PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2006) **273**, 2565–2570
doi:10.1098/rspb.2006.3600
Published online 5 July 2006

Evolutionary games and population dynamics: maintenance of cooperation in public goods games

Christoph Hauert[1,*], Miranda Holmes[2,†] and Michael Doebeli[2]

[1]*Program for Evolutionary Dynamics, Harvard University, One Brattle Square, Cambridge, MA 02138, USA*
[2]*Departments of Zoology and Mathematics, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada*

The emergence and abundance of cooperation in nature poses a tenacious and challenging puzzle to evolutionary biology. Cooperative behaviour seems to contradict Darwinian evolution because altruistic individuals increase the fitness of other members of the population at a cost to themselves. Thus, in the absence of supporting mechanisms, cooperation should decrease and vanish, as predicted by classical models for cooperation in evolutionary game theory, such as the Prisoner's Dilemma and public goods games. Traditional approaches to studying the problem of cooperation assume constant population sizes and thus neglect the ecology of the interacting individuals. Here, we incorporate ecological dynamics into evolutionary games and reveal a new mechanism for maintaining cooperation. In public goods games, cooperation can gain a foothold if the population density depends on the average population payoff. Decreasing population densities, due to defection leading to small payoffs, results in smaller interaction group sizes in which cooperation can be favoured. This feedback between ecological dynamics and game dynamics can generate stable coexistence of cooperators and defectors in public goods games. However, this mechanism fails for pairwise Prisoner's Dilemma interactions and the population is driven to extinction. Our model represents natural extension of replicator dynamics to populations of varying densities.

Keywords: evolutionary game theory; cooperation; Prisoner's Dilemma; population dynamics

1. INTRODUCTION

Cooperative behaviour is abundant in animal and human societies (Binmore 1994; Colman 1995; Dugatkin 1997; Doebeli & Hauert 2005). Well-known examples include vampire bats regurgitating blood to feed hungry conspecifics (Wilkinson 1984), sticklebacks inspecting predatory pikes preferably in pairs (Milinski 1987), alarm calls from watchful sentinels warning other meerkats from predators (Clutton-Brock *et al.* 1999), musk oxen defending their young against wolves in groups (Wilkinson & Shank 1977), etc. In all these examples, cooperative individuals provide a benefit to one or more individuals at some cost to themselves. However, this behaviour is prone to exploitation by defectors that readily accept support but avoid the costs of assisting others. In the complex society of humans, social interactions lead to multifaceted dilemmas of cooperation. This is most apparent in the consumption of various kinds of public resources, which include public transportation, social welfare, drinking water or clean air. All such resources are prone to exploitation and overuse, as exemplified by the metaphor of the *The tragedy of the commons* (Hardin 1968). Over the past decades, a number of mechanisms have been suggested which are capable of supporting cooperation in absence of genetic relatedness. Most notably, this includes repeated interactions and direct reciprocity (Trivers 1971; Axelrod & Hamilton 1981), punishment (Clutton-Brock & Parker 1995; Fehr & Gächter 2002), spatially structured populations (Nowak & May 1992; Hauert & Doebeli 2004) or voluntary participation in social interactions (Hauert *et al.* 2002*b*). Unique to humans is apparently the capacity for indirect reciprocity (Alexander 1987; Nowak & Sigmund 1998, 2005) and the internalization of benefits as a foundation for moral systems.

Traditionally, the problem of cooperation in social dilemmas (Dawes 1980; Hauert *et al.* 2006) is investigated by means of the game theoretical models of the Prisoner's Dilemma for pairwise interactions and, more generally, public goods games for groups of interacting individuals (Kagel & Roth 1995). In a typical public goods experiment, an experimenter endows, for example, four players with 10 dollars each. All players then have the opportunity to invest their money into a common pool knowing that the experimenter will double the total amount and divide it equally among all participants, irrespective of whether they contributed. Thus, if everybody invests their money, each player ends up with 20 dollars, i.e. doubles the invested money. However, every player faces the temptation to defect, because each invested dollar returns only 50 cents to the investor. Consequently, the rational, selfish solution is to withhold the money and attempt to free ride on the other players'

* Author for correspondence (christoph_hauert@harvard.edu).
† Present address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA.

Received 7 February 2006
Accepted 28 April 2006

© 2006 The Royal Society

contributions—but if everybody follows this reasoning, no one increases the initial capital and foregoes the benefits of the public good.

