



By-product mutualism and the ambiguous effects of harsher environments – A game-theoretic model



Kris De Jaegher^{a,*}, Britta Hoyer^b

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

^b Paderborn University, Department of Economics, Warburger Str. 200, 33098 Paderborn, Germany

HIGHLIGHTS

- We provide a game-theoretic model of by-product mutualism.
- Using this model, we study the effect of harsher environments on cooperation.
- In variants of the model, harsh environments are interpreted in different ways.
- Harsher environments may either encourage or discourage cooperation.

ARTICLE INFO

Article history:

Received 10 June 2015

Received in revised form

23 December 2015

Accepted 28 December 2015

Available online 15 January 2016

Keywords:

By-product mutualism

Game theory

Harsh environments

Common enemies

ABSTRACT

We construct two-player two-strategy game-theoretic models of by-product mutualism, where our focus lies on the way in which the probability of cooperation among players is affected by the degree of adversity facing the players. In our first model, cooperation consists of the production of a public good, and adversity is linked to the degree of complementarity of the players' efforts in producing the public good. In our second model, cooperation consists of the defense of a public, and/or a private good with by-product benefits, and adversity is measured by the number of random attacks (e.g., by a predator) facing the players. In both of these models, our analysis confirms the existence of the so-called boomerang effect, which states that in a harsh environment, the individual player has few incentives to unilaterally defect in a situation of joint cooperation. Focusing on such an effect in isolation leads to the "common-enemy" hypothesis that a larger degree of adversity increases the probability of cooperation. Yet, we also find that a sucker effect may simultaneously exist, which says that in a harsh environment, the individual player has few incentives to unilaterally cooperate in a situation of joint defection. Looked at in isolation, the sucker effect leads to the competing hypothesis that a larger degree of adversity *decreases* the probability of cooperation. Our analysis predicts circumstances in which the "common enemy" hypothesis prevails, and circumstances in which the competing hypothesis prevails.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Explaining the evolution of cooperation is one of the central themes in biology (Dugatkin, 1997, 2002a; Sachs et al., 2004; Lehmann and Keller, 2006; Nowak, 2006). A key explanation for the evolution of cooperation is *kin selection*, where cooperation occurs among genetically related organisms (Hamilton, 1964; West et al., 2002). For cooperation among unrelated organisms (Dugatkin, 2002b; Clutton-Brock, 2009), *reciprocal altruism* has been offered as an explanation

(Trivers, 1971; Pfeiffer et al., 2005), where organisms interact repeatedly and are able to enforce cooperation by punishing defectors. Yet, empirical support for reciprocal altruism among organisms is so far limited (Hammerstein, 2003).

Among several alternative explanations for cooperation among unrelated organisms (most prominently *group selection* (Wilson, 1980)), a straightforward explanation is found in *by-product mutualism* (West Eberhard, 1975; Brown, 1983): unrelated organisms cooperate¹

* Corresponding author at: Utrecht University School of Economics, Utrecht University, Kriekenpitplein 21-22, 3584 EC Utrecht, The Netherlands. Tel.: +31 30 2539964; fax: +31 30 2537373.

E-mail address: k.dejaegher@uu.nl (K. De Jaegher).

<http://dx.doi.org/10.1016/j.jtbi.2015.12.034>

0022-5193/© 2016 Elsevier Ltd. All rights reserved.

¹ Following Mesterton-Gibbons and Dugatkin (1992, 1997), in this paper the noun "cooperation" broadly refers to an outcome that is good for a group, but requires costly collective action of the members of this group. The verb "cooperate" refers to individual cooperative behavior of a player, independently of the behavior of other players (where cooperation is only achieved if a sufficient number of other players cooperate as well).

because it is in their individual interests to do so, and the benefits that this conveys to other organisms arise as a by-product. Potential examples include cooperative hunting among lions (*Panthera leo*) (Scheel and Packer, 1991), collective defense by male lions of territories to maintain exclusive access to females (Grinnell et al., 1995), and instances of food sharing (Stevens and Gilby, 2004; e.g., house sparrows call each other towards food sources to avoid predation risk (Elgar, 1986)) and collective breeding (Clutton-Brock, 2002; e.g., birds feed unrelated young to avoid predators being attracted to their own young (Brown, 1987)). Using a simple two-player two-strategy game-theoretic model, Mesterton-Gibbons and Dugatkin (1992, 1997) argue that, in particular, the mechanism that triggers by-product mutualism is a harsh environment, so that any individual that defects from joint cooperation becomes the victim of its own defection, an effect which they term the *boomerang effect* (or boomerang factor). The argument that the boomerang effect is specifically at work in harsh environments, suggests a “common-enemy” hypothesis, predicting that cooperation is more frequent the higher the degree of adversity faced by the organisms. In a two-player setting, this hypothesis is clearly valid if a harsh environment not only gives the individual player fewer incentives to unilaterally defect in a situation of joint cooperation, but also more incentives to unilaterally cooperate in a situation of joint defection, as Mesterton-Gibbons and Dugatkin assume.²

The key argument underlying our paper is that, while the boomerang effect systematically exists (i.e., under adversity, the individual player has few incentives to unilaterally deviate from joint cooperation), what we call a *sucker effect* may also exist, saying that under adversity the individual player has few incentives to unilaterally cooperate in a situation of joint defection. The coexistence of the boomerang and the sucker effect makes the effect of adversity on the probability of cooperation ambiguous, as the boomerang effect suggests that adversity makes it more difficult to escape the basin of attraction of joint cooperation, while the sucker effect suggests that adversity makes it more difficult to escape the basin of attraction of joint defection. When the boomerang effect is the dominant effect, the common-enemy hypothesis is maintained. Yet, when the sucker effect is instead the dominant effect, a competing hypothesis is obtained, stating that a harsher environment leads to a lower probability of cooperation.³ Our purpose is then to, *first*, identify in what circumstances the sucker effect does/does not apply, and *second*, when the sucker effect applies, to identify the circumstances in which the boomerang effect, or on the contrary the sucker effect is dominant.

At a more general level, our paper contributes to the theory of by-product mutualism by providing *micro-foundations* for the effect of adversity on the probability of cooperation. Indeed, in the current literature, the underlying mechanism by which adversity favors cooperation often remains unclear (Sandoval and Wilson, 2012), and lacks theoretical underpinnings (Smaldino et al., 2013). Given the possibility of ambiguous effects of harsher environments on cooperation, providing such underpinnings is all the

more important. As a first step in providing underpinnings, our paper focuses on simple two-player two-strategy games, allowing direct comparison to the original two-player model by Mesterton-Gibbons and Dugatkin (1992, 1997).

When attempting to construct a micro-foundations model of the effect of adversity on cooperation, one hits upon the problem that the literature interprets adversity in a myriad of ways. First, adversity has been likened to the risk of predation in case of collective defense against a predator (Mesterton-Gibbons and Dugatkin, 1992, p. 274; Spieler, 2003), or in case of cooperative breeding (Krams et al., 2010). Alternatively, adversity has been likened to the size of the prey in case of group hunting (Scheel and Packer, 1991; Mesterton-Gibbons and Dugatkin, 1992; Dugatkin, 2002b), and to relative group size in case of intergroup conflict (Bonanni et al., 2010). More generally, adversity has been linked to the degree of interdependence between potentially cooperating animals (Roberts, 2005), to the scarcity of the available resources (Strassman et al., 2000; Callaway et al., 2002), as well as to bad weather conditions (Dugatkin, 1997, p. 84).

In order to reflect the diversity of interpretations of adversity in the literature, we consider variants of our model where cooperating means contributing to the *production* of a public good (Model 1), and where alternatively cooperating means contributing to the *defense* of a public and/or private good (where the private good generates by-product benefits) (Model 2). In Model 1, the public good is, e.g., a large prey, where predators “produce” the catch of this prey by means of their hunting efforts (cf. Scheel and Packer, 1991). In Model 2, the public good may be thought of as a common territory, of which animals maintain the value by means of their defensive efforts (cf. Grinnell et al., 1995). The private good may be thought of as a food source, where the defensive effort consists of calling other animals to the food source, in order to reduce predation risk (cf. Elgar, 1986).

In the contexts of these models, we refer to a changed environment as more adverse (or: harsher) if, for any given number of cooperating players in the group, the change in environment makes players weakly worse off. In particular, in Model 1 we proxy the degree of adversity by the extent to which each player's cooperative effort is critical in producing the public good, referred to as the *degree of complementarity*. E.g., in the example of cooperative hunting, where adversity appears in the form of a larger prey, for the same mix of cooperative efforts among the predators, the probability of a successful hunt may be lower (particularly in any case where not all predators exert equal effort). This is because a larger prey may make each individual predator's contribution to a group hunt become more critical in ensuring success. For instance, it may be that a larger prey can only be successfully caught if each cooperating predator fulfills a specific task (Packer and Ruttan, 1988).⁴ In Model 2, we measure the degree of adversity by the number of random attacks faced by the players defending a public good, and/or private good with by-product benefits. For the same mix of defensive efforts among the players, a higher number of attacks decreases the probability of maintaining the value of the public good and/or private good, and the number of attacks functions as a direct measure of the degree of adversity.

Additionally, in both Models 1 and 2, we look at the effect of an increase in the value of the public good on the probability of cooperation, and do the same in Model 2 for an increase in the value of the private good. Clearly, such higher values cannot be

² In the literature on cooperative breeding, the *ecological constraints hypothesis* (Emlen, 1982; Hatchwell and Komdeur, 2000) states that cooperative breeding takes place under the ecological constraint of few alternative breeding grounds being available, making dispersal a bad option. E.g., Sandoval and Wilson (2012) interpret the ecological constraints hypothesis as a general effect of adversity on the probability of cooperation, and Bergmüller et al. (2007) connect it to theories of cooperation.

³ The boomerang effect (respectively: sucker effect) refers to the fact that, when the other player cooperates (defects), the individual player is better (worse) off when cooperating rather than defecting. When this effect operates only under a sufficiently harsh environment (but not under “normal” environment), this suggests a positive (negative) impact of the degree of adversity on the probability of cooperation. We refer to this positive impact as the “common-enemy hypothesis” (respectively: competing hypothesis).

⁴ Note that the degree of complementarity does not constitute a direct measure of adversity: adversity literally takes the form of a larger prey, and a larger degree of complementarity could be said to only be a side-effect of facing a larger prey. Yet, the degree of complementarity proxies for the degree of adversity, as it is positively correlated with it.

considered as constituting a harsher environment, as for a given number of cooperating players, an increase in value makes the players strictly better off, rather than worse off. The reason for still looking at the effect of an increase in the value of the public or private good in the context of adversity is an expositional one. At a basic level, one could argue that if adversity favors cooperation, it is simply because it decreases the individual player's cost-benefit ratio of cooperating. Indeed, it is this straightforward argument that seems to underlie the prediction of Mesterton-Gibbons and Dugatkin (1992, 1997) that a harsh environment facilitates cooperation. A key point in our paper is that the effect of increased adversity cannot always be equated to a decrease in the cost-benefit ratio of cooperation. The analysis of the effect of an increase in the value of the public or private good (which is tantamount to a decrease in the cost-benefit ratio of cooperating), serves as a benchmark to identify the manner in which the effect of increased adversity operates in a different manner.

The paper is structured as follows. Before we construct our micro-foundations model, in Section 2, we set out a general framework for two-player two-strategy games. The advantage of this framework is that it enables us to describe in a succinct manner four scenarios for the effect of adversity that consistently appear in all variants of our micro-foundations model, allowing us to refer back to these scenarios once we describe our results. In the micro-foundations model of Section 3, cooperating consists of contributing to the production of a public good, and adversity is proxied by the degree of complementarity between players' efforts in contributing to the public good. In Section 4, cooperating consists of the defense of a public good, and/or private good with by-product benefits, and adversity is measured by the number of random attacks. In Section 5, we compare our model of adversity to models that study respectively the effect of the degree of synergy between players' cooperative efforts, and of the degree of interdependence between players. We end with some conclusions in Section 6.

2. Two-player two-strategy games without micro-foundations for the effect of adversity

We start by setting out the general framework for the simple two-player two-strategy games we treat (cf., e.g., Doebeli and Hauert, 2005; Archetti et al., 2011; Archetti and Scheuring, 2012), with strategies C (=cooperate) and D (=defect). The payoff matrix is given in (1). R is the reward each player receives from joint cooperation. T is the temptation payoff obtained from defecting when the other player cooperates. S is the sucker payoff obtained when cooperating even though the other player defects. Finally, P is the punishment payoff received when both players defect.

$$\begin{matrix} & C & D \\ \begin{matrix} C \\ D \end{matrix} & \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{matrix} \quad (1)$$

These payoffs are assumed to satisfy $R > P$, so that the sum of the players' payoffs is larger if both play C than if both play D . Define $(R - T)$ as the added payoff to the individual player of cooperating rather than defecting when the other player is cooperating as well, or in short the *added payoff of cooperating jointly*. Define $(S - P)$ as the added payoff to the individual player of cooperating rather than defecting when the other player defects, or in short the *added payoff of cooperating alone*.

If $(R - T) > 0$, there is a so-called *boomerang effect* (Mesterton-Gibbons and Dugatkin, 1992, 1997): the unilateral defector suffers so much from unilaterally defecting that it is better to cooperate. If $(R - T) < 0$, there is a *free-rider effect*, and the unilateral defector is

better off defecting. If $(S - P) < 0$, there is a *sucker effect*: a player who unilaterally cooperates is worse off than in case he defects. If $(S - P) > 0$, there is what we call a *hero effect*: if the other player defects, the individual player is still better off when unilaterally cooperating. The effects as a function of the added payoffs are summarized in Table 1.

As a function of the four possible combinations of effects, we obtain four games, as summarized in Table 2 (cf. Sigmund, 2010). If $(R - T) < 0$ and $(S - P) < 0$, the game is a Prisoner's Dilemma (PD) (Tucker, 1950), and joint defection is the unique evolutionary stable strategy or ESS (Maynard Smith and Price, 1973). If $(R - T) > 0$ and $(S - P) < 0$, the game is a Stag Hunt (SH) (Skyrms, 2004), and both joint defection and joint cooperation are ESS's (the game additionally has a mixed Nash equilibrium, which is not an ESS). In order to make predictions about the probability with which each of the two ESS's in the Stag Hunt is played, we consider the relative sizes of their basins of attractions. Conceptually, any two-player game is then played by generation upon generation, where in each generation a large population of players is randomly matched in pairs to play the game. Consider the proportion of cooperating players at any given point of time. Then under replicator dynamics (Taylor and Jonker, 1978), the growth rate of the proportion of cooperating players in the population equals the difference between the average payoff obtained from cooperating, and the average overall payoff of the population. Therefore, in the Stag Hunt, as soon as the proportion of cooperating players exceeds a certain threshold p^* , the population lies in the basin of attraction of joint cooperation, and this ESS evolves (see Mesterton-Gibbons and Dugatkin (1992, p. 281)). Oppositely, if the proportion of cooperating players is lower than p^* , the population lies in the basin of attraction of joint defection, and this ESS evolves. To calculate the threshold, note that the individual player is indifferent between cooperating and not cooperating when $p^*R + (1 - p^*)S = p^*T + (1 - p^*)P$, so that

$$p^* = (P - S) / (R - T + P - S). \quad (2)$$

Considering any initial population state as equally likely, it follows that in the Stag Hunt, joint cooperation evolves with probability $(1 - p^*)$, and joint defection with probability p^* . We consider these as the probabilities that each of the ESS's is played.

If $(R - T) > 0$ and $(S - P) > 0$, the game is a game of no conflict or Harmony Game (HG), and joint cooperation is the unique ESS. Finally, if $(R - T) < 0$ and $(S - P) > 0$, the game is a Snowdrift game

Table 1
Four effects as function of sign of added payoff of cooperating jointly/alone ($(R - T)$, respectively $(S - P)$).

		$(S - P)$	
		< 0	> 0
$(R - T)$	< 0	Free-rider Sucker	Free-rider Hero
	> 0	Boomerang Sucker	Boomerang Hero

Table 2
Four effects as function of sign of added payoff of cooperating jointly/alone ($(R - T)$, respectively $(S - P)$).

		$(S - P)$	
		< 0	> 0
$(R - T)$	< 0	Prisoner's Dilemma (PD)	Snowdrift game (SD)
	> 0	Stag Hunt (SH)	Harmony Game (HG)

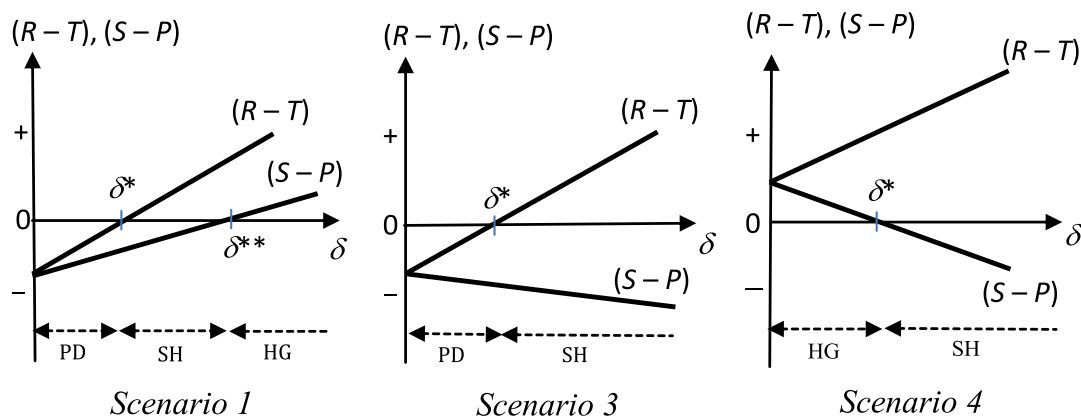


Fig. 1. Four scenarios for the effect parameter δ (mainly interpreted as the degree of adversity) on the added payoff of joint/alone cooperation. **Scenario 2** (not represented) is identical to **Scenario 1**, except that the labels $(R-T)$ and $(S-P)$ are reversed, and for intermediate δ the game is a Snowdrift game rather than a Stag Hunt game.

(SD) (Sugden, 1986; also called Chicken game (Russell, 1959), or Hawk–Dove game (Maynard Smith and Price, 1973)). The Snowdrift game has three Nash equilibria, namely two strict equilibria where one player cooperates and the other player defects, and a mixed Nash equilibrium where players randomize between cooperating and defecting. Looking at the Snowdrift game as an evolutionary game, if players cannot condition their strategies on their current position or role, the mixed equilibrium is the unique ESS.⁵ In the mixed equilibrium, a proportion p of the players cooperate such that $pR + (1-p)S = pT + (1-p)P$, meaning that

$$p = (S-P)/(T-R+S-P) \quad (3)$$

Compared to the value p^* in (2), note that p in (3) is not the probability that a joint defection equilibrium is reached rather than a joint cooperation equilibrium, but instead is the proportion of cooperating players within the population (who coexist with defecting players).

Following Hauert et al. (2006), we do not see the games in Table 2 as separate and isolated categories, but as a single game, the characteristics of which change as a function of a parameter denoted δ – in the first instance interpreted as the degree of adversity. The common-enemy hypothesis applies when a higher degree of adversity leads to a larger basin of attraction for joint cooperation ($\partial p^*/\partial \delta < 0$), or to a higher proportion of cooperating players in a mixed population of cooperating and defecting players ($\partial p/\partial \delta > 0$). The competing hypothesis applies when a higher degree of adversity leads to a smaller basin of attraction for joint cooperation ($\partial p^*/\partial \delta > 0$), or to a lower proportion of cooperating players in a mixed population of cooperating and defecting players ($\partial p/\partial \delta < 0$).

In our analysis, we find four main scenarios (referred to as Scenarios 1–4) for the manner in which a higher degree of adversity (proxied by the degree of complementarity, or measured by the number of attacks) affects the game played, summarized in Fig. 1. We present these here without providing the micro-foundations we will provide later on. This allows us to refer back to the scenarios, and to show that, across the different models we treat, the effects of adversity we find typically fit one of these four scenarios (Sections 3.1 and 4.1). More generally, our analysis will also look at the effect when the parameter δ does not take the form of adversity, but where the comparison with adversity is

insightful. In particular, we will look at the case where δ is the value of the public or private good (Sections 3.2 and 4.2). Additionally, in Section 5 we shortly treat two other models in the cooperation literature where δ either equals the degree of synergy between the players' efforts (following Hauert et al., 2006), or equals the degree of interdependence between the players (following Roberts, 2005).

All models treated have in common that the added payoff of cooperating jointly ($R-T$) is positively affected by δ , but differ, first, according to the relation between the added payoff of cooperating jointly and the added payoff of cooperating alone ($S-P$) and, second, according to the way in which the added payoff of cooperating alone is affected by δ .

Scenario 1. This scenario replicates the single scenario considered by Mesterton-Gibbons and Dugatkin (1992, 1997). The distinguishing features of Scenario 1 are that (except possibly for non-minimal δ) we have $(R-T) > (S-P)$, and that $(R-T)$ and $(S-P)$ both increase in δ . Fig. 1 represents a particular example where $(R-T)$ and $(S-P)$ are linear in δ , and where $(R-T)$ is steeper than $(S-P)$ – but we consider any case with the listed distinguishing features as an instance of Scenario 1. In Scenario 1, the added payoffs are both negative for a small δ (meaning that both the free-rider effect and the sucker effect apply), but both positive for a large δ (meaning that both the boomerang effect and the hero effect apply). An increase in δ first turns the free-rider effect into a boomerang effect, and the latter effect becomes stronger as δ is further increased. In the same manner, the sucker effect turns into a hero effect, and the latter effect is further reinforced as δ becomes larger. As illustrated in Fig. 1, critical levels δ^* and δ^{**} (with $\delta^* < \delta^{**}$) exist such that the following applies: the game played is a Prisoner's Dilemma for low δ ($\delta < \delta^*$), a Stag Hunt for intermediate δ ($\delta^* < \delta < \delta^{**}$), and a Harmony Game for high δ ($\delta > \delta^{**}$). When δ is interpreted as the degree of adversity, the manner in which the game changes as the degree of adversity increases illustrates the common-enemy hypothesis that a harsher environment increases the probability of cooperation (even though it is not specified in this simple framework how the basin of attraction of joint cooperation is affected by adversity in the Stag Hunt).

Scenario 2. This scenario is similar to Scenario 1, in that both $(S-P)$ and $(R-T)$ continue to increase in δ . Yet, for any non-minimal δ , we now have $(S-P) > (R-T)$. The only difference with Scenario 1 is that for intermediate δ , the Snowdrift game is played rather than the Stag Hunt. Because of its similarity with Scenario 1, this case is not separately depicted in Fig. 1. Also, an intermediate case between Scenarios 1 and 2 will appear in our analysis, where

⁵ If players are able to condition their strategies on their current position or role, the two strict Nash equilibria are ESS's, and the mixed Nash equilibrium is not an ESS (Weibull, 1995, pp. 67–68). It is now easy to see that the two strict equilibria have equal basins of attraction, so that if each initial population state is equally likely, the two strict equilibria are equally likely to evolve. In the paper, we focus on the case where players cannot take on roles.

$(R-T)$ and $(S-P)$ coincide and both increase in δ . In this case an increase in δ leads to a direct transition from the Prisoner's Dilemma to the Harmony Game, without the intermediate case of a Stag Hunt or Snowdrift game.

Scenario 3. This scenario is similar to [Scenario 1](#), in that $(R-T) > (S-P)$ for any non-minimal δ , and in that $(S-P)$ and $(R-T)$ are both negative for low δ . Yet, the added payoff of cooperating alone now *decreases* in δ , and is therefore always negative. [Fig. 1](#) again presents a particular instance of this scenario. It continues to be the case that, as δ is increased, the free-rider effect turns into a boomerang effect, and the latter effect becomes stronger; yet, at the same time, the sucker effect is always present, and becomes stronger as δ is increased. As $(R-T)$ increases in δ , a critical degree δ^* exists such that for $\delta < \delta^*$, players play a Prisoner's Dilemma, and for $\delta > \delta^*$, they play a Stag Hunt. This switch from a Prisoner's Dilemma to a Stag Hunt again is in line with the common-enemy hypothesis, however, the effect of a common enemy is not strong enough to lead players to play a Harmony Game.

Scenario 4. This scenario is identical to [Scenario 3](#), with the exception that for low δ , the two added payoffs are positive. This may be due to, e.g., a lower cost of cooperating. [Fig. 1](#) shows one particular instance of this scenario. As δ is increased, there initially is both a boomerang effect and a hero effect. As δ is further increased, the boomerang effect becomes stronger; yet, at the same time, the hero effect eventually turns into a sucker effect, and the latter effect is further reinforced as δ becomes larger. It follows that $(R-T)$ is always positive, so that players never play a Prisoner's Dilemma. As $(S-P)$ decreases in δ , a critical δ^* exists such that for $\delta < \delta^*$, players play a Harmony Game, and for $\delta > \delta^*$, they play a Stag Hunt. When δ is interpreted as the degree of adversity, in this scenario, the effect of a harsh environment is completely reversed: a higher degree of adversity means that joint defection also becomes an ESS, and a competing hypothesis to the common-enemy hypothesis is obtained.

We now show that these four main scenarios systematically arise when modeling the micro-foundations of how the payoffs R , S , T and P are affected by the degree of adversity. Our results are summarized in [Table 3](#). In Model 1 ([Section 3](#)), cooperating means

Table 3
Summary of the different cases obtained in our analysis. The cases that arise in Model 1 (see [Section 3](#)) and in Model 2 (see [Section 4](#)) for the effect of the parameter δ on the added benefit of cooperating jointly/alone ($(R-T)$, respectively $(S-P)$), follow four possible Scenarios (introduced in [Section 2](#), and summarized in [Fig. 1](#)). \uparrow (\downarrow) denotes an increasing (decreasing) function of δ . V is the value of the public good, G the value of the private good, c the cost of cooperating, k the degree of complementarity, A the number of attacks, and α the weight put on the public good. The parameter δ equals either the degree of complementarity (k) (Model 1a, [Section 3.1](#)), the value of the public good (V) (Model 1b, [Section 3.2](#)), the number of attacks (A) (Model 2a, [Section 4.1](#)), or both the value of the public good (V) and of the private good (G) (Model 2b, [Section 4.2](#)).

	Scenario 1: $(R-T)\uparrow$ $(S-P)\uparrow$ $(R-T) > (S-P)$	Scenario 2: $(R-T)\uparrow$ $(S-P)\uparrow$ $(S-P) > (R-T)$	Scenario 3: $(R-T)\uparrow$ $(S-P)\downarrow$ Large c	Scenario 4: $(R-T)\uparrow$ $(S-P)\downarrow$ Small c
Model 1: Public good production	Case 1b.I: $\delta=V$, $k > 0.5$	Case 1b.II: $\delta=V$, $k < 0.5$	Case 1a.I: $\delta=k$, $k > 0.5$	Case 1a.II: $\delta=k$, $k > 0.5$
Model 2: Public/private good defense	Case 2a.III: $\delta=A$, $0.5 < k < 2/3$, or $k > 2/3$ and α small Case 2b.I: $\delta=V$ or G , $k > 0.5$	Case 2a.IV: $\delta=A$, $k < 0.5$	Case 2a.I: $\delta=A$, $k > 2/3$, α large	Case 2a.II: $\delta=A$, $k > 2/3$, α large
	Case 2b.II: $\delta=V$ or G , $k < 0.5$			

producing a public good (where adversity is proxied by the degree of complementarity between the players' efforts in producing the public good), and in Model 2 ([Section 4](#)) cooperation means defending an existing public good, and/or a private good with by-product benefits against adversity (where adversity is measured by the number of random attacks the players face). We provide both intuitions in the body of the paper, and formal results in the [Appendix](#), where the [Appendix](#) also shows how the basin of attraction of cooperation in the Stag Hunt, and the proportion of cooperating players in the mixed population in the Snowdrift game, is affected by the parameter δ .

3. Model 1: production of a public good

In Model 1, a player who cooperates (plays C) incurs a cost c , and contributes to the joint production of a public good (e.g., the catch of a large prey).⁶ Playing D (defecting) comes at no cost and means not providing any effort to the production of the public good. When both players cooperate, they jointly produce a single public good. The benefit obtained from the public good is always the same for both players (i.e., the public good is non-excludable, see [Dionisio and Gordo \(2006\)](#)); e.g., in case the players are predators that can catch a large prey, it may be that none of the predators can be excluded from consuming the prey.

Following [Hirschleifer's \(1983\)](#) taxonomy of possible "technologies" for the production of a public good, we consider three benchmark cases (see [Table 4](#)), and all cases in between ([Table 4](#) also refers to our model of the defense of a public good, for which [Table 4](#) is equally relevant, as will be explained in [Section 4](#) below). For all technologies, it holds that if both players cooperate, the value of the public good is V , and if both defect the value of the public good is zero. The technologies differ only according to the value of the public good that is obtained when one player cooperates and the other player defects. In order to distinguish between the technologies by means of a single parameter, we write down the value of the public good when one player cooperates and the other player defects as $(1-k)V$, where k ($0 \leq k \leq 1$) is interpreted as the *degree of complementarity* between the players' cooperative efforts.⁷ Considering C (respectively D) as an effort with a value 1 (respectively 0), the production function of the public good can be written as $f(x_1, x_2)V$, where x_1 and x_2 are the values of Player 1's, respectively Player 2's efforts, and where $f(\cdot)$ is the production function. Then $f(1, 1) = 1$, $f(0, 0) = 0$, and $f(1, 0) = f(0, 1) = (1-k)V$.

⁶ In the standard example attached to a Stag Hunt, cooperating means contributing to the hunt of a big prey, whereas defecting means individually hunting a smaller prey ([Skyrms, 2004](#)). Put otherwise, in this standard example, cooperating means contributing to the production of a public good, while defecting means individually producing a private good. Yet, in game theory, the name Stag Hunt more generally refers to the payoff combination in the bottom-left corner of [Table 2](#), and not specifically to this example. In our model, defecting does not mean that a private good is produced, but that one does not contribute to the production of the public good. Still, our analysis is not qualitatively changed if defecting instead means producing a private good, as the added payoff of cooperating jointly and alone then simply shift up by the same amount, effectively making cooperating relatively more costly.

⁷ For a more general production function, with n players and where the typical player i chooses an effort from a continuum of efforts, the so-called constant-elasticity-of-substitution production function may be considered, taking the general form

$$\left[\sum_{i=0}^n x_i^{1-\sigma} \right]^{1/(1-\sigma)} \quad (\text{see, e.g., } \text{Ray et al. (2007)}). \quad \sigma \text{ is the degree of complementarity, with the best-shot (weakest-link) production function obtained for } \sigma \text{ approaching minus (plus) infinity, and the summation production function reached for } \sigma=0. \text{ Just as is the case in our simplified production function, } \sigma \text{ does not affect the value obtained when all players do the same effort, but decreases the value when not all players do the same effort.}$$

Table 4

Hirschleifer production functions for the production of a public good (Model 1: $x_i=0$ if player i defects; $x_i=1$ if player i cooperates), and for the defense of a public good (Model 2: $x_i=0$ if player i defects and is attacked; $x_i=1$ if player i cooperates, or defects but is not attacked); corresponding degree of complementarity (k) between the players' efforts, and corresponding value obtained from the public good when only one player cooperates.

Technology	Production function	Degree of complementarity	Value public good when one player puts $x_i=1$: $(1-k)V$
Best-shot	$\text{Max}(x_1, x_2)V$	$k=0$	V
Summation	$[(x_1 + x_2)/2]V$	$k=0.5$	$V/2$
Weakest-link	$\text{Min}(x_1, x_2)V$	$k=1$	0

The following three benchmark technologies can now be considered. With a *best-shot technology*, the degree of complementarity between efforts is minimal ($k=0$) and one cooperating player is sufficient to produce the maximal possible value of the public good (e.g., one predator attacking a weak, large prey may be sufficient to generate a public good for a group of predators). Thus, $(1-k)V=V$. The production function of the public good in this case can be written as $\text{max}(x_1, x_2)V$. With a *summation technology*, each additional cooperating player adds the same value to the public good ($k=0.5$). In this case, when only one player cooperates, $(1-k)V=0.5V$ and exactly half of the value of the public good gets produced. The production function of the public good can now be written as $[(x_1+x_2)/2]V$ (e.g., the more predators join a hunt, the more likely it is that a large prey is caught).⁸ Finally, with a *weakest-link technology*, the degree of complementarity between efforts is maximal ($k=1$) and both players need to cooperate for any value of the public good to be generated (e.g., only the joint efforts of all allows a group of predators to catch a large, strong prey). The payoff when only one player cooperates now equals $(1-k)V=0$. The production function of the public good can in this case be written as $\text{min}(x_1, x_2)V$. One way in which the weakest-link technology may apply is when there is a division of labor between the players in producing the public good (cf. [Leimar and Connor, 2003](#)).⁹ Such a division of labor may be most plausible in small groups (for a similar argument, see [Maynard Smith \(1983, p. 448\)](#), and [Nunn and Lewis \(2001, p. 60\)](#)), so that in this sense, apart from the advantage of simplicity, focusing on a small group of two players makes sense.

We obtain the following values for R , T , S and P ¹⁰:

$$R = V - c \quad (4)$$

⁸ In the standard public goods game (see, e.g., [Archetti and Scheuring \(2012\)](#)), the value of the public good obtained by the individual player equals $r^*c^*n_c/n$. n_c is the number of cooperating players. The cost of cooperating c is seen as an investment. All investments are added, multiplied by a factor r , and divided over all players. Putting $r^*c^*=V$, the standard public goods game fits one to one with the case $k=1/2$ in our model.

⁹ As noted by one of the referees, the interpretation of the weakest-link technology as modeling a division of labor imposes an asymmetric interpretation on a symmetric model. If cooperation takes place when each of the two players specializes in doing a particular type of effort, then missing in our model is a process determining which player does which effort. We do not undertake this exercise here, but note that in any case, the weakest-link technology equally applies in a symmetric model where each player's effort takes exactly the same form, but is critical for the production of the public good.

¹⁰ In an earlier version of this paper, by symmetry of the model in [Section 4](#), cooperating also meant producing a private good with value G , receiving a weight of $(1-\alpha)$, and creating a by-product benefit eG to the other player (with $0 \leq e \leq 1$). Yet, as pointed out by one of the referees, the value obtained from private good is not affected by the degree of complementarity, and the presence of a private good therefore plays a passive role in the analysis of Model 1. We therefore omit the private good from the analysis of Model 1, even though the production of a private good with a by-product benefit falls under the definition of by-product mutualism.

$$S = (1-k)V - c \quad (5)$$

$$T = (1-k)V \quad (6)$$

$$P = 0. \quad (7)$$

This means that the added payoff of cooperating jointly and the added payoff of cooperating alone, respectively, take on the form:

$$(R-T) = kV - c \quad (8)$$

$$(S-P) = (1-k)V - c. \quad (9)$$

Finally, we can calculate the difference between the added payoff of cooperating jointly, and the added payoff of cooperating alone, as

$$(R-T) - (S-P) = (2k-1)V. \quad (10)$$

It follows that in case of a large degree of complementarity between the players' efforts ($k > 1/2$), the added payoff of cooperating jointly is larger than the added payoff of cooperating alone; in case of a small degree of complementarity ($k < 1/2$), the opposite is the case. When the degree of complementarity is neither large nor small ($k = 1/2$), the two added payoffs are exactly equal.

When the game is a Stag Hunt ($(R-T) > 0$, $(S-P) < 0$), using Eqs. (2) and (4)–(7), the probability p^* that joint defection is played equals:

$$p^* = [c - (1-k)V] / [(2k-1)V] \quad (11)$$

When we have a Snowdrift game ($(R-T) < 0$, $(S-P) > 0$), using Eqs. (3)–(7), the probability p that joint action is played in the corresponding mixed equilibrium equals:

$$p = [(1-k)V - c] / [(1-2k)V] \quad (12)$$

In the following, we use Model 1 to look at the effect on the probability of joint cooperation and on the proportion of cooperating players in the population, first, of the degree of complementarity between the players' cooperative efforts (Model 1a, $\delta=k$), and second, of the value of the public good (Model 1b, $\delta=V$).

3.1. Model 1a: effect of degree of complementarity between cooperative efforts

We first look at the effect of a higher degree of complementarity between the players' cooperative efforts ($\delta=k$). As summarized in [Fig. 2](#), we obtain four cases, which we here treat at an intuitive level, and of which we then formally show the existence in the [Appendix](#). In order to provide intuitions, we use the example of cooperative hunting by lions ([Scheel and Packer, 1991](#); [Stander, 1992](#)), even though Model 1a is not restricted to this particular application.

Case 1a.I (cf. Scenario 3). Take as a starting point the case where the value lions obtain from cooperative hunting is produced according to the summation technology ($k=0.5$). The first lion to join the hunt then contributes as much to the value of the hunt as the second lion joining the hunt (e.g., because one extra lion increases the probability of a successful hunt by a fixed amount), meaning that the added payoff of cooperating jointly and alone are equal ($(R-T)=(S-P)$). Assume in [Case 1a.I](#) that the cost of cooperation is sufficiently large for both these added payoffs to be negative for the summation technology, so that both the free-rider effect and the sucker effect apply, meaning that the lions play a Prisoner's Dilemma. As we now increase the degree of complementarity between the lions' hunting efforts ([Fig. 2\(a\)](#)), to the right of $k=0.5$, and move closer to the weakest-link technology ($k=1$) where the hunt can only be successful if both lions join the hunt, it is clear that the free-rider effect will turn into a

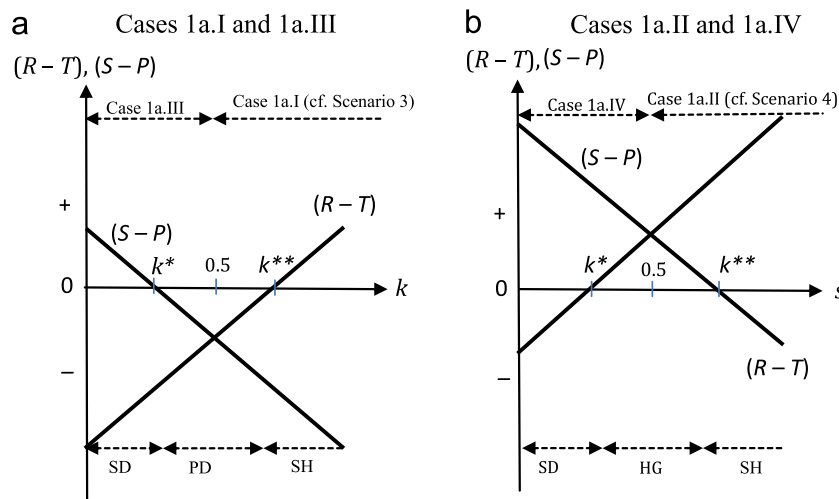


Fig. 2. Model 1a. Added payoff of joint/loose cooperation ($(R-T)$, respectively $(S-P)$) as function of degree of complementarity ($\delta=k$), for large cooperation costs (left) and small cooperation costs (right).

boomerang effect. As the degree of complementarity is increased, the individual lion has fewer incentives to unilaterally deviate from joint cooperation. At the same time, the sucker effect is maintained and even reinforced as we increase the degree of complementarity: if it is already the case for the summation technology that the individual lion does not want to hunt alone, it will certainly be the case for a higher degree of complementarity between hunting efforts. This means that the added payoff of cooperating alone decreases in the degree of complementarity. Yet, because the free-rider effect turns into a boomerang effect as the degree of complementarity is increased, a higher degree of complementarity increases the probability of joint cooperation. When the degree of complementarity is positively correlated with the size of the prey, which is itself interpreted as the degree of adversity, this is in accordance with the common-enemy hypothesis that a higher degree of adversity fosters cooperation. Case 1a.I is in line with Scenario 3 in Fig. 1.

Case 1a.II (cf. Scenario 4). Take again as a starting point the case where the value lions obtain from cooperative hunting is produced according to the summation technology ($k=0.5$). Assume, however, that the cost of cooperation is sufficiently low for both the added payoffs to be positive for the summation technology, so that both the boomerang effect and the hero effect apply. As we now increase the degree of complementarity (Fig. 2(b), to the right of $k=0.5$) and move closer to the weakest-link technology ($k=1$), the hero effect turns into a sucker effect: as we make it more critical that a second lion joins the hunt, the added payoff of hunting by oneself becomes smaller. Again, the added payoff of cooperating alone decreases in the degree of complementarity. At the same time, the boomerang effect is maintained as we increase the degree of complementarity, and is even reinforced: the higher the degree of complementarity, the lower the benefit of unilaterally deviating from joint hunting. It follows that, for the specific case of low cooperation costs, and intermediate to high degrees of complementarity, as we increase the degree of complementarity, the game changes from a Harmony Game into a Stag Hunt. When the degree of complementarity is again positively correlated with the size of the prey, which in turn measures the degree of adversity, a competing hypothesis to the one of Mesterton-Gibbons and Dugatkin (1992, 1997) is obtained, namely that a higher degree of

adversity causes a lower probability of cooperative hunting. Case 1a.II is in line with Scenario 4 in Fig. 1.

We also treat two cases where the hunting technology is such that the contribution of the first lion joining the hunt is larger than the contribution of the second lion ($(S-P) > (R-T)$), giving rise to the possibility that lions play a Snowdrift game. While these two additional cases do not fit any of our four main scenarios in Fig. 1, they are of interest given the emphasis that part of the literature on the evolution of cooperation puts on Snowdrift games (e.g. Doebeli and Hauert, 2005).

Case 1a.III. Just as in Case 1a.I, we again start from the case where the value lions obtain from the prey is produced according to the summation technology ($k=0.5$), and assume that the cost of cooperation is large. The free-rider effect and the sucker effect again both apply, and the lions play a Prisoner's Dilemma. Rather than increasing the degree of complementarity between the lions' hunting efforts, we now decrease this degree (Fig. 2(a), to the left of $k=0.5$), and move closer to the best-shot technology ($k=0$), where a single hunting lion suffices to produce a maximal value obtained from the prey. It follows now that as we decrease the degree of complementarity, the free-rider effect is further reinforced, as there are even more incentives to unilaterally deviate from joint cooperation. At the same time, as we decrease the degree of complementarity, the sucker effect eventually turns into a hero effect, and the lions play a Snowdrift game. For the specific case of large cooperation costs, and low to intermediate degrees of complementarity, lions therefore play a Snowdrift game for a degree of complementarity close to the best-shot technology, and play a Prisoner's dilemma as the degree of complementarity is increased. Interpreting the degree of complementarity as a proxy for the degree of adversity ($\delta=k$), Case 1a.III predicts that a higher degree of adversity decreases the probability of cooperation.

Case 1a.IV. Just as in Case 1a.II, we start from the case where the value lions obtain from the prey is produced according to the summation technology ($k=0.5$), and assume that the cost of cooperation is small. The boomerang effect and the hero effect both apply, and the lions play a Harmony Game. As we decrease the degree of complementarity (Fig. 2(b), to the left of $k=0.5$) and move closer to the best-shot technology ($k=0$), the hero effect is reinforced, as the individual lion has even more incentives to unilaterally deviate from joint defection. Yet, at the same time, the boomerang effect turns into a free-rider effect, as close to a best-

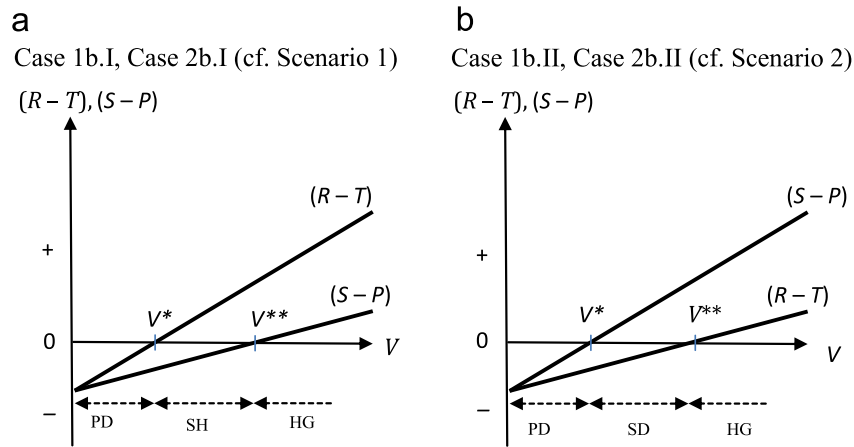


Fig. 3. Models 1b and 2b. Added payoff of joint/alone cooperation ($(R-T)$, respectively $(S-P)$) as function of value of public good (V), for large degree of complementarity (Cases 1b.I and 2b.I) and small degree of complementarity (Cases 1b.II and 2b.II).

shot technology the individual lion will find it optimal to defect when the starting position is joint hunting. It follows that for the specific case of low to intermediate degrees of complementarity and low cooperation costs, for a degree of complementarity close to the best-shot technology, the game is a Snowdrift game, and that as we increase the degree of complementarity, the game turns into a Harmony Game. Therefore, interpreting the degree of complementarity as a proxy for the degree of adversity ($\delta=k$), Case 1a.IV predicts that a higher degree of adversity increases the probability of cooperation.

3.2. Model 1b: effect of the value of the public or the private good

All else equal, a larger value of the public good increases both the added payoff of cooperating alone, and the added payoff of cooperating jointly. For a small value of the public good, both the free-rider and the sucker effect may apply, and players may therefore play a Prisoner's Dilemma. Yet, for a large value of the public good, the free-rider effect may turn into a boomerang effect, and the sucker effect into a hero effect, so that players play a Harmony Game. The type of game players play for intermediate values of the public good, depends on the level of the fixed degree of complementarity, leading to two cases (see Fig. 3). We here treat these cases at an intuitive level, and formally show their existence in the Appendix. Fig. 3 also refers to our model of the defense of a public good, because the analysis there is very similar, as will be explained in Section 4.2 below.

Case 1b.I (cf. Scenario 1). In this case, the degree of complementarity is fixed at a high level ($k > 0.5$), so that a second cooperating player contributes more than a first cooperating player. The added payoff of cooperating jointly is therefore everywhere larger than the added payoff of cooperating alone ($(R-T) > (S-P)$), meaning that as we increase the value of the public good, the former added payoff becomes positive before the latter does. Put otherwise, for a value of the public good large enough to turn the free-rider effect into a boomerang effect, the sucker effect still exists. It follows that for intermediate values of the public good, players play a Stag Hunt. As the value of the public good is increased, we therefore move from a Prisoner's Dilemma, to a Stag Hunt, to finally a Harmony Game, in line with Scenario 1 in Fig. 1, and with the picture proposed by Mesterton-Gibbons and Dugatkin (1992, 1997) for the effect of adversity (see Fig. 3(a)).

Case 1b.II (cf. Scenario 2). In this case, the degree of complementarity is fixed at a low level ($k < 0.5$), so that a second cooperating player contributes less than a first cooperating player.

The added payoff of cooperating alone is therefore everywhere larger than the added payoff of cooperating jointly ($(S-P) > (R-T)$), meaning that as we increase the value of the public good, the former added payoff becomes larger than zero before the latter added payoff does. Put otherwise, for a value of the public good that turns the sucker effect into a hero effect, the free-rider effect may still exist. Therefore, for intermediate values of the public good, players play a Snowdrift game, rather than a Stag Hunt as in Case 1b.I. It follows that Case 1b.II is in line with Scenario 2 for the effect of adversity in Fig. 1 (see Fig. 3(b)).

Exactly in between Cases 1b.I and 1b.II, the added payoff of cooperating jointly and alone is always equal when the degree of complementarity is exactly intermediate ($k = 1/2$), and both identical added payoffs increase in the value of the public good. In this intermediate case, one directly moves from a Prisoner's Dilemma to a Harmony Game.

A comparison between the effect of the degree of complementarity on the probability of cooperation, and the effect of an increase in the value of the public good reveals a fundamental difference. While both an increase in the degree of complementarity and an increase in the value of the public good increase the added payoff of cooperating jointly, the added payoff of cooperating alone increases in the value of the public good, but decreases in the degree of complementarity. Because of this decrease, the effect of the degree of complementarity cannot be likened to a simple decrease in the cost-benefit ratio of cooperating. A higher degree of complementarity decreases the cost-benefit ratio of cooperating when the other player cooperates, but increases this ratio when the other player does not cooperate. For this reason, as seen in the first row of Table 3, the effect of an increase in the value of the public good fits Scenarios 1 and 2, while the effect of the degree of complementarity fits Scenarios 3 and 4.

4. Model 2: defense of public and private goods

In Model 2, Nature or one or more outside players, randomly and repeatedly attack our two players, where the number of attacks is exogenously given, and measures the degree of adversity facing the players. Playing C (cooperating) now means exerting effort to defend an existing public good against these random attacks, rather than exerting effort to produce a public good, as in Model 1; again, playing C comes at cost c . Playing D (defecting) means not doing any effort to defend; this again comes at zero cost. The public good is, e.g., a common territory that the players collectively defend. To be able to reflect in our model some key

examples of by-product mutualism, we also consider the case where the individual player who cooperates alternatively defends a private good with by-product benefits,¹¹ or where cooperating means defending a public good and a private good with by-product benefits at the same time. In case cooperating only means defending a private good, e.g., an individual player, after having found a food source, produces for himself the private benefit of being safeguarded from predation risk by issuing a call to attract conspecifics, who obtain a by-product benefit in the form of a share in the food source. In case cooperating generates both a public good and a private good, the cooperating player may be thought of as a sentinel, who both produces the public good of his group being safeguarded from predation, and a private benefit as he is able to escape the predator first.¹²

The model is adapted from De Jaegher and Hoyer (forthcoming), who only consider the extreme case of a single attack whereas here we consider any number of attacks. The sampling process by which a random player is repeatedly selected to be attacked is one with replacement, so that by coincidence it may happen that only one player is ever attacked, because the same player is sampled again and again. In particular, with a number of attacks A , the probability that only a single player is attacked equals $1/(2^A)$; for a higher number of attacks it is therefore more and more likely that both players are attacked.

The players provide defensive inputs with value $x_i=0$ or 1 in a defensive production function of the form $f(x_1, x_2)V$ for the public good, and of the form x_iG for the private good, where G is the value of the private good. A player provides a defensive input with value 1 when cooperating (=exerting cooperative effort), but also when defecting (=not providing cooperative effort) and never being attacked (this contrasts with Model 1, where defecting always means providing an input with value zero in the production function of the public good). For instance, if defending means guarding one's side of a common territory, then failing to guard does not have any consequences for the common territory if one's side happens not to be attacked. A player provides defensive input with value 0 when defecting, and when additionally being attacked at least one time. Note that the defensive inputs do not produce the public good and/or private good, but rather maintain their values: if the number of attacks would be zero, the players would obtain the full value of the private and/or public good even when they defect.

For all defensive production functions of the public good that we consider, it is the case that $f(1, 1)=1$, and $f(0, 0)=0$: if there is no player who at the same time defects and is attacked, then the full value of the public good is maintained; if both players defect and are attacked, then the entire value of the public good is lost. Defensive production functions of the public good only differ according to the value that is obtained when one player defects and is attacked, whereas the other player either cooperates, or defects but is not attacked. In particular, $f(1, 0)=f(0, 1)=(1-k)V$, where k is the degree of complementarity between the players'

defensive inputs. Reinterpreting x_1 and x_2 in Table 1 as the values of the players' defensive inputs, we can consider the same benchmark production functions as listed there. For $k=0$ (best-shot technology), a single player who defends, or who does not defend but is not attacked, suffices to maintain the full value of the public good. For $k=1/2$ (summation technology), each player who defends, or who does not defend but is not attacked, adds additively to the value of the public good. For $k=1$ (weakest-link technology), a single player who does not defend and who is attacked suffices for the entire value of the public good to be lost.

We assume that each player attaches weight α to the value of the public good and weight $(1-\alpha)$ to the value of the private good. Cooperating comes at a cost c . The value x_iG which player i obtains from the defense of his own private good, creates a by-product benefit $e x_iG$ to the other player, independent of whether the other player also defends his own private good, where e , with $0 \leq e \leq 1$, is a measure of the strength of the by-product benefit. A nonzero e justifies using the term cooperating even when the players only put value on the private good. The following values are now obtained for R , S , T and P , and are further explained below:

$$R = \alpha V + (1-\alpha)G(1+e) - c \quad (13)$$

$$S = \alpha V \{1/(2^A) + [1-1/(2^A)](1-k)\} + (1-\alpha)G[1+e/(2^A)] - c \quad (14)$$

$$T = \alpha V \{1/(2^A) + [1-1/(2^A)](1-k)\} + (1-\alpha)G[1/(2^A) + e] \quad (15)$$

$$P = \alpha V(1-k)/(2^{A-1}) + (1-\alpha)G(1+e)/(2^A) \quad (16)$$

When both players cooperate (reward payoff R), the individual player obtains the full value of the public good and of his private good (plus the by-product benefit from the other player's private good). The sucker payoff S is constructed as follows. With probability $1/(2^A)$, only the sucker is ever attacked, and the sucker obtains the full value from the public good and from his own private good (as an attack on a cooperating player is never successful), and obtains eG from the value of the other player's private good, because the other player is never attacked (so that it does not matter that the other player defects). With complementary probability $[1-1/(2^A)]$, the other (defecting) player is attacked at least once, so that the value of the public good is reduced to $(1-k)V$, and so that additionally the sucker does not obtain any by-product benefit from the other player's private good; however, the sucker continues to obtain the value of his own private good.

The public good part of the temptation payoff T obtained by a unilateral defector is by definition identical to the public good part of the sucker payoff, as players always obtain the same value from the public good. The defector always obtains the by-product benefit from the private good of the other player, simply because the other player cooperates. With probability $1/(2^A)$, only the other (cooperating) player is ever attacked, and the defector obtains the full value of his own private good as well; with the complementary probability the defector is attacked at least once, and he obtains a zero value from his own private good.

Finally, to see how the punishment payoff P is constructed, note first that with a single attack ($A=1$), as no player cooperates, value $(1-k)V$ is automatically obtained from the public good. If two attacks take place ($A=2$), with probability 0.25 Player 1 is attacked twice, and with probability 0.25 Player 2 is attacked twice, so that with probability 0.5 value $(1-k)V$ is obtained from the public good; with the complementary probability 0.5 two different players are attacked, and the value obtained from the public good is 0 . For general $A \geq 1$, the probability that either only one, or only the other player is ever attacked, and that $(1-k)V$ is obtained, equals $1/(2^{A-1})$. To find the value obtained from the private good by the individual player, note that with probability $1/(2^A)$ only the other player is ever attacked, so that the individual player obtains

¹¹ By-product mutualism is generally defined as "ordinarily selfish behavior incidentally benefiting neighbors" (West Eberhard, 1975, p. 19). It thus in general involves a selfish act, which creates a by-product benefit. The production of a public good also fits into this. If, given the behavior of the other players, the individual player contributes to the public good, he does this because it is in his own interest, and the fact that the public good also generates value for other players, arises as a by-product. The distinguishing feature of the public good is that in any circumstance, the players obtain the same value of the public good.

¹² While the benefit from the public good is non-excludable, in comparison, the benefit Player 1 obtains from producing his own private good need not be the same as the by-product benefit Player 2 obtains from the private good produced by Player 1. E.g., when an animal calls a conspecific to a newly found food source to decrease predation risk, the benefit to the calling animal (who was first at the food source) may be larger than the benefit to the conspecifics.

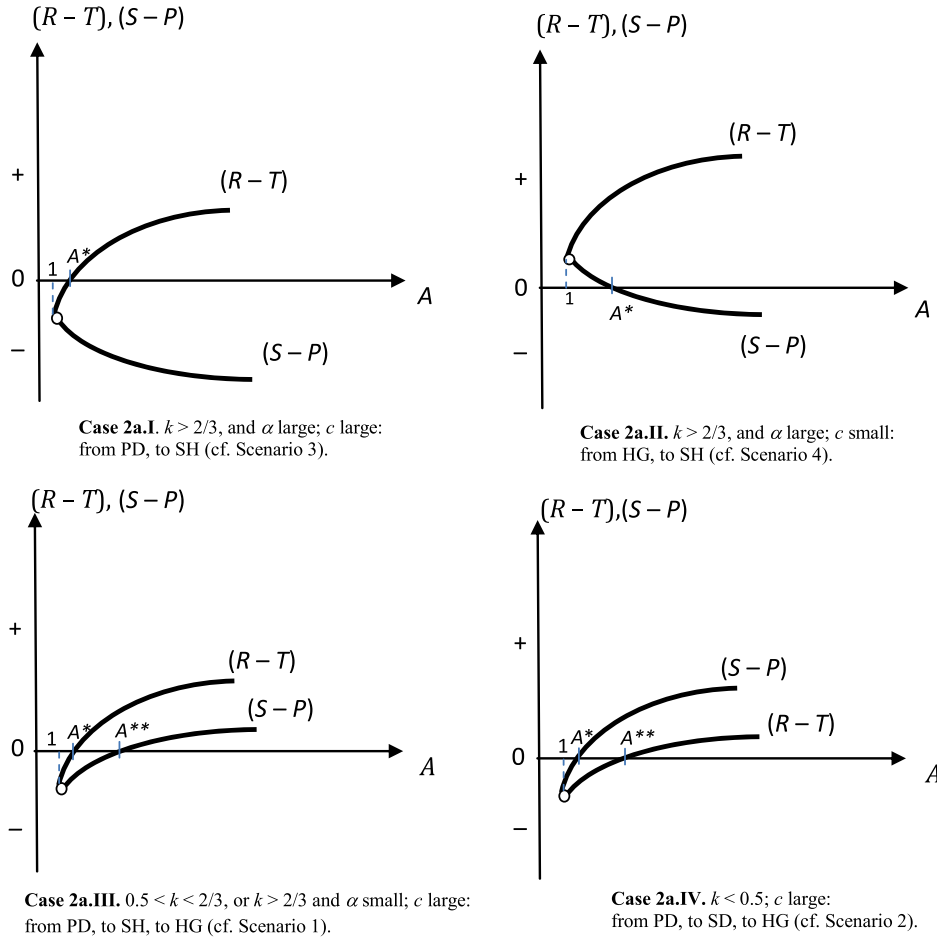


Fig. 4. Model 2a. Effect of the degree of adversity (δ) measured as the number of attacks ($\delta=A$), on added payoff of joint/lone cooperation (respectively $(R-T)$, $(S-P)$); depends on weight given to the public good (α), and on degree of complementarity (k).

payoff G from his own private good, but no by-product benefit from the other player's private good. With the same probability $1/(2^A)$, only the individual player himself is ever attacked, and does not obtain any value from his own private good, but obtains by-product benefit eG from the other player's private good. With the remaining probability, both players are attacked, and no values or by-product benefits are obtained from the private goods.

Using payoffs (13)–(16), we obtain the following values for the added payoff of cooperating jointly and the added payoff of cooperating alone:

$$(R-T) = [1 - 1/(2^A)][k\alpha V + (1-\alpha)G] - c \quad (17)$$

$$(S-P) = \alpha V\{(2k-1)/(2^A) + [1 - 1/(2^A)](1-k)\} + [1 - 1/(2^A)](1-\alpha)G - c \quad (18)$$

Note that these added payoffs do not depend on e (=the degree to which one player's private good generates a by-product benefit to the other player), simply because the individual player does not internalize the by-product benefit he produces for the other player.

The difference between the added payoff of cooperating jointly and the added payoff of cooperating alone, can be calculated to equal

$$(R-T) - (S-P) = [1 - 2/(2^A)](2k-1)\alpha V. \quad (19)$$

This difference does not depend on the value of the private good, as from the perspective purely of the private good, there is no difference between the added payoff of cooperating jointly or alone. Finally, using (2) and (13)–(16), the probability of joint

defection being played rather than joint cooperation in the Stag Hunt equals

$$\begin{aligned} p^* &= (P-S)/(R-T+P-S) \\ &= [c - \alpha V\{(2k-1)/(2^A) + [1 - 1/(2^A)](1-k)\} \\ &\quad - [1 - 1/(2^A)](1-\alpha)G] / \{[1 - 2/(2^A)](2k-1)\alpha V\}. \end{aligned} \quad (20)$$

Also, using (3) and (13)–(16), the proportion of cooperating players in the mixed equilibrium of the Snowdrift game equals

$$\begin{aligned} p &= (S-P)/(T-R+S-P) = \{ \alpha V\{(2k-1)/(2^A) + [1 - 1/(2^A)](1-k)\} \\ &\quad + [1 - 1/(2^A)](1-\alpha)G - c \} / \{ [1 - 2/(2^A)](1-2k)\alpha V \}. \end{aligned} \quad (21)$$

4.1. Model 2a: effect of the number of random attacks

We here measure the degree of adversity as the number of random attacks the two players face ($\delta=A$). Again, we provide intuitions for the different cases we obtain by means of potential examples, and derive them formally in the Appendix.

Case 2a.I (cf. Scenario 3). Consider the potential example of defense of a common territory by male lions (Grinnell et al., 1995; more generally, see Port et al. (2011)), in particular two lions (for other potential examples of territorial defense see, e.g., Gese (2001) on coyotes (*Canis latrans*), and Rubenstein and Nunez (2009) on equids). Let each lion be positioned on either side of the common territory, and let cooperation by the individual lion mean

being attentive to potential intruders at his own side of the territory, and let defection mean not being attentive. Intruders randomly pick one of the two sides of the territory to attack, where such a random attack may take place any number of times. While [Case 2a.I](#) is more general, the intuition is best understood for the extreme case where, when a lion's side of the territory is attacked and he is not attentive, the common territory is completely lost (=weakest-link technology).

Let it be the case that the cost of being attentive is high (e.g., because being attentive to intruders involves chasing them, and therefore causes a risk of being hurt). Then, if the lions face only *one* attack, both the free-rider effect and the sucker effect apply, and no lion wants to cooperate, whatever the other lion does ($(R-T) < 0$, $(S-P) < 0$). The relatively low probability of the intruder attacking one's side does not justify the cost of cooperating. We therefore have a Prisoner's Dilemma. At the other extreme, let the lions face many attacks. If both lions are initially attentive, and one lion defects, then it is very likely that he is attacked at least once, and that with a weakest-link technology the common territory is completely lost. The boomerang effect applies here ($(R-T) > 0$). If neither lion is initially attentive, and one lion decides to be attentive, then with many attacks it is very unlikely that only the attentive lion is attacked, and with a weakest-link technology very likely that the common territory is completely lost. Therefore, there continues to be a sucker effect ($(S-P) < 0$). It follows that as the number of attacks is increased, the free-rider effect turns into a boomerang effect, while the sucker effect is maintained, so that the game becomes a Stag Hunt (see [Fig. 4](#), [Case 2a.I](#)). Measuring the degree of adversity by the number of attacks, a higher degree of adversity therefore increases the probability of joint cooperation, and we obtain a case in line with [Scenario 3](#) in [Fig. 1](#).

It is clear that [Case 2a.I](#) applies more generally in less extreme cases, where the degree of complementarity between defensive inputs is high. In such cases, a higher number of attacks has little impact on the payoffs obtained when no player defends (punishment payoff P), but decreases the payoff of a player who defends unilaterally (sucker payoff S), so that the added payoff of cooperating alone ($S-P$) decreases in the number of attacks.

Case 2a.II (cf. Scenario 4). Consider exactly the same case as in [Case 2a.I](#), except that the cost of being attentive is now low (assume, e.g., that being attentive simply involves roaring at intruders). Then as long as this cost is sufficiently low, the boomerang effect will apply even if the lions face one attack ($(R-T) > 0$). Moreover, because of the low cost, also the hero effect will then apply ($(S-P) > 0$). It follows that the lions play a Harmony Game in case of a low number of attacks. Note now that as the number of attacks is increased, it becomes more and more likely that a lion who unilaterally deviates from joint cooperation is attacked. Therefore, a higher number of attacks makes the boomerang effect even stronger. Yet, at the same time, as the number of attacks is increased, it will be less and less likely that a lion who unilaterally deviates from joint defection will be the only one who is attacked. The payoff of a lion who is the only one to defend (sucker payoff S) decreases in the number of attacks, while the payoff obtained when no lion defends (punishment payoff P) under a sufficiently high degree of complementarity between defensive inputs changes little. The added payoff of cooperating alone ($S-P$) therefore decreases in the number of attacks, from which it follows that with a larger number of attacks, the hero effect turns into a sucker effect ($(S-P) < 0$). Therefore, as the number of attacks is increased, the game changes from a Harmony Game into a Stag Hunt (see [Fig. 4](#), [Case 2a.II](#)). Considering the number of attacks as the degree of adversity, this means that a

higher degree of adversity makes joint cooperation less likely, and [Case 2a.II](#) is in line with [Scenario 4](#) in [Fig. 1](#).

[Cases 2a.I and 2a.II](#) are qualitatively similar to [Cases 1a.I and 1a.II](#) in Model 1 (see columns for [Scenarios 3 and 4](#) in [Table 3](#)). Intuitively, when the degree of complementarity between defensive inputs is high, an increase in the number of attacks in Model 2 has a similar effect to an increase in the degree of complementarity in Model 1. This is because a higher number of attacks makes each player's defensive efforts more complementary, as the individual player is more likely to be attacked.

Case 2a.III (cf. Scenario 1). For a potential example of this case, consider fiddler crabs (*Uca mjoebergi*) who each hold private, neighboring territories ([Detto et al., 2010](#)). Cooperating means assisting in defending if a neighboring crab's territory is attacked by an intruder. The private benefit obtained from this for the cooperating crab is to avoid having to renegotiate territorial boundaries once the intruder becomes the new neighbor (the so-called dear enemy effect). Assisting clearly creates a by-product benefit for the neighboring crab that is assisted.¹³ If cooperating only generates private benefits and by-product benefits, a crab's benefit of assisting a neighboring crab does not depend in any way on whether the neighboring crab also assists ($(R-T) = (S-P)$). For a high cost of assisting, the crabs play a Prisoner's Dilemma ($(R-T) = (S-P) < 0$). As the number of attacks is increased, a switch occurs directly from a Prisoner's Dilemma to a Harmony Game ($(R-T) = (S-P) > 0$). Other potential examples may be house sparrows (*Passer domesticus*) producing chirrup calls to attract other sparrows to a divisible food resource, because this makes it easier for the sparrow to defend the food resource against larger birds ([Elgar, 1986](#)); or, feeding of others' offspring by birds living in groups, in order to stop others' offspring from begging and thereby attracting predators ([Caraco and Brown, 1986](#)).

These examples at first sight exclusively involve private goods, so that contrary to what is the case in [Fig. 4](#) for [Case 2a.III](#), the curves for the added payoff of cooperating jointly and the added payoff of cooperating alone coincide. Yet, as soon as the defensive inputs also have a public good aspect, and as soon as a second defending player contributes more than a first defending player to the public good ($(R-T) > (S-P)$), the picture in [Fig. 4](#) for [Case 2a.III](#) is obtained. For instance, in the example of feeding others' young, there may be benefits of maintaining a group that have a public good aspect, and feeding may not exclusively have private good aspects. Measuring the degree of adversity as the number of attacks ($\delta=A$), [Case 2a.III](#) is in line with [Scenario 1](#) in [Fig. 1](#).

[Case 2a.III](#) may also occur when players exclusively value the defense of a public good. The case is then obtained under two conditions. First, the contribution of a second defending player exceeds the contribution of the first defending player ($(R-T) > (S-P)$), i.e., the degree of complementarity between the players' defensive inputs is sufficiently high. Second, the degree of complementarity between players' defensive inputs is at the same time low enough for the added payoff of cooperating alone to increase with a higher number of attacks. Intuitively, we know that with a summation technology, the added payoffs of cooperating jointly and alone are equal, and both increase in the number of attacks. With a weakest-link technology, we know that the added payoff of cooperating alone decreases in the number of attacks: the punishment payoff P does not change, but the sucker payoff S decreases. Therefore, for technologies between the summation

¹³ Note that as the individual crab's individual benefit consists of maintaining the same neighbor, an attack of the individual crab's private good, means that its neighbor is attacked, and an attack of the other player's private good means that it is attacked itself.

technology and the weakest-link technology, but closer to the summation technology, the added payoff of cooperating alone ($S-P$) increases in the number of attacks.

Case 2a.IV (cf. Scenario 2). As a potential example of this case, consider sentinel behavior in groups of meerkats (*Suricata suricatta*) (Clutton-Brock et al., 1999; more generally, see Bednekoff (2001)). Cooperating means acting as a sentinel (=being attentive to predators, and giving an alarm call if a predator is present), defecting means not acting as a sentinel. Keeping the group safe generates a public benefit. At the same time, acting as a sentinel may also yield a private benefit, as the sentinel is then the first who can escape. Consider the extreme case where one sentinel suffices to keep the entire group safe. Assume now that acting as a sentinel is relatively costly. Then, when the meerkats face only one attack, both a free-rider effect and a sucker effect apply ($(R-T) < 0$, $(S-P) < 0$), and the game played is a Prisoner's Dilemma. As the number of attacks is increased, given that one sentinel suffices to keep the group safe, the sucker effect will turn into a hero effect ($(S-P) > 0$), but the free-rider effect will be maintained (the same applies more generally if the degree of complementarity between meerkats' defensive efforts in the form of sentinel behavior is low). The game therefore turns into a Snowdrift game (see Fig. 4, Case 2a.IV). If the private benefits of acting as a sentinel are sufficiently high, as the number of predators is further increased, the free-rider effect may eventually turn into boomerang effect ($(R-T) > 0$), while the hero effect is maintained, so that a Harmony Game is played. Measuring the degree of adversity by the number of attacks ($\delta=A$), Case 2a.IV is in line with Scenario 2.

4.2. Model 2b: effect of the values of the public and the private good

In Model 2b, we look again at the effect of increases in the value of the public good ($\delta=V$), where this value is this time interpreted as the value of the common resource that the players defend (one may here equally well look at the effect of an increase in the value of the private good, which has a similar effect). In spite of the different setting, it is easy to see that the analysis is completely analogous to the one in Model 1b, and that Fig. 3 directly applies to this case as well. Both the added payoff of a first defending player, and the added payoff of a second defending player, increase in the value of the public good. We here provide the intuitions, and derive the cases formally in the Appendix.

Case 2b.I (cf. Scenario 1). In this case, the degree of complementarity between the players' defensive inputs is fixed at a high level ($k > 0.5$), so that the added payoff of a second defending player is larger than the added payoff of a first defending player ($(R-T) > (S-P)$). As the value of the public good is increased, we move from a Prisoner's Dilemma, to a Stag Hunt, and finally to a Harmony Game (see left part Fig. 3). It follows that Case 2a.I is fully in line with Scenario 1 in Fig. 1.

Case 2b.II (cf. Scenario 2). In this case, the degree of complementarity between the players' defensive inputs is instead fixed at a low level ($k < 0.5$), and the added payoff of a second defending player is smaller than the added payoff of a first defending player ($(R-T) < (S-P)$). We now move from a Prisoner's Dilemma, to a Snowdrift game, and finally to a Harmony Game as the value of the public good is increased (see right part Fig. 3). It follows that Case 2b.II is a variant of Scenario 2.

We finally compare for Model 2 the effect of an increase in the value of the public good and/or private good, to the effect of an increase in the number of attacks. For sufficiently high weight on the public good and for a high degree of complementarity between defensive inputs, the comparison between Cases 2a.I/2a.II

and 2b.I/2b.II again reveals that the effect of the number of attacks differs from a simple change in the cost-benefit ratio of cooperating. Yet, for a low degree of complementarity and/or high weight on the private good, the effect of a higher number of attacks and a simple decrease in the cost-benefit ratio of cooperating is similar (see columns for Scenarios 1 and 2 in Table 3). For a private good, this is straightforward; for a public good, it is intuitive as, e.g., with a summation technology, the number of attacks affects the added payoff of cooperating no matter how many other players cooperate.

We now compare our micro-foundations model to two apparently related models of cooperation, namely one where synergies between players' efforts induce cooperation, and one where interdependence between the players' benefits induces cooperation, to clarify the differences with our model.

5. Comparison to models of synergy/discounting and of interdependence

Consider first the model of synergy/discounting of cooperation by Hauert et al. (2006), which we apply here to the simple two-player setting.¹⁴ We again denote the production function of the public good as $f(x_1, x_2)V$, where x_1 and x_2 can take on values 0 or 1 reflecting respectively that a player does not or does cooperate. In a model of synergy/discounting, $f(1, 1)=V(1+w)/2$, $f(0, 0)=0$, and $f(1, 0)=f(0, 1)=V/2$, where w is the degree of synergizing. For $w < 1$ (discounting), just as in our case of a low degree of complementarity, the second player contributes less than the first; for $w > 1$ (synergy), just as in our case of a high degree of complementarity, the second player contributes more than the first. Yet, in the model of Hauert et al. starting from joint defection, a first cooperating player always contributes $V/2$ to the public good ($(S-P)=(V/2)-c$), whereas a second cooperating player contributes $wV/2$ to the public good ($(R-T)=(wV/2)-c$). This reflects a crucial difference with our model: the degree of synergy does not affect the value of the public good when only one player cooperates, but increases the value when both cooperate. In our model, a higher the degree of complementarity instead negatively affects the value of the public good when only one player cooperates, but does not change the value when both cooperate. Our paper focuses on the degree of complementarity rather than on synergy/discounting, because of our interest in the common-enemy hypothesis, broadly saying that a harsh environment makes players cooperate more because they depend more on each other's cooperative efforts. A higher degree of complementarity can be linked to a more adverse environment, because in case only one player cooperates, it reduces the payoff of players – whereas the opposite is true for a higher degree of synergy.¹⁵

It is instructive to consider the model of synergy/discounting of cooperation within our framework. In this model, the added payoff of cooperating jointly increases as we increase w to move from discounting to synergy, whereas the added payoff of cooperating alone is unaffected. This leads to a variant on Fig. 2, with the difference that the added payoff of cooperating alone is flat. For $c > (V/2)$, as we increase w , we move from a Prisoner's Dilemma to a Stag Hunt, where it can be checked that in the Stag Hunt the basin of attraction of joint cooperation increases in w . The case is similar to Fig. 2(a), except that the game never becomes a Snowdrift game

¹⁴ A wider perspective on synergism is recently found in Corning and Szathmáry (2015).

¹⁵ As synergy can induce by-product mutualism (or, *synergistic mutualism* (Maynard Smith, 1983)), and as a higher degree of synergy cannot be considered as a harsher environment, the synergy/discounting model further illustrates that an adverse environment is not a necessary condition for by-product mutualism.

for low w . For $c < (V/2)$, as we increase w , we move from a Snowdrift game to a Harmony Game, where it can be checked that in the Snowdrift game the probability of cooperating in the mixed equilibrium increases in w . The case is similar to Fig. 2(b), except that the game never becomes a Stag Hunt for high w . Thus, the essential difference between the effect of the degree of synergy, and the effect of the degree of complementarity is that the former leaves the added payoff of cooperating alone unchanged, whereas the latter decreases the added payoff of cooperating alone. The consequence is that while a higher degree of synergy can never negatively affect the probability of cooperation, the effect of a higher degree of complementarity is ambiguous.

Second, consider the model of interdependence by Roberts (2005), who points out the links of this model to the common-enemy hypothesis.¹⁶ When a player cooperates, he provides a benefit b to the other player, at a cost c and because of interdependence with the other player obtains a benefit sb from providing this benefit. In this case, $(R - T) = (S - P) = sb - c$.¹⁷ An increase of s does not fit the manner in which we have defined adversity, as for any number of cooperating players, a higher degree of interdependence makes players better off. Still, the effect of interdependence can be considered in the framework of the scenarios considered in Section 2. In particular, the effect of interdependence s in the model of interdependence is in line with the case exactly in between Scenarios 1 and 2 (the contributions to the public good are additive, or only the private good receives weight). Increased interdependence increases the probability of joint cooperation, as we move from a Prisoner's Dilemma to a Harmony Game. Thus, an increase in the degree of interdependence as modeled by Roberts (2005), unambiguously increases the probability of cooperation.

6. Conclusions

The argument that harsh environments serve as a common enemy encouraging cooperation can be dated back to at least Kropotkin (1902), and appeals to our common sense. As formulated by Mesterton-Gibbons (2001, p. 211): "How many times have you heard it said that the only way to accomplish wholesale cooperation among humans would be to invite aliens from outer space to invade our planet?" Several papers have in diverse contexts confirmed the hypothesis that adversity favors cooperation (e.g., Callaway et al., 2002; Spieler, 2003; Krams, Berzins et al., 2010; Krams, Krama et al. 2010; Sandoval and Wilson, 2012). Yet, the number of such papers is still not sufficiently large to claim that the hypothesis is systematically confirmed. Moreover, comparative studies have not yet identified ecological constraints that

differ systematically between cooperative and non-cooperative species (Hatchwell and Komdeur, 2000).

Appealing to similar common sense as Mesterton-Gibbons, social scientists hypothesize that a repressive government comes to be seen as a common enemy, and induces collective action against this government by activists. Yet, interestingly, social science includes a competing hypothesis that repression instead stifles such collective action (see Earl (2011), for an overview), which at an intuitive level seems equally plausible. The question then is not so much whether either the common-enemy hypothesis or its competing hypothesis applies; the question is rather in which circumstances each hypothesis applies.¹⁸

The basic insights from our model are the following. Let cooperating consist of contributing to the production of a public good, and let adversity make players' cooperative efforts in producing the public good more complementary. Then on the one hand, such an increase in the degree of complementarity gives the players fewer incentives to unilaterally defect in a situation of joint cooperation (boomerang effect). Yet, on the other hand, it also gives players fewer incentives to unilaterally cooperate in a situation of joint defection (sucker effect). When the cost of cooperation is sufficiently high such that for a low degree of complementarity, it is a dominant strategy for players to defect, because of the boomerang effect, an increase in the degree of complementarity increases the probability that players cooperate. This is in line with the common-enemy hypothesis. When, on the contrary, the cost of cooperation is sufficiently high such that for a low degree of complementarity, it is a dominant strategy for players to cooperate, because of the sucker effect, an increase in the degree of complementarity decreases the probability that players cooperate. This is in line with the competing hypothesis to the common-enemy hypothesis.

Next, let cooperating take the form of defending a public good, and/or a private good with by-product benefits, and let adversity consist of random attacks on the players. Then from the perspective of the public good, as long as the degree of complementarity between players' defensive inputs is fixed at a sufficiently high level, a larger number of attacks has the same effect as an increase in the degree of complementarity in the production of a public good, and the circumstances in which the common-hypothesis does/does not apply are analogous to the case where the public good is produced rather than defended. Yet, if the degree of complementarity is instead fixed at a low level, while the boomerang effect continues to apply, a hero effect exists rather than a sucker effect, meaning that a larger number of attacks gives the individual player more incentives to unilaterally cooperate in a situation of joint defection. The consequence is that the common-enemy hypothesis unambiguously applies in this case. The same is true from the perspective of the private good, meaning that as long as the private benefits of defensive inputs receive sufficient weight compared to the public benefits, the common-enemy hypothesis unambiguously applies.

We have studied these effects in the simplest possible setting, namely the setting of two players, with discrete rather than continuous efforts. The basic insights from this paper should equally operate in a multi-player setting with continuous efforts, but studying the multi-player case may lead to additional insights and deserves separate attention.

¹⁶ Citing Roberts (2005, p. 905): "Mesterton-Gibbons and Dugatkin (1997) consider what they call by-product mutualism and what Connor calls pseudo-reciprocity as resulting from the common enemy of an adverse environment. This argument is similar to that of interdependence, but the stakeholder framework adds an explicit feedback mechanism which shows how any given interaction may fit a range of payoff matrices according to the magnitude of the stake."

¹⁷ Connor (1986) defines the concept of *pseudo-reciprocity*: at a cost c , agent 1 provides a benefit $f(c)$ to agent 2, because of the by-product benefit that this gives to agent 1. Considering the by-product benefit as $sf(c)$, with s the degree of interdependence, this bears close resemblance to the model of interdependence of Roberts (2005) (in Brown and Vincent (2008), there is additionally cost sharing between the players). Framing the example differently, it resembles the production of a private good (see Footnote 8): one could say that agent 1 "produces" private good $sf(c) = G$, and that this yields a by-product benefit $f(c) = eG$ to agent 2. While Leimar and Connor, (2003) and Leimar and Hammerstein (2010) argue that pseudo-reciprocity is a different type of cooperation, based on the mentioned analogies, we follow Mesterton-Gibbons and Dugatkin (1997) in only using the term by-product mutualism.

¹⁸ As one referee points out, it is doubtful that even evolutionary theorists consider the common-enemy hypothesis as the norm.

Acknowledgments

We would like to thank four anonymous referees and the editor for helpful comments.

Appendix

Model 1a: formal analysis of the effect of the degree of complementarity.

Looking at the extreme cases, by (8) and (9), for $k=0$, $(S-P)=V-c$, and for $k=1$, $(R-T)=V-c$ as well. Also, for $k=0$, $(S-P)=-c$, and for $k=1$, $(R-T)=-c$ as well.

By (8)–(10), for $k=0.5$, we have $(S-P)=(R-T)=0.5V-c$. For $c > 0.5V$ (“large cooperation costs”), $(S-P)=(R-T) < 0$ for $k=0.5$, whereas for $c < 0.5V$ (“small cooperation costs”), $(S-P)=(R-T) > 0$ for $k=0.5$. More generally, it is easy to check from (8) and (9) that $(S-P)$ decreases linearly in k , whereas $(R-T)$ increases linearly in k . It follows that for $c > 0.5V$ (Fig. 2(a)), two critical degrees of complementarity k^* and k^{**} exist, with $k^* < k^{**}$, such that for $k < k^*$ the game is a Snowdrift game, for $k^* < k < k^{**}$ it is a Prisoner's Dilemma, and for $k > k^{**}$ it is a Stag Hunt. For $c < 0.5V$ (Fig. 2(b)), two critical degrees of complementarity k^* and k^{**} exist, with $k^* < k^{**}$, such that for $k^* < k < k^{**}$ the game is a Harmony Game, with the other cases identical to those for large costs of cooperation.

In Case 1a.I, the switch from the Prisoner's Dilemma to the Stag Hunt as k is increased suggests that a higher degree of complementarity promotes cooperation. This is confirmed by looking at the basin of attraction of joint cooperation in the Stag Hunt. Using (11), the partial derivative of the basin of attraction of joint defection with respect to k equals:

$$\begin{aligned} \partial p^* / \partial k &= \{V(2k-1)V - [c-(1-k)V]2V\} / [(2k-1)V]^2 \\ &= V\{V-2c\} / [(2k-1)V]^2 \end{aligned} \quad (\text{A.1})$$

This expression is negative for $c > 0.5V$, showing that for $k > k^{**}$, the basin of attraction of joint cooperation increases in the degree of complementarity.

In Case 1a.II, the switch from the Harmony Game to the Stag Hunt as k is increased suggests that a higher degree of complementarity in this case instead discourages cooperation. This is confirmed by looking at the basin of attraction of joint cooperation in the Stag Hunt. As is clear from (A.1), the partial derivative of the basin of attraction of joint defection with respect to k is positive for $c < 0.5V$. This shows that for $k > k^{**}$, the basin of attraction of joint cooperation decreases in the degree of complementarity.

In Case 1a.III, the switch from the Snowdrift game to the Prisoner's Dilemma as k is increased suggests that a higher degree of complementarity discourages cooperation. This is confirmed by looking at the proportion of cooperating players in the mixed equilibrium of the Snowdrift game. Using (12), the partial derivative of this proportion with respect to k equals:

$$\begin{aligned} \partial p / \partial k &= \{-V[(1-2k)V] + 2V[(1-k)V-c]\} / [(1-2k)V]^2 \\ &= \{V[V-2c]\} / [(1-2k)V]^2 \end{aligned} \quad (\text{A.2})$$

This expression is negative for $c > 0.5V$, showing that for $k > k^*$, the proportion of cooperating players in the mixed equilibrium of the Snowdrift game decreases in the degree of complementarity.

In Case 1a.IV, the switch from the Snowdrift game to the Harmony Game as k is increased on the contrary suggests that a higher degree of complementarity promotes cooperation. This is confirmed by looking at the proportion of cooperating players in the mixed equilibrium of the Snowdrift game. Indeed, (A.2) is positive for $c < 0.5V$, showing that for $k > k^*$, the proportion of

cooperating players in the mixed equilibrium of the Snowdrift game decreases in the degree of complementarity.

Model 1b: formal analysis of the effect of the value of the public and private good.

As is easy to check from (8) and (9), both $(S-P)$ and $(R-T)$ always increase linearly in V . As can be seen from (10), for $k > 0.5$ (Case 1b.I, Fig. 3(a)), it is the case for any positive V that $(R-T) > (S-P)$. It follows that for this case, critical values of the public good V^* and V^{**} (with $V^* < V^{**}$) exist such that the following is true. For $V < V^*$, the game played is a Prisoner's Dilemma; for $V^* < V < V^{**}$, it is a Stag Hunt; for $V > V^{**}$, it is a Harmony Game. In the Stag Hunt, by (11), the partial derivative of p^* with respect to V equals:

$$\partial p^* / \partial V = \{-(2k-1)c\} / [(2k-1)V]^2. \quad (\text{A.3})$$

Given that $(2k-1) > 0$ in Case 1b.I, it follows that $\partial p^* / \partial V < 0$, meaning that the basin of attraction of joint cooperation increases in the value of the public good.

As can be seen from (10), for $k < 0.5$ (Case 1b.II in Fig. 3(b)), it is the case for any positive V that $(R-T) < (S-P)$. Thus, