(\bar{x}) > 0 \end{aligned} \quad (32)$$

Since $Q(x)$ increases monotonically in $]0, \bar{x}[$ and decreases monotonically in $]\bar{x}, 1[$, it follows automatically that $Q(x) > 0$ for all $x \in [0, 1]$.

3) For $\lambda < 1/M$, then $\bar{x} < 0$. Therefore, $Q(x)$ decreases monotonically between 0 and 1. Since $Q(0) = c(\lambda - 1) < 0$ and $Q(1) = cr(\lambda - 1) < 0$, it follows that $Q(x) < 0$ for all $x \in [0, 1]$. $\square$

1.2.2. Analysis for $M=N$

Let us derive an expression for $Q(x)$, which is valid for $M = N$, starting from Equation (6). By calculating the fitness differences

$$\Pi_{AD}(k+1) - \Pi_{AD}(k) = \begin{cases} \lambda c & \text{if } k < N-1 \\ \lambda c + \lambda c(r-1)N & \text{if } k = N-1 \end{cases} \quad (33)$$

and

$$\Pi_{R_M}(k+1) - \Pi_{AD}(k) = \begin{cases} (\lambda - 1)c & \text{if } k < N-1 \\ \lambda c(1 + (r-1)N) - cr & \text{if } k = N-1 \end{cases} \quad (34)$$

we arrive at the following expression for $Q(x)$:

$$\begin{aligned} Q(x) &= x^{N-1} [\lambda c(1 + (r-1)N) - cr] + \sum_{k=0}^{N-2} \binom{N-1}{k} x^k (1-x)^{N-k-1} c(\lambda - 1) \\ &= x^{N-1} [\lambda c(1 + (r-1)N) - cr] + c(\lambda - 1) - c(\lambda - 1)x^{N-1} \\ &= (\lambda - 1)c + c(r-1)(\lambda N - 1)x^{N-1} \end{aligned} \quad (35)$$

Proposition b.1 Let $M = N$. $Q(x)$ satisfies the following properties:

1) If $1/N < \lambda < 1$, then there is a critical number of rounds $\bar{r} \equiv 1 + \frac{1-\lambda}{\lambda N - 1}$ which

determines the behavior of $Q(x)$:

i) If $r < \bar{r}$, then $Q(x) < 0$ for all $x \in [0, 1]$.

ii) If $r = \bar{r}$, then $Q(1) = 0$ and $Q(x) < 0$ for all $x \in [0, 1[$.

iii) If $r > \bar{r}$, then $Q(x)$ has one simple root x^* in $]0,1[$. $Q(x) < 0$ for $x \in [0, x^*[$
and $Q(x) > 0$ for $x \in]x^*, 1]$.

2) If $\lambda > 1$, then $Q(x) > 0$ for al $x \in [0, 1]$.

3) If $\lambda < 1/N$, then $Q(x) < 0$ for al $x \in [0, 1]$.

Proof: The last two properties follow directly from Equation (35). We now prove the first property. Note that $Q(x)$ increases monotonously if $1/N < \lambda < 1$. Therefore, $Q(x)$ has one single root if and only if $Q(0) < 0$ and $Q(1) > 0$. This condition $Q(0) < 0$ is always true for $\lambda < 1$. The other condition, $Q(1) > 0$, holds in case $r > \bar{r} \equiv 1 + \frac{1 - \lambda}{\lambda N - 1}$.

□

1.2.3. Analysis for $M=1$

For $M=1$, Equation (4) reduces to

$$\Pi_{R_M}(k+1) - \Pi_{AD}(k) = (\lambda - 1)cr. \quad (36)$$

Therefore,

$$\begin{aligned} Q(x) &= \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-k-1} (\lambda - 1)cr \\ &= (\lambda - 1)cr(x + 1 - x)^{N-1} \\ &= (\lambda - 1)cr \end{aligned} \quad (37)$$

which proves the proposition below.

Proposition c.1 Let $M=1$. $Q(x)$ satisfies the following properties:

1) If $\lambda > 1$, then $Q(x) > 0$ for al $x \in [0, 1]$.

2) If $\lambda < 1$, then $Q(x) < 0$ for al $x \in [0, 1]$.

2. Evolutionary dynamics in finite populations

2.1. Overview

We consider a well-mixed population of constant size Z . Suppose that only two strategies are present in the population: AD and R_M , M being fixed at one specific value.

The expected payoff associated with each of these two strategies is given by [3, 4]

$$\begin{aligned} f_{R_M}(k) &= \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k-1}{j} \binom{Z-k}{N-j-1} \Pi_{R_M}(j+1) \\ f_{AD}(k) &= \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k}{j} \binom{Z-k-1}{N-j-1} \Pi_{AD}(j) \end{aligned} \quad , \quad (38)$$

where k denotes the number of R_M players in the population.

Strategies evolve according to a mutation-selection process defined in discrete time. At each time step, the strategy of one randomly selected individual A is updated. With probability μ , A undergoes a mutation. He/she adopts a strategy drawn randomly from the space of available strategies, which includes the strategy AD and the N strategies R_M ($M \in \{1, \dots, N\}$). With probability $1 - \mu$, another randomly selected individual B acts as a role model for A . The probability that A adopts the strategy of B equals $p = [1 + e^{\beta(f_A - f_B)}]^{-1}$. A sticks to his/her former strategy with probability $1 - p$.