In formal terms, the payoffs for cooperators and defectors in a group with k cooperators is given by

$$P_{\mathrm{D}}(k) = r\frac{kc}{N}, \tag{1.1a}$$

$$P_{\mathrm{C}}(k) = P_{\mathrm{D}}(k) - c, \tag{1.1b}$$

where r denotes the multiplication factor, c the cooperative investment and N the size of the group engaging in public goods interactions. For simplicity, the costs c are set to unity in the remainder of the text. Note that for pairwise interactions ($N=2$), the public goods formalism can be easily mapped onto the traditional formulation of the Prisoner's Dilemma in terms of costs that a cooperator incurs and benefits that accrue exclusively to their interaction partners (Hauert & Szabó 2003).

In populations of interacting individuals, the dynamics of cooperators and defectors in the public goods game is determined by their respective payoffs obtained in randomly formed groups of N individuals. Thus, any given focal individual finds itself in a group with $N-1$ other individuals. If x is the frequency of cooperators in the population, then the chance that k of those other individuals are cooperators is

$$\binom{N-1}{k} x^k (1-x)^{N-1-k}.$$

This probability is independent of whether the focal individual is a cooperator or a defector. Therefore, every focal individual encounters the same expected number of cooperators, and hence the same expected payoff from the other players during game interactions. From this it follows that the only determinant of success in the well-mixed public goods game is the payoff that the focal individual receives from itself. This payoff is zero if the focal individual is a defector, and is equal to $(r/N-1)c$ if the focal individual is a cooperator. The traditional formulation of public goods games requires $1<r<N$, such that defectors are better off and $x=0$ is globally stable. Conversely, for $r>N$, the social dilemma posed by the public goods game is relaxed and cooperation dominates. However, also note that even in this case, defectors are better off than cooperators in any group consisting of both types. The fact that evolution nevertheless favours cooperation represents an instance of Simpson's paradox (Simpson 1951; Hauert *et al.* 2002*b*). This follows by noting that in this case cooperators receive, on average, higher payoffs than defectors according to the argument given above.

The basic idea of the present paper is that if the public goods game is played in populations of varying densities, then the effective group size S of the public goods interactions also varies. Small population densities result in small effective group sizes and vice versa if population densities are large. Assuming that birth rates are proportional to payoffs, population growth is small or negative if defectors abound, because payoffs in groups with many defectors are low according to equations (1.1*a*) and (1.1*b*). However, if population densities decrease, then the effective interaction group size S also decreases until eventually $r>S$ holds and cooperation is favoured. Thus, we expect scenarios where large population densities favour defection, leading to a decrease in population density and hence to a decrease in S, which, in turn, favours cooperation. Here, we show that this feedback between game dynamics and ecological dynamics can maintain cooperation and lead to stable coexistence of cooperators and defectors.

2. GAME DYNAMICS AND POPULATION DYNAMICS

In order to combine game dynamics and population dynamics in a replicator equation, we assume that x denotes the fraction of cooperators, y the fraction of defectors and $z=1-x-y$ the fraction of empty space. Thus, $x+y$ denotes a normalized population density, such that 0 corresponds to extinction and 1 is the maximal population density. The fractions x and y determine the average payoffs of cooperators, f_{C}, and of defectors, f_{D}, at any given point in time, as detailed below. To determine the dynamics of x, y and z, we assume that cooperators and defectors die at a constant rate d and give birth at a rate proportional to their average payoffs. We also assume that reproduction can only occur into empty space, so that birth rates are proportional to z. This leads to the following population dynamic model:

$$\dot{x} = x(zf_{\mathrm{C}} - d), \tag{2.1a}$$

$$\dot{y} = y(zf_{\mathrm{D}} - d), \tag{2.1b}$$

$$\dot{z} = -\dot{x} - \dot{y} = (x+y)d - z(xf_{\mathrm{C}} + yf_{\mathrm{D}}). \tag{2.1c}$$

This system of equations represents a natural extension of the replicator dynamics (Taylor & Jonker 1978; Hofbauer & Sigmund 1998). If the population density $x+y$ is kept constant ($\dot{z}=0$) by adjusting the death rate accordingly, i.e. by setting $d=z\bar{f}$ (where $\bar{f}=(xf_{\mathrm{C}}+yf_{\mathrm{D}})/(1-z)$ denotes the mean fitness), then the traditional replicator dynamics is recovered.

The average payoffs or fitnesses f_{C} and f_{D} are determined by public goods interactions in randomly formed groups. These interaction groups are formed by interpreting the densities x, y, z as probabilities for drawing a particular strategy and for failing to find a participant, respectively. The general idea is that for large proportions of empty space z individuals typically find themselves in small groups of interacting individuals, and vice versa if z is small. Specifically, we envisage a scenario in which interaction groups are formed randomly in such a way that the available N places are randomly filled with either cooperators or defectors, or are left empty, according to the probabilities x, y and z, respectively; in particular, if $z>0$ individuals typically find themselves in interaction groups of size $S<N$. Thus, the chance that an individual finds itself in a group of size $S \leq N$ is given by the chance to find $S-1$ interaction partners:

$$\binom{N-1}{S-1}(1-z)^{S-1} z^{N-S}. \tag{2.2}$$

If an individual finds itself in a group of size S, it faces m cooperators and $S-1-m$ defectors among its $S-1$ interaction partners with probability

$$\left(\frac{x}{1-z}\right)^m \left(\frac{y}{1-z}\right)^{S-1-m} \binom{S-1}{m}.$$

It follows that the average payoffs to defectors and cooperators, $P_D(S)$ and $P_C(S)$, in a group of size S are

$$P_D(S) = \frac{r}{S}\sum_{m=0}^{S-1}\left(\frac{x}{1-z}\right)^m\left(\frac{y}{1-z}\right)^{S-1-m}\binom{S-1}{m}m$$
$$= r\frac{x}{1-z}\left(1-\frac{1}{S}\right), \tag{2.3a}$$

$$P_C(S) = P_D(S) + \frac{r}{S} - 1. \tag{2.3b}$$

Taking the weighted average over all possible group sizes S, with the weights given by the probabilities described by equation (2.2), yields the average fitness of defectors and cooperators as

$$f_D = \sum_{S=2}^{N}\binom{N-1}{S-1}(1-z)^{S-1}z^{N-S}P_D(S),$$

$$f_C = \sum_{S=2}^{N}\binom{N-1}{S-1}(1-z)^{S-1}z^{N-S}P_C(S).$$

Evaluating the sums, this yields

$$f_D = r\frac{x}{1-z}\left(1-\frac{1-z^N}{N(1-z)}\right), \tag{2.4a}$$

$$f_C = f_D - F(z), \tag{2.4b}$$

where

$$F(z) = 1 + (r-1)z^{N-1} - \frac{r}{N}\frac{1-z^N}{1-z}. \tag{2.5}$$

We note that f_C could in principle become negative, in which case it could not be interpreted as a birth rate. However, positivity of f_C can be achieved by adding a constant baseline birth rate to both f_D and f_C. As long as the death rate is assumed to be larger than this baseline birth rate, populations cannot persist in the absence of public goods interactions, and the results will be qualitatively unchanged. Therefore, we carry out the analysis based on equations (2.4).

We note that we have excluded groups with $S=1$, in which only a single individual joins the public goods game, because this no longer refers to social interactions and cooperation becomes trivial for $r>1$. Thus, the exclusion of such groups is a conservative assumption that nevertheless does not affect the qualitative conclusion that population dynamics can promote cooperation. However, we note that due to this restrictive assumption, cooperation cannot persist for $N=2$, i.e. in the Prisoner's Dilemma. In this case, there is no smaller effective group size $S<r<N$ that would favour cooperation. The parameter N simply acts as an upper bound for the interaction group size. We also note that since interaction groups are formed at random, there is no spatial structure in this model.

(a) *Homogenous populations*

When analysing equations (2.1*a*)–(2.1*c*), it is useful to first consider the special cases where one of the two strategies is absent. In the absence of cooperators ($x=0$), the average payoff of defectors is $f_D=0$ and hence $\dot{y}<0$, such that defectors decrease in frequency and eventually vanish and the system converges to the steady state $z=1$.

Conversely, in absence of defectors ($y=0$), the dynamics becomes more interesting. In this case, $f_C=(r-1)(1-z^{N-1})$, and depending on the values of the parameters N, r and d, the system has one to three fixed points determined by the roots of $\dot{x}=x(zf_C-d)$ (see figure 1). If the death rate d is larger than the maximum of the function zf_C, i.e. if $d>d_{max}=(r-1)(N-1)N^{-N/(N-1)}$, then the population goes extinct ($x\to 0$). However, if $d<d_{max}$, then cooperation can persist for sufficiently large initial densities x_0 (see figure 1). If x_0 is too small, fitness benefits from public goods interactions are insufficient to prevent extinction, because individuals are not encountering sufficiently many interaction partners. The fact that population persistence hinges on sufficiently high population densities relates to the Allee effect (Stephens 1999), which describes a positive correlation between population density and *per capita* growth rate, such that at low densities, populations become prone to extinction, typically due to increasing difficulties in finding mating partners.

