Theoretical Population Biology 113 (2017) 1–12

Contents lists available at ScienceDirect

Theoretical Population Biology

journal homepage: www.elsevier.com/locate/tpb

Harsh environments and the evolution of multi-player cooperation

Kris De Jaegher

Utrecht University School of Economics, Utrecht University, Kriekenpitplein 21-22, 3584 EC, Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 2 March 2016
Available online 21 September 2016

Keywords:
By-product mutualism
Evolutionary game theory
Harsh environments
Multi-player public goods games

ABSTRACT

The game-theoretic model in this paper provides micro-foundations for the effect a harsher environment on the probability of cooperation among multiple players. The harshness of the environment is alternatively measured by the degree of complementarity between the players' cooperative efforts in producing a public good, and by the number of attacks on an existing public good that the players can collectively defend, where it is shown that these two measures of the degree of adversity facing the players operate in a similar fashion. We show that the effect of the degree of adversity on the probability of cooperation is monotonous, and has an opposite sign for smaller and for larger cooperation costs. For intermediate cooperation costs, we show that the effect of a harsher environment on the probability of cooperation is hill-shaped.

1. Introduction

Among several evolutionary explanations of cooperation (Dugatkin, 2002; Sachs et al., 2004; Lehmann and Keller, 2006; Nowak, 2006), a particular simple explanation is found in *by-product mutualism* (West Eberhard, 1975; Brown, 1983): players cooperate because it is in their individual interests to do so, and any benefit that this produces for other players is a mere by-product. Mesterton-Gibbons and Dugatkin (1992, 1997) argue that by-product mutualism arises particularly in harsh environments, leading to the so-called common-enemy hypothesis of by-product mutualism. The purpose of this paper is to provide micro-foundations for this common-enemy hypothesis in the setting of cooperation between multiple players.

Mesterton-Gibbons and Dugatkin (1992, 1997) formalize the common-enemy hypothesis by means of a two-player game where each player can either cooperate, or defect. Define R as the reward from jointly cooperating, T as the temptation payoff of unilaterally deviating from joint cooperation, S as the sucker payoff obtained when unilaterally cooperating, and P as the punishment payoff of joint defection. Let it be the case that the harshness of the environment can be reflected by a single measure, referred to as the degree of adversity. Mesterton-Gibbons and Dugatkin assume that the degree of adversity positively affects both $(R-T)$ (i.e., the added payoff of cooperating jointly) and $(S-P)$ (i.e., the added payoff of cooperating alone), and that $(R-T) > (S-P)$. The consequence is that, as the degree of adversity is increased, one moves from a game where joint defection is the only evolutionary stable state (henceforth ESS; Maynard Smith and Price, 1973), to a game where both joint defection and joint cooperation are ESS's, to finally a game where joint cooperation is the only ESS, illustrating the common-enemy hypothesis. Yet, micro-foundations of how the degree of adversity affects the payoffs are not provided. At a more general level, in literature linking cooperation to harsh environments, the mechanism by which the degree of adversity affects cooperation is not clear (Sandoval and Wilson, 2012), and theoretical underpinnings are missing (Smaldino et al., 2013).

For the two-player case, De Jaegher and Hoyer (2016a) provide micro-foundations for the effect of the degree of adversity by modeling cooperation in two ways. In their first model, cooperation consists of the production of a public good, and a higher degree of adversity is linked to a higher degree of complementarity between players' contributions to the public good. For instance, in a cooperative hunt (see, e.g., Scheel and Packer, 1991 on cooperative hunting by lions), when facing the harsher environment of a larger prey (Mesterton-Gibbons and Dugatkin, 1992; Dugatkin, 2002), each predator's effort may become more pivotal in ensuring a successful hunt. In particular, a division of labor may be needed where each predator takes on a specific role (see e.g. Stander, 1992; Leimar and Connor, 2003). What Mesterton-Gibbons and Dugatkin (1992, 1997) call the *boomerang effect* now applies, where a player who unilaterally deviates from joint cooperation is the victim of his own cheating: as he takes on a specific, pivotal role in the cooperative group, unilateral deviation means that little of the public good is produced. One would therefore expect the common-enemy hypothesis to apply. Yet, as argued by De Jaegher and Hoyer (2016a), the harsher environment of, e.g., a larger prey, and the attached

E-mail address: k.dejaegher@uu.nl.

http://dx.doi.org/10.1016/j.tpb.2016.09.003
0040-5809/

higher degree of complementarity between players' contributions, also make the added payoff of cooperating alone decrease. A *sucker effect* applies, where unilateral deviation from joint defection is unattractive; if every player takes on a specific role in the cooperative group, then unilaterally cooperating does not produce much value of the public good. Looking at the sucker effect in isolation, rather than the common-enemy hypothesis applying, one would expect a competing hypothesis to apply, where a harsher environment makes cooperation less likely.

The same mechanisms apply in the second model treated in De Jaegher and Hoyer, where cooperation consists of the defense of an existing public good, and where a harsher environment is measured by a larger number of attacks on the players. For instance, male lions defend territories in order to keep exclusive access to females (Grinnell et al., 1995; other possible examples of territorial defense include Gese, 2001, and Rubenstein and Nuñez, 2009; see Port et al., 2011 for a more general treatment). Also, prey collectively defend against predators (Mesterton-Gibbons and Dugatkin, 1992, p. 274; Spieler, 2003). In De Jaegher and Hoyer's (2016a) model of collective defense, as long as players' contributions to collective defense are complementary to a sufficient extent, a harsher environment also causes a boomerang and a sucker effect: a larger number of attacks on the one hand makes it less attractive to deviate from joint cooperation (as a unilaterally defecting player is more likely to be attacked), and on the other hand makes it less attractive to deviate from joint defection (as a unilaterally cooperating player is less likely to make a difference).

For both models, De Jaegher and Hoyer show that the boomerang effect is the dominant effect for large cooperation costs, in which case the common-enemy hypothesis applies. For small cooperation costs, the competing hypothesis applies instead. Yet, a weakness of their model is that it only considers two players, and not the empirically more relevant case of multiple players. It is not clear then, first, whether similar results apply in the case of multiple players (in general, as pointed out by e.g. Peña et al., 2014, Gokhale and Traulsen, 2014, and Broom and Rychtář, 2016 having more than two players makes the selection gradient non-linear and may change the number of fixed points); second, whether the cases of large and small cooperation costs both remain equally relevant in the case of multiple players. The present paper analyzes the multi-player case, and not only identifies similarities to, but also critical differences with, the two-player case.

2. General setting

We start with a general setting for public goods games, which fits both a public goods game where players produce a public good (Section 3), and a public goods game where players defend an existing public good (Section 4). The public goods games we consider are one-shot, have n players ($n \geq 2$), are binary, and have constant costs. Our players face the binary choice of either investing in the public good (= cooperate), or not investing (= defect). We assume an infinitely large, well-mixed population that reproduces asexually. At any given point of time, the population contains a fraction x of cooperators, and a fraction $(1-x)$ of defectors. The population is repeatedly and randomly matched in groups of n players. In line with these assumptions, the change in the fraction of cooperators x follows the continuous replicator dynamics (Hofbauer and Sigmund, 1998), and is determined by the performance of cooperators relative to defectors:

$$\dot{x} = x\,(1-x)\,[f_C\,(x) - f_D\,(x)], \tag{1}$$

where $[f_C\,(x) - f_D\,(x)]$ is the gain function of cooperating rather than defecting, and where $f_C\,(x)$ denotes the average fitness of cooperating and $f_D\,(x)$ denotes the average fitness of defecting, as a function of the fraction of cooperators x. These fitnesses equal respectively

$$f_C\,(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} x^k\,(1-x)^{n-1-k}\,b_{k+1} - c \tag{2}$$

$$f_D\,(x) = \sum_{k=0}^{n-1} \binom{n-1}{k} x^k\,(1-x)^{n-1-k}\,b_k. \tag{3}$$

Eqs. (2) and (3) can be understood as follows. First, when exactly k other players cooperate in a group, this generates a benefit b_k to the focal player in this group who defects, and a benefit b_{k+1} to the focal player in this group who cooperates. These benefits constitute a public good to the given group, as they are non-excludable (Dionisio and Gordo, 2006): each player in a group always obtains the same benefit. As we limit ourselves to constant cost games, cooperating comes at a constant cost c. Second, players in the population are randomly matched in groups of size n, with $n \geq 2$. It follows that from the perspective of a focal player, the number k of cooperators among the $(n-1)$ other players in his current group follows a binomial distribution (cf. Archetti and Scheuring, 2012; Peña et al., 2014). The individual cooperator will have at least one cooperator (namely himself) in his group, so that the number of cooperators ranges from 1 to n. The individual defector will have at most $(n-1)$ cooperators in his group, so that the number of cooperators ranges from 0 to $(n-1)$.

We now use concepts introduced in Peña et al. (2014), which allow for a simple characterization. The benefit sequence is the sequence of all benefits, $\boldsymbol{b} = (b_0, b_1, \ldots, b_n)$. For k such that $0 \leq k \leq (n-1)$, denote the first forward difference of b_k as $\triangle b_k = b_{k+1} - b_k$ (which is the equivalent of the first derivative of a real function). For k such that $0 \leq k \leq (n-2)$, denote the second forward difference of b_k as $\triangle^2 b_k = \triangle b_{k+1} - \triangle b_k$ (which is the equivalent of the second derivative of a real function). $\triangle b_k$ is the added benefit (or incremental benefit) of cooperating rather than defecting in a group where k other players cooperate, where the added benefit sequence is the sequence $\triangle \boldsymbol{b} = (\triangle b_0, \triangle b_1, \ldots, \triangle b_{n-1})$. The shape of $\boldsymbol{b}$ and $\triangle \boldsymbol{b}$ together characterizes the "technology" through which players' investments get turned into value of the public good (a taxonomy of extreme cases of such technologies is found in Hirshleifer, 1983). We limit ourselves to technologies where $\boldsymbol{b}$ is an increasing sequence, meaning $\triangle \boldsymbol{b} > \boldsymbol{0}$; simply, the value of the public good generated increases in the number of investing players.

The technologies we consider are distinguished by the sign of $\triangle^2 \boldsymbol{b}$ (i.e., the sign of $\triangle^2 b_k$, for k such that $0 \leq k \leq (n-2)$; cf. Motro, 1991). First, with a convex technology, it is the case that $\triangle^2 \boldsymbol{b} > \boldsymbol{0}$, meaning that $\boldsymbol{b}$ is convex (and that $\triangle \boldsymbol{b}$ is increasing). Starting from a situation with only defectors in a group, if one consecutively adds cooperators to the group, each additional cooperator adds more and more value to the public good. A limit case of the convex technology is the weakest-link technology, where $b_0 = b_1 = \cdots = b_{n-1} = 0$, and $b_n > 0$ (meaning that $\triangle b_0 = \triangle b_1 = \cdots = \triangle b_{n-2} = 0$, $\triangle b_{n-1} > 0$), so that benefits are only obtained in case all players in a group cooperate; as soon as at least one player defects, zero benefits are obtained. In this limit case, the public goods game is a *weakest-link game* (Hirshleifer, 1983).

Second, with a concave technology, it is the case that $\triangle^2 \boldsymbol{b} < \boldsymbol{0}$, meaning that $\boldsymbol{b}$ is concave (and that $\triangle \boldsymbol{b}$ is decreasing). Starting from a situation with only defectors in a group, if one consecutively adds cooperators to the group, each additional cooperator adds less and less to the value of the public good. A limit case of the concave technology is the best-shot technology, where $b_1 = b_2 = \cdots = b_n > 0$, and $b_0 = 0$ (meaning that $\triangle b_1 = \triangle b_2 = \cdots = \triangle b_{n-1} = 0$, $\triangle b_0 > 0$), so that maximal benefits of the public good are obtained as soon as at least one player in a group cooperates. In this limit case, the public goods game is a so-called *volunteer's dilemma* (Diekmann, 1985).