This update rule is known as the pairwise comparison rule [6, 7]. We used f_A and f_B to denote the fitness of individual A and B , respectively. The parameter $\beta \geq 0$, in EGT called the *intensity of selection*, measures the contribution of fitness to the update process. In the limit of strong selection ($\beta \rightarrow \infty$), the probability p is either zero or one, depending on f_A and f_B . In the limit of weak selection ($\beta \rightarrow 0$), p is always equal to $1/2$, irrespective of the fitness of A and B .

If the mutation probability μ is sufficiently small, the population will never contain

more than two different strategies simultaneously. The time between two mutation events is usually so large that the population will always evolve to a homogeneous state, i.e., to a state in which all individuals adopt the same strategy, before the next mutation occurs. The dynamics can now be approximated by means of an embedded Markov chain whose states correspond to the different homogeneous states of the population [8-11]. Let us denote the list of available strategies as S_i ($i \in \{1, \dots, N+2\}$). The transition matrix $\Lambda = [\Lambda_{ij}]_{i,j=1}^{N+2}$ collects the different (transition) probabilities for the population to move from one state to the other. Specifically, Λ_{ij} is the probability that a population in state S_i will end up in state S_j after the occurrence of one single mutation.

This probability is given by $\Lambda_{ij} = \frac{\rho_{ji}}{N+1}$ ($j \neq i$), where ρ_{ij} is the probability that a S_j mutant takes over a resident population of S_i individuals. The diagonal of the transition matrix is defined by $\Lambda_{ii} = 1 - \frac{1}{N+1} \sum_{\substack{k=1 \\ k \neq i}}^{N+2} \rho_{ki}$. The normalized left eigenvector associated

with eigenvalue 1 of matrix Λ determines the stationary distribution, i.e, the fraction of time the population spends in each of the homogeneous states of the population [12, 13].

The fixation probability ρ_{ij} can be calculated analytically as follows. Let us assume a population with k S_i individuals and $Z-k$ S_j individuals. The probability that the number of S_i individuals increases/decreases by one is given by

$$T^{\pm}(k) = \frac{k}{Z} \frac{Z-k}{Z} \left[1 + e^{m\beta(f_{S_i} - f_{S_j})} \right]^{\pm 1}, \quad (39)$$

establishing also the probability ρ_{ij} [6, 7, 14]

$$\rho_{ij} = \left[1 + \sum_{l=1}^{Z-1} \prod_{k=1}^l \frac{T^-(k)}{T^+(k)} \right]^{-1}. \quad (40).$$

Finally, it is important to highlight the difference between behavioural mutations and execution or implementation errors. In this context, the latter turns the analytical computation of the fitness values of each strategy cumbersome (see Eqs. (2) and (38)), being however easy to compute numerically. In fact, all results portrayed in Fig. 2 of the main text remain qualitatively unaltered if one adopts a small probability (ϵ) associated with a limited fraction of implementation errors.

2.2. Computer simulations

All individual-based computer simulations start from a randomly initialized population of size $Z = 100$. The population dynamics is implemented following the rules described above. We calculate the fitness of an individual by averaging the payoff he/she acquires in 1000 **NPD**'s with randomly selected partners. We have checked that the obtained values provide a good approximation of the actual expected payoff values, which are given by Equation (38) in case of just two strategies. The stationary distributions shown in Fig. 3 of the main text are computed as the configuration of the population averaged over the entire simulation time (10^9 iterations). We do 30 independent runs for each value of μ . The value of $\beta=0.05$ reflects a convenient choice in terms of computation time, but similar results are obtained if we adopt the selection of strength ($\beta=1.0$) used in the rest of the figures (see stationary distributions in Fig. 2 and dashed lines in Fig. 3b).

3. References

- [1] J. Hofbauer, and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998).
- [2] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, 2010).
- [3] C. Hauert *et al.*, Journal of Theoretical Biology **239**, 195 (2006).
- [4] J. M. Pacheco *et al.*, Proc. R. Soc. B **276**, 315 (2008).
- [5] C. S. Gokhale, and A. Traulsen, Proc Natl Acad Sci U S A **107**, 5500 (2010).

- [6] A. Traulsen, J. M. Pacheco, and M. A. Nowak, *J Theor Biol* **246**, 522 (2007).
- [7] A. Traulsen, M. A. Nowak, and J. M. Pacheco, *Phys Rev E* **74** 011090 (2006).
- [8] D. Fudenberg, and L. Imhof, *J Econ Theory* **131**, 251 (2005).
- [9] L. A. Imhof, D. Fudenberg, and M. A. Nowak, *Proc Natl Acad Sci U S A* **102**, 10797 (2005).
- [10] C. Hauert *et al.*, *Science* **316**, 1905 (2007).
- [11] S. Van Segbroeck *et al.*, *Phys Rev Lett* **102**, 58105 (2009).
- [12] S. Karlin, and H. E. Taylor, *A First Course in Stochastic Processes* (Academic Press, 1975).
- [13] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, 2007).
- [14] A. Traulsen, N. Shores, and M. A. Nowak, *Bull Math Biol* **70**, 1410 (2008).