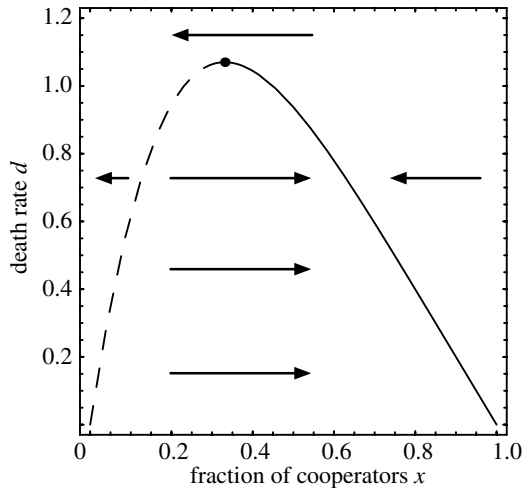

Figure 1. Population dynamics of cooperators engaging in public goods interactions in absence of defectors for different death rates d. Cooperators are unable to survive for $d>d_{max}=(r-1)(N-1)N^{-N/(N-1)}$ and the only stable equilibrium is $x_0=0$. However, for $d<d_{max}$, the system undergoes a bifurcation and two interior equilibria appear: one stable branch at higher x (solid line) and one unstable branch at lower x (dashed line). Consequently, cooperators thrive at sufficiently high densities but go extinct otherwise, i.e. approach x_0. For $d=0$, the equilibrium x_0 becomes unstable and the system converges to $x=1$. The dynamics is illustrated for $N=5$ and $r=3$.

(b) *Heterogenous populations*

Interestingly, population dynamics enables cooperators to survive even when facing exploitation by defectors. In order to analyse this scenario, we introduce the new variable $f=x/(x+y)$, i.e. the relative proportion of cooperators, and rewrite equations (2.1*a*)–(2.1*c*) as

$$\dot{f} = \frac{\dot{x}y-\dot{y}x}{(1-z)^2} = -zf(1-f)F(z), \tag{2.6a}$$

$$\dot{z} = -(1-z)(fz(r-1)(1-z^{N-1})-d). \tag{2.6b}$$

Here, $F(z)$ is given by equation (2.5). Thus, the dynamics unfold in a rectangle determined by $f\in[0,1]$ and $z\in[0,1]$. The dynamics on the boundaries of this rectangle are easily understood. The boundaries $f=0$ and 1 represent the two homogenous scenarios discussed in §2*a*. Thus, z converges to 1 on the boundary $f=0$ (no cooperators), and the dynamics on the boundary $f=1$ has one to three equilibria, corresponding to extinction or possible persistence of cooperators (see figure 2).

The boundary $z=1$ is attracting, as it is easy to see that for sufficiently small population densities $x+y$, $\dot{z}>0$

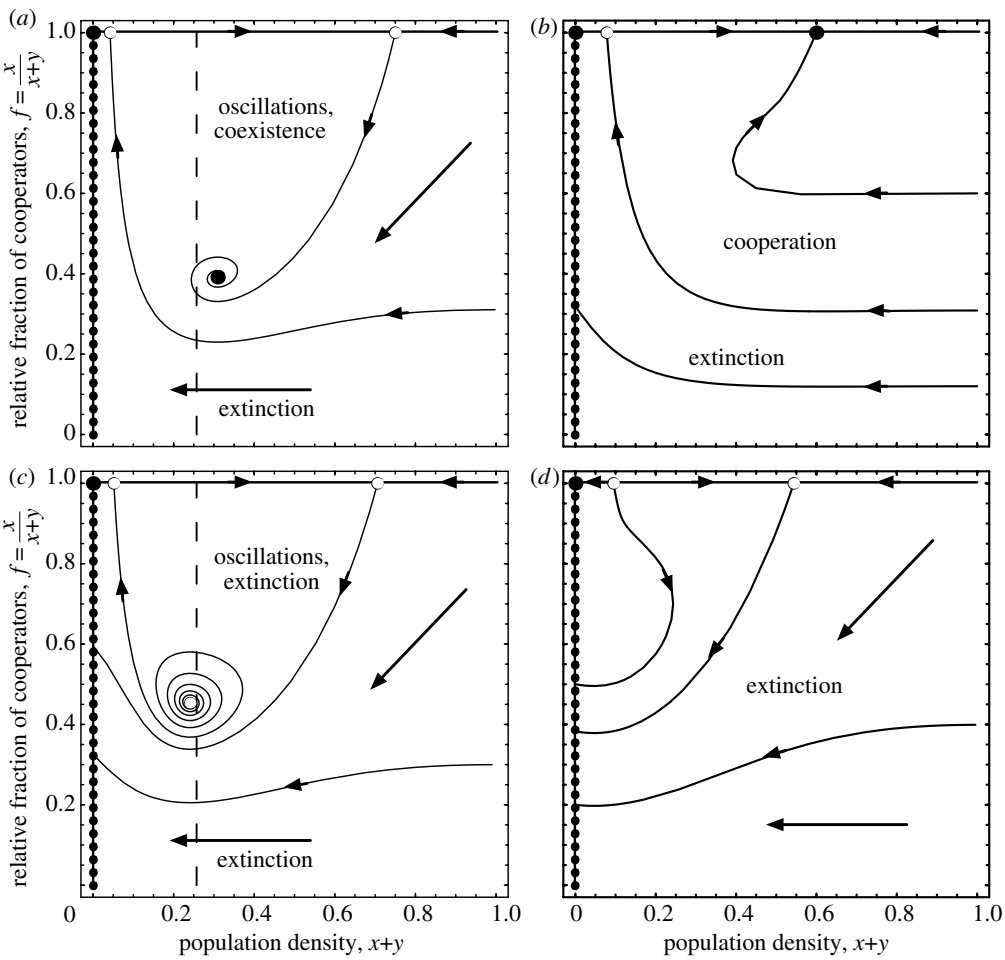

Figure 2. Full analysis of the population dynamics of cooperators and defectors engaging in public goods games in groups of N individuals. Four panels illustrate the different dynamical scenarios. The phase space is spanned by the population density $x+y$ (or $1-z$) and the relative fraction of cooperators $f = x/(x+y)$. The left boundary ($z=1$) is attracting and consists of a line of stable fixed points (filled circles), which represent states where the population cannot maintain itself and disappears. Conversely, the right boundary, which denotes the maximal population density ($z=0$), is repelling. Along the bottom boundary, i.e. in absence of cooperators ($f=0$), population densities decrease and eventually vanish. Finally, along the top boundary, i.e. in absence of defectors ($f=1$), there are two fixed points that are either both saddle points (open circles) as in (*a*), (*c*), (*d*) or a saddle point and a stable equilibrium (filled circle) as in (*b*). In addition, there may be an interior focus $\boldsymbol{Q}$ present. $\boldsymbol{Q}$ is stable if it lies to the right of the dashed line, which marks $\tilde{z} = N^{-1/(N-1)}$. (*a*) $\boldsymbol{Q}$ is stable and cooperators and defectors can coexist. Trajectories spiral towards $\boldsymbol{Q}$, except for low initial population densities or abundant defection, in which case the population goes extinct. (*b*) Increasing death rates push the unstable $\boldsymbol{Q}$ upwards until it disappears. This leads to persistence of cooperators and elimination of defectors. The population vanishes for low initial population densities or abundant defection. (*c*) For smaller multiplication factors r, $\boldsymbol{Q}$ becomes unstable and the population always goes extinct. Trajectories originating in the vicinity of $\boldsymbol{Q}$ approach $z=1$ in an oscillatory manner with increasing amplitude. (*d*) When eliminating the unstable focus $\boldsymbol{Q}$ by increasing d, the population continues to go extinct but the oscillatory dynamics has disappeared. In summary, for increasing d, the dynamics changes from the left to the right column, whereas for decreasing r, it changes from the top to the bottom row. The different scenarios are illustrated for $N=8$ and (*a*) $r=3,\ d=0.5$; (*b*) $r=5,\ d=1.6$; (*c*) $r=2.7,\ d=0.5$; (*d*) $r=2.1,\ d=0.5$.

holds. Thus, states where the population goes extinct are locally attracting, essentially because at low population densities, there are not enough public goods interactions to sustain positive growth rates. In contrast, the opposite boundary with maximal population density ($z=0$) is repelling, $\dot{z}|_{z=0} > 0$ (see equations (2.6*a*) and (2.6*b*)), because offspring production is limited by available space, i.e. proportional to z.

Our analysis of the behaviour in the interior of the phase space reveals four qualitatively different dynamic regimes, which are illustrated in figure 2. These regimes depend largely on the existence of interior equilibrium points. Such points are of the form $\boldsymbol{Q} = (\hat{z}, \hat{f})$, where $\hat{z}$ is solution of $F(\hat{z}) = 0$ and where $\hat{f} = d/(\hat{z}(r-1)(1-\hat{z}^{N-1}))$. It follows from the results in Hauert *et al.* (2002*a*) that $F(\hat{z}) = 0$ has at most one solution in the interval $(0, 1)$, and hence the dynamical system (2.6*a*) and (2.6*b*) admits at most one equilibrium point $\boldsymbol{Q}$ in the interior of the phase space.

In general, if $\boldsymbol{Q}$ exists, its stability depends on parameter values. Specifically, standard stability analysis shows that the equilibrium is locally stable whenever $\hat{z} < \tilde{z} = N^{-1/(N-1)}$ (i.e. whenever the total population density at the equilibrium is higher than the population density that can sustain the highest death rate in the absence of defectors, see figure 1). Moreover, the equilibrium is always a focus (i.e. the two eigenvalues of the corresponding Jacobian matrix are never real).

Figure 2*a* illustrates the first regime, in which $\boldsymbol{Q}$ exists and is a stable focus ($\hat{z} < \tilde{z}$). The basin of attraction of $\boldsymbol{Q}$ is substantial, but it does not span the entire phase space. Whenever defectors abound or population densities are too small, populations remain unable to recover from

exploitation and disappear. However, inside the basin of attraction, the system exhibits oscillatory behaviour with decreasing amplitude as the system converges to $\boldsymbol{Q}$, which represents stable coexistence of cooperators and defectors.

Increasing the death rate d generates the second regime, which is illustrated in figure 2*b*. Larger d increases $\hat{f}$ and eventually $\boldsymbol{Q}$ crosses the boundary $f=1$. Once $\boldsymbol{Q}$ has disappeared, the pure cooperator saddle point, which is stable along the boundary $f=1$, becomes a stable equilibrium. In this case, cooperators and defectors cannot coexist, but cooperators may survive depending on initial conditions. Otherwise, the whole population goes extinct. Note that for further increases of d, the fixed points along $f=1$ collide and disappear ($d > d_{\max}$), as explained in §2*a* (cf. figure 1).

A stable interior equilibrium $\boldsymbol{Q}$ may be destabilized by decreasing the multiplication factor r, which generates the third dynamic regime (figure 2*c*). In this case, $\boldsymbol{Q}$ is an unstable focus ($\hat{z} > \tilde{z}$) and the system invariably converges to the line of fixed points with $z=1$, corresponding to extinction of the population. However, owing to the presence of the unstable focus, the system may exhibit oscillatory behaviour, depending on initial conditions. Owing to the instability of the equilibrium, the amplitude of the oscillations keeps growing and the population eventually disappears.

Increasing the death rate d in this third scenario generates the final regime, which is illustrated in figure 2*d*. Once $\boldsymbol{Q}$ has disappeared, the low-density saddle point, which is unstable along $f=1$, becomes unstable in all directions. This change precludes oscillatory behaviour, but does not affect the evolutionary outcome, and the system keeps converging to the line $z=1$. This scenario also applies to public goods interactions with $r<2$, for which no solution to $F(z)=0$ exists (Hauert *et al.* 2002*a*) and cooperators are always worse off than defectors ($f_{\mathrm{C}} < f_{\mathrm{D}}$). For $r<2$, cooperation cannot be favoured, because the effective group size is $S \geq 2$ and thus never falls below r. Note that this scenario includes pairwise Prisoner's Dilemma interactions because $N=2$ implies $r<2$. Also note that decreasing r in the second regime (figure 2*b*) again destabilizes cooperation and generates this final regime where populations eventually disappear.

3. CONCLUSIONS

The public goods game is a basic and general mathematical metaphor for the problem of cooperation in groups of N interacting individuals. For any multiplication factor $r<N$, defectors dominate and cooperators vanish. In contrast, we have demonstrated that by combining game dynamics and ecological dynamics, cooperators may thrive if the average population payoff determines the population growth rate. Cooperation can be maintained because decreasing population densities can reduce the effective interaction group size S to a point where $r > S$. In these groups, the social dilemma is relaxed and cooperation is favoured. Conversely, high population densities generate large effective interaction group sizes S, in which $r<S$, so that defection is favoured. Thus, if populations are large and defection is favoured, average fitness is low and populations decrease, while if populations are small a high frequency of cooperation leads to high average fitness and triggers an increase in population density. The main result of this paper is that this eco-evolutionary feedback can stabilize cooperation at intermediate frequencies (figure 2). At equilibrium, the population density is such that the expected size of the interaction groups is exactly equal to the multiplication factor r, so that cooperators and defectors have equal payoffs.

Similar effects promoting cooperation based on variable sizes of the group engaging in public goods interactions were observed in Hauert *et al.* (2002*b*). These authors introduced a third strategic type, the loners. Loners are risk averse and do not participate in the public goods game, and instead rely on small but fixed earnings. Thus, the presence of many loners implies small interaction groups for the public goods game. This generates a rock–scissors–paper type cyclic dominance of cooperators, defectors and loners: if the population mainly consists of defectors, then the loner's payoff is higher than the average fitness, and hence the frequency of the loners increases. As the loner frequency increases, the size of the interaction groups decreases, so that eventually cooperation becomes advantageous and takes over. This, in turn, creates the conditions for defectors to thrive and the cycle continues. The cyclic dynamics has actually been verified in experiments where students opted for the three behavioural options in a periodically alternating manner (Semmann *et al.* 2003). In contrast to the ecological model presented here, the inclusion of loners cannot create a locally stable interior equilibrium at which cooperators and defectors persist. Nevertheless, the two scenarios are similar in that in both cases, a third variable is introduced (empty space and loners, respectively), and the dynamics of this third quantity mediates the transitions between the different regimes in the public goods game. However, also note that both scenarios fail to promote cooperation in pairwise Prisoner's Dilemma interactions—in one case only loners survive and in the other the population goes extinct altogether.

Figure 2 gives a complete classification of the configuration and stability of equilibrium states occurring in the model given by equations (2.1*a*)–(2.1*c*). If the interior equilibrium $\boldsymbol{Q}$ at which cooperators and defectors coexist is present, it is always a focus, i.e. small perturbation away from $\boldsymbol{Q}$ generates oscillatory dynamics. In principle, destabilization of $\boldsymbol{Q}$ would therefore open up the possibility of a Hopf bifurcation and of stable limit cycles, leading to cyclic coexistence between cooperators and defectors. However, extensive numerical simulations indicate that such stable limit cycles do not occur in this model. It would be interesting to investigate suitable extensions and modifications of the model that would generate limit cycles. One extension that seems particularly worthwhile is the continuously varying cooperative strategies, rather than just the all-or-nothing strategies cooperate and defect, and how ecological feedback mechanisms affect the evolutionary dynamics of continuous cooperative investments. Continuously varying cooperative investments can lead to interesting new insights and scenarios (Killingback *et al.* 1999; Wahl & Nowak 1999), including evolutionary diversification of cooperative investments levels (Doebeli *et al.* 2004).

Biologically, the models studied here correspond to scenarios in which the size of interaction groups only varies as a consequence of variation in the overall population density. However, one could also imagine scenarios in

which the size of an interaction is influenced by how many defectors and cooperators it contains. For example, the tendency of individuals to join an interaction group might depend on the expected benefit from joining the group, and hence on the number of cooperators it contains. Moreover, this tendency might itself be an evolving trait (Aviles 2002). Even if joining the group is not a behavioural choice, the marginal benefit obtained from having an additional cooperator added to an interaction group could depend on how many cooperators the group already contains. Thus, benefits from cooperation could be saturating or synergistic (Hauert *et al.* 2006). It would be interesting to extend the model presented here to such scenarios.

Furthermore, it would also be interesting to develop empirical tests of the role of ecological dynamics for the evolution of cooperation. There are a number of experimental model systems that could be used for such studies (Doebeli & Hauert 2005). For example, Greig & Travisano (2004) reported that in the yeast *Saccharomyces cerevisiae*, selfish strategies that do not produce an enzyme needed for digestion of sucrose do well when population densities are large, but fare poorly compared to cooperators producing the enzyme when population densities are small. This indicates that there is the potential for an ecological feedback to maintain cooperation as envisaged in our models.

In conclusion, our model represents a step towards understanding how the interplay between population dynamics and game dynamics can affect the evolution of cooperation. The results reported here demonstrate that ecological feedback mechanisms can facilitate the origin and maintenance of cooperation in public goods interactions.

REFERENCES

Alexander, R. D. 1987 *The biology of moral systems*. New York, NY: Aldine de Gruyter.

Aviles, L. 2002 Solving the freeloaders paradox: genetic associations and frequency-dependent selection in the evolution of cooperation among nonrelatives. *Proc. Natl Acad. Sci. USA* **99**, 14 268–14 273. (doi:10.1073/pnas.212408299)

Axelrod, R. & Hamilton, W. D. 1981 The evolution of cooperation. *Science* **211**, 1390–1396.

Binmore, K. G. 1994 *Playing fair: game theory and the social contract*. Cambridge, MA: MIT Press.

Clutton-Brock, T. H. & Parker, G. A. 1995 Punishment in animal societies. *Nature* **373**, 209–216. (doi:10.1038/373209a0)

Clutton-Brock, T. H., O'Riain, M. J., Brotherton, P. N. M., Gaynor, D., Kansky, R., Griffin, A. S. & Manser, M. 1999 Selfish sentinels in cooperative mammals. *Science* **284**, 1640–1644. (doi:10.1126/science.284.5420.1640)

Colman, A. M. 1995 *Game theory and its applications in the social and biological sciences*. Oxford, UK: Butterworth-Heinemann.

Dawes, R. M. 1980 Social dilemmas. *Annu. Rev. Psychol.* **31**, 169–193. (doi:10.1146/annurev.ps.31.020180.001125)

Doebeli, M. & Hauert, C. 2005 Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecol. Lett.* **8**, 748–766. (doi:10.1111/j.1461-0248.2005.00773.x)

Doebeli, M., Hauert, C. & Killingback, T. 2004 The evolutionary origin of cooperators and defectors. *Science* **306**, 859–862. (doi:10.1126/science.1101456)

Dugatkin, L. A. 1997 *Cooperation among animals: an evolutionary perspective*. Oxford, UK: Oxford University Press.

Fehr, E. & Gächter, S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)

Greig, D. & Travisano, M. 2004 The Prisoner's Dilemma and polymorphism in yeast *SUC* genes. *Biol. Lett.* **271**, S25–S26. (doi:10.1098/rsbl.2003.0083)

Hardin, G. 1968 The tragedy of the commons. *Science* **162**, 1243–1248.

Hauert, C. & Doebeli, M. 2004 Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646. (doi:10.1038/nature02360)

Hauert, C. & Szabó, G. 2003 Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interactions. *Complexity* **8**, 31–38. (doi:10.1002/cplx.10092)

Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. 2002*a* Replicator dynamics in optional public goods games. *J. Theor. Biol.* **218**, 187–194. (doi:10.1006/jtbi.2002.3067)

Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. 2002*b* Volunteering as red queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132. (doi:10.1126/science.1070582)

Hauert, C., Michor, F., Nowak, M. & Doebeli, M. 2006 Synergy and discounting of cooperation in social dilemmas. *J. Theor. Biol.* **239**, 195–202. (doi:10.1016/j.jtbi.2005.08.040)

Hofbauer, J. & Sigmund, K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.

Kagel, J. H. & Roth, A. E. (eds) 1995 *The handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Killingback, T., Doebeli, M. & Knowlton, N. 1999 Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proc. R. Soc. B* **266**, 1723–1728. (doi:10.1098/rspb.1999.0838)

Milinski, M. 1987 Tit for tat in sticklebacks and the evolution of cooperation. *Nature* **325**, 433–435. (doi:10.1038/325433a0)

Nowak, M. A. & May, R. M. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)

Nowak, M. A. & Sigmund, K. 1998 Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577. (doi:10.1038/31225)

Nowak, M. A. & Sigmund, K. 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298. (doi:10.1038/nature04131)

Semmann, D., Krambeck, H.-J. & Milinski, M. 2003 Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature* **425**, 390–393. (doi:10.1038/nature01986)

Simpson, E. H. 1951 The interpretation of interaction in contingency tables. *J. R. Stat. Soc. B* **13**, 238–241.

Stephens, P. A. 1999 What is the Allee effect? *Oikos* **87**, 185–190.

Taylor, P. & Jonker, L. 1978 Game dynamics and evolutionarily stable strategies. *Math. Biosci.* **40**, 145–156. (doi:10.1016/0025-5564(78)90077-9)

Trivers, R. L. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)

Wahl, L. M. & Nowak, M. A. 1999 The continuous prisoner's dilemma. I. Linear reactive strategies. *J. Theor. Biol.* **200**, 307–321. (doi:10.1006/jtbi.1999.0996)

Wilkinson, G. S. 1984 Reciprocal food-sharing in the vampire bat. *Nature* **308**, 181–184. (doi:10.1038/308181a0)

Wilkinson, G. S. & Shank, C. C. 1977 Rutting-fight mortality among musk oxen on Banks Island, Northwest Territories, Canada. *Anim. Behav.* **24**, 756–758. (doi:10.1016/S0003-3472(76)80004-8)