Interim Report   IR-02-041

Volunteering as Red Queen Mechanism for Co-operation in Public Goods Games

Christoph Hauert (christoph.hauert@univie.ac.at)
Silvia De Monte (silvia@fysik.dtu.dk)
Josef Hofbauer (josef.hofbauer@univie.ac.at)
Karl Sigmund (karl.sigmund@univie.ac.at)

Approved by

Abstract

The evolution of co-operation among non-related individuals is one of the fundamental problems in biology and social sciences. Reciprocal altruism fails to provide a solution if interactions are not repeated often enough or groups are too large. Punishment and reward can be very effective but require that defectors can be traced and identified. Here we present a simple but effective mechanism operating under full anonymity. Optional participation can foil exploiters and overcome the social dilemma. In voluntary public goods interactions, co-operators and defectors will coexist. We show that this result holds under very diverse assumptions on population structure and adaptation mechanisms. Thus, voluntary participation offers an escape hatch out of some social traps. Co-operation can subsist in sizeable groups even if interactions are not repeated, defectors remain anonymous, players have no memory and assortment is purely random.

About the Authors

Christoph Hauert
Institute for Mathematics
University of Vienna
Strudlhofgasse 4
A-1090 Vienna, Austria
and
Department of Zoology
University of British Columbia
6270 University Boulevard
Vancouver, B.C., Canada V6T 1Z4

Silvia De Monte
Institute for Mathematics
University of Vienna
Strudlhofgasse 4
A-1090 Vienna, Austria
and
Department of Physics
Danish Technical University
DK-2800 Kgs. Lyngby, Denmark

Josef Hofbauer
Institute for Mathematics
University of Vienna
Strudlhofgasse 4
A-1090 Vienna, Austria

Karl Sigmund
Institute for Mathematics
University of Vienna
Strudlhofgasse 4
A-1090 Vienna, Austria
and
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

Volunteering as Red Queen Mechanism for Co-operation in Public Goods Games

Christoph Hauert
Silvia De Monte
Josef Hofbauer
Karl Sigmund

Public goods are defining elements of all societies. Collective efforts to shelter, protect and nourish the group form the backbone of human evolution from prehistoric time to global civilisation. They confront individuals with the temptation to defect, i.e. to take advantage of the public good without contributing to it. This is known as 'Tragedy of the Commons', 'Free Rider Problem', 'Social Dilemma' or 'Multi-person Prisoner's Dilemma' – the diversity of the names underlines the ubiquity of the issue (*1–7*).

Theoreticians and experimental economists investigate this issue by public goods games (*8–11*), which are characterised by the fact that groups of co-operators do better than groups of defectors, but defectors always outperform the co-operators in their group. In typical examples, the individual contributions are multiplied by a factor r and then divided equally among all players (*12*). With r smaller than the group size, this is an example of a social dilemma (*13, 14*): every individual player is better off defecting than co-operating, no matter what the other players do. Groups would therefore consist of defectors only, and forego the public good. For two-player groups, this is the prisoner's dilemma game. It is well known that in this case co-operation based on direct or indirect reciprocation can get established, provided the probability of another round is sufficiently high (*15, 16*). But retaliation does not work if many players are engaged in the game (*17*), because players intending to punish a defector can do so only by refraining from co-operation in subsequent rounds, thereby also punishing the co-operators in the group.

If players are offered, after each round, the possibility to fine specific co-players, co-operation gets firmly established. This happens even if punishment is costly to the punisher (*18, 19*), and if players believe that they will never meet again (*20*). But such fining, or alternatively rewarding (*21*), requires that players can discriminate individual defectors. While reward and punishment must be major factors in human co-operation, we want to draw attention to a simpler mechanism. It consists in allowing the players not to participate, and to fall back on a safe 'side income' which does not depend on others. Such risk-averse optional participation can foil exploiters and relax the social dilemma, even if players have no way of discriminating against defectors (*22*).

We consider three strategic types: co-operators and defectors, both willing to engage in the public goods game and speculate (though with different intentions) on the success of a joint enterprise, and 'loners' who rely on some autark way of life. Co-operators will not stably dominate the population in such a voluntary public goods game, but neither
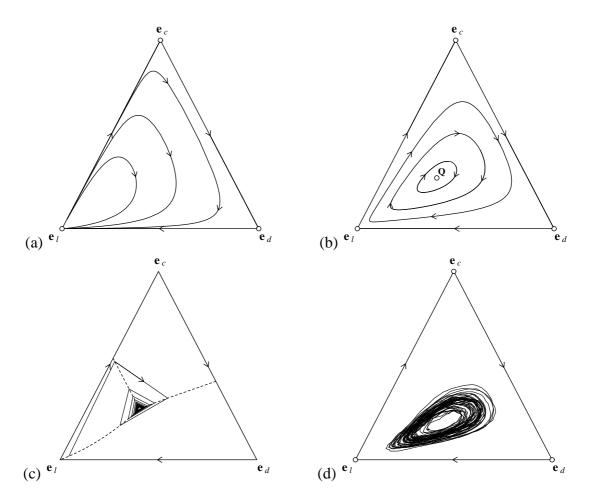
Figure 1: Optional public goods games in large well mixed populations. The three equilibria $\mathbf{e_c}$, $\mathbf{e_d}$ and $\mathbf{e_l}$ are saddle points, denoting homogenous populations of co-operators, defectors and loners. (a) and (b) describe the replicator dynamics $\dot{x}_i = x_i(P_i - \bar{P})$, where $\bar{P}$ is the average payoff in the population. For $r \leq 2$ (a), the interior of the simplex S_3 consists of orbits issued from and returning to $\mathbf{e}_l$. Only brief intermittent bursts of co-operation are observed. (b) For $r > 2$, an equilibrium point $\mathbf{Q}$ appears, surrounded by closed orbits. (c) With perfect information, i.e. best-reply dynamics, $\mathbf{Q}$ becomes a stable fixed point. The dashed lines divide S_3 into three regions where co-operation, defection and loners dominate. (d) Individual based simulations confirm the stability of the cycles in finite populations, if the strategy of a randomly picked individual is imitated whenever it performs better. Parameters: $N = 5$, (a) $r = 1.8, \sigma = 0.5$, (b)-(d) $r = 3, \sigma = 1$, (d) population size: 5000, number of interactions: 10^6.

will exploiters. Their frequencies oscillate, because the public good becomes unattractive if free riders abound.

To model this with evolutionary game theory, we assume a large population consisting of co-operators, defectors and loners. From time to time, a random sample of N individuals is offered to engage in a public goods game. The loners will refuse. They each get a payoff $P_l = \sigma$. The remaining group of S players of the sample consist of n_c co-operators and $S - n_c$ defectors. If $S = 1$, we assume that this single player has to act like a loner. We normalise the individual investment to 1. The defectors' payoff is then $P_d = r \cdot n_c / S$ and the co-operator's payoff is $P_c = P_d - 1$ (due to the cost of co-operation). Hence, in every group, defectors do better than co-operators. We assume $r > 1$ (if all co-operate, they are better off than if all defect) and $0 < \sigma < r - 1$ (better to be a loner than in a group of defectors; but better still to be in a group of co-operators). We stress that players' strategies are decided before the samples are selected, and do not depend on the composition of the group. No anticipation, preferential assortment or conditional response is involved. Co-operation persists in this minimalistic scenario under a wide variety of assumptions concerning population structure or adaptation mechanisms. The results are extremely robust, and do not depend on any particular brand of evolutionary game theory.

In a well-mixed population, analytic expressions for the payoff values can be derived (*23*). The strategies display a rock-scissors-paper cycle. If most players co-operate, it pays to defect. If defectors are prevalent, it is better to stay out of the public goods game and resort to the loners' strategy. But if most players are loners, groups of small size S can form. For such groups, the public goods game is no longer a social dilemma: although defectors do always better than co-operators, in any given group, the payoff for co-operators, when averaged over all groups, will be higher than that of defectors (and loners), and so co-operation will increase. This is an instance of the well-known Simpson's paradox (*24*). Thus, group size S divides the game into two parts. For small size, co-operation is dominant, and for large size, defection; but the mere option to drop out of the game keeps the balance between the two options, in a very natural way.

The game dynamics describing the frequencies of the strategies depends on how players imitate others and learn (*25, 26*). If, for instance, they occasionally update their strategy by picking another player at random, and adopting that model's strategy with a probability proportional to the payoff difference (provided it is positive), then this yields the usual replicator dynamics (*27*). It can be fully analysed in spite of the highly non-linear payoff terms (*28*). For $r < 2$ we observe brief recurrent bursts of co-operation interrupting long periods of prevalence of the loner's strategy. For $r > 2$, a mixed equilibrium appears and all orbits are periodic. The time averages of the frequencies correspond to the equilibrium values, and the time average of the payoff is the same for all strategies, and hence equal to the loner's payoff σ. Other imitation mechanisms may lead to other oscillatory dynamics. In particular, if players always adopt the strategy of their randomly chosen 'model' whenever that model has a higher payoff, then individual-based simulations display stable oscillations for the frequencies of the three strategies (*29*). This finding is very robust, and little affected by additional effects like hyperbolic discounting, random changes of strategies or occasional errors leading to the adoption of strategies with lower payoffs. The oscillations persist if σ, r and N are random variables. Another updating mechanism is the best reply dynamics based on the assumption that from time to time, individuals switch to whatever is the best strategy, given the current composition
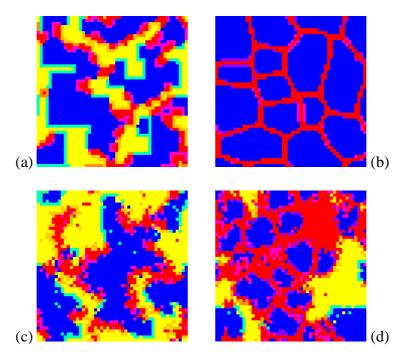
Figure 2: Representative snapshots of the optional public goods games on a square lattice with synchronous updates. In (a), (b) the deterministic rule applies where each site is taken over by the best strategy within its 3×3 neighborhood. In (c), (d) the stochastic rule prescribes that 80% of all sites adopt more successful neighboring strategies, with a probability proportional to the payoff difference. Blue refers to co-operators, red to defectors and yellow to loners. Intermediate colors indicate players that have just changed their strategy. For low multiplication rates ($r = 2.2$ in (a) and (c)) persistent travelling waves are observed regardless of the details of the update rules. In (b), for $r = 3.8$, co-operators thrive on their own and loners go extinct. But in (d), for the same high value of r, co-operators would go extinct in absence of loners, due to the randomness. In a typical configuration, clusters of co-operators are surrounded by defectors and the latter again surrounded by loners. Every now and then co-operators manage to break through the defectors clutch and invade domains of loners. Parameters: 50×50 lattice, periodic boundaries, $\sigma = 1$.
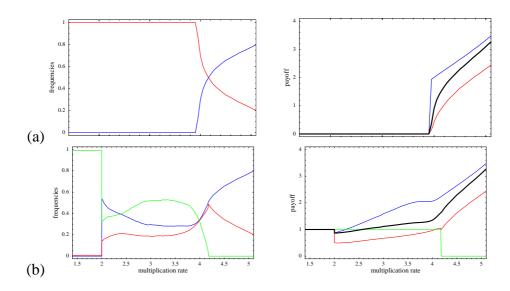
Figure 3: Average frequencies and payoffs in the spatial public goods for (a) compulsory and (b) voluntary participation with a loner's payoff of $\sigma = 1$. Individuals imitate more successful neighboring strategies with a probability proportional to the payoff difference. In (a) co-operators (blue line) persist for sufficiently high interest rates $r \gtrsim 3.90$ through cluster formation i.e. by minimizing interactions with defectors (red line). Interestingly they achieve always significantly higher payoffs than defectors. In (b) the additional protection against exploitation provided by loners (green line) enables co-operators to persist for all $r > \sigma + 1$. For $r \gtrsim 4.17$, the loner strategy no longer represents a valuable alternative and goes extinct – co-operators thrive on their own. As in (a), the payoff of co-operators is significantly higher than for defectors but, somewhat surprisingly, for low r, the average population payoff $\bar{P}$ (blue line) drops even below σ and hence the population would be better off without the opportunity to participate in a public goods game.

of the population. The best-reply dynamics displays damped oscillations converging to a stable polymorphism.

So far we have considered well-mixed populations: groups form randomly, and potential 'role models' are chosen randomly. But the option to withdraw from the game boosts co-operation also for other population structures. For instance, we may assume that individuals are bound to a rigid spatial lattice, and interact only with their nearest neighbors (*30*). As in the related prisoner's dilemma game (*31*), co-operators tend to fare better in the spatial than in the well-mixed case. In the optional public goods game, this is even more pronounced: co-operators persist for all values of $r > \sigma + 1$, whereas in the compulsory game (i.e. without the loner's option), co-operation can persist only for considerably larger values of r (*32*). Thus, loners protect co-operation. The dynamics displays travelling waves driven by the rock-scissors-paper succession of co-operators, defectors and loners (*29, 33*).

In the public goods game, the drop-out option allows groups to form on a voluntary basis, and thus to re-launch co-operation again and again. But each additional player brings a diminishing return and an increased threat of exploitation. As in the land of the Red Queen 'it takes all the running you can do, to keep in the same place'. Individuals keep adjusting their strategies but in the long run do no better than if the public goods option had never existed. On the other hand, voluntary participation avoids the deadlock of mutual defection which threatens any public enterprise in larger groups.

5

References and Notes

1. W. D. Hamilton, *Biosocial Anthropology*, R. Fox, ed. (Malaby Press, London, 1975), pp. 133–153.

2. G. Hardin, *Science* **162**, 1243 (1968).

3. R. L. Trivers, *Q. Rev. Biol.* **46**, 35 (1971).

4. J. Maynard Smith, E. Szathmáry, *The Major Transitions in Evolution* (W. H. Freeman & Co, Oxford, 1995).

5. K. G. Binmore, *Playing Fair: Game Theory and the Social Contract* (MIT Press, Cambridge, 1994).

6. R. Sugden, *The Economics of Rights, Co-operation and Welfare* (Blackwell, Oxford and New York, 1986).

7. B. J. Crespi, *Trends Ecol. Evol.* **16**, 178 (2001).

8. H. Gintis, *Game Theory Evolving* (Princeton University Press, Princeton, 2000).

9. A. M. Colman, *Game Theory and its Applications in the Social and Biological Sciences* (Butterworth-Heinemann, Oxford, 1995).

10. J. H. Kagel, A. E. Roth, eds., *The Handbook of Experimental Economics* (Princeton University Press, Princeton, 1995).

11. E. Fehr, S. Gächter, *Am. Econ. Rev.* **90** 980 (2000)

12. For example, an experimenter endows six players with one dollar each. The players are offered the opportunity to invest their dollar into a common pool. The experimenter then triples the amount in the pool and divides it equally among all six participants, irrespective of their investments. If everybody invests, they triple their fortune, and hence earn two dollars each as an additional income. However, players face the temptation to free ride on the other players' contributions, since each invested dollar yields only a return of 50 cents to the player. Thus, the selfish solution is to invest nothing. But if all players adopt this 'dominant' strategy, it leads to economic stalemate. In actual experiments, many players invest substantially. If the game is repeated, however, the investment drops to zero within a few rounds (*10, 11*).

13. R. M. Dawes, *Ann. Rev. Psychol.* **31**, 169 (1980).

14. T. C. Schelling, *J. Conflict Resolution* **17**, 381 (1973).

15. R. Axelrod, W. D. Hamilton, *Science* **211**, 1390 (1981).

16. C. Wedekind, M. Milinski, *Science* **288**, 850 (2000).

17. R. Boyd, P. J. Richerson, *J. Theor. Biol.* **132**, 337 (1988).

18. R. Boyd, P. J. Richerson, *Ethol. Sociobiol.* **13**, 171 (1992).

19. H. Gintis, *J. Theor. Biol.* **206**, 169 (2000).

20. E. Fehr, S. Gächter, *Nature* **415**, 137 (2002).

21. Milinski, M., Semmann, D. and Krambeck, H.-J., *Nature* **415**, 424 (2002).

22. As noted in (*35*), most analyses of the prisoner's dilemma have tacitly built on the fact that the two partners in the original story are prisoners, whereas in most real-life examples, individuals do have the freedom to chose between playing and not playing. In the few examples where the option of not playing the game was explicitly offered to test persons playing the prisoner's dilemma game, the social welfare increased.

23. If x_c, x_d and x_l specify the frequencies of co-operators, defectors and loners, respectively (with $x_c + x_d + x_l = 1$), then their average payoff values are

$$P_d = \sigma x_l^{N-1} + r \frac{x_c}{1-x_l} \left(1 - \frac{1 - x_l^N}{N(1-x_l)} \right)$$

$$P_c = P_d - (r-1)x_l^{N-1} + \frac{r}{N} \frac{1 - x_l^N}{1 - x_l} - 1$$

$$P_l = \sigma.$$

24. As an example, let us consider two groups A and B, with $N = 10$ and $r = 5$. Assume that A consists of eight co-operators and two defectors. The co-operators obtain \$3 and the defectors \$4. Group B contains two co-operators and eight defectors. Co-operators get nothing and defectors \$1. In both groups, defectors earn one dollar more than co-operators. Yet on average, defectors get only \$1.6 while co-operators earn \$2.4. Whenever the payoff values allow Simpson's paradox to operate in the small groups made possible through the loner's option, rock-scissors-paper dynamics can be expected. The assumption that payoff is linear in the number n_c of co-operators is used only for the sake of simplicity, and because this is the traditional way to model public goods games. The relevance of Simpson's paradox in the evolution of co-operation has been pointed out in (*36*).

25. J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

26. D. Fudenberg, D. Levine, *The Theory of Learning in Games* (MIT Press, Cambridge MA, 1998).

27. K. Schlag, *J. Econ. Theor.* **78**, 130 (1998).

28. An extended version of this will appear in Journal of Theoretical Biology under the title 'Replicator Dynamics for Optional Public Goods Games'.

29. In order to simulate and visualize the evolution in well-mixed populations as well as in populations arranged on rigid regular lattices, we provide interactive virtual laboratories at http://www.univie.ac.at/virtuallabs/PublicGoods. These Java Applets allow to change various parameters, including the geometry of the lattice, and observe the resulting dynamics.

30. In spatial public goods game, we assume that the players invited to participate in the public goods game are for example those in a chess kings' neighborhood of a given site (in which case $N = 9$). In each 'generation' every site of the lattice is the center of one game, such that each player is invited to N games. Subsequently, players update synchronously, either by adopting the strategy of the most successful neighbor, or by adopting the strategy of a more successful neighbor with a probability proportional to the difference in accumulated payoff. We note that an individual's update depends on the strategies in a 7×7 neighborhood.

31. M. A. Nowak, R. M. May, *Nature* **359**, 826 (1992).

32. The threshold value r_c is determined by geometrical configurations. If players imitate the most successful neighbor (including itself), r_c is close to 3; one can check that $r > 3$ corresponds to the condition that a half-plane of co-operators can advance along a straight front into the defectors' regime.

33. The same holds for asynchronous updating of the strategies. A related behavior occurs in the case of the prisoner's dilemma (with two-player interactions) if sites are allowed to go empty (*34*).

34. M. A. Nowak, S. Bonhoeffer, R. M. May, *Int. J. Bifurcation Chaos* **4**, 33 (1994).

35. J. H. Orbell and R. M. Dawes, *Am. Soc. Rev.* **58**, 787 (1993).

36. E. Sober and D. S. Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Harvard University Press, Cambridge, MA, 1999).

37. Ch.H. acknowledges support of the Swiss National Science Foundation; K.S. acknowledges support of the Wissenschaftskolleg WK W008 *Differential Equation Models in Science and Engineering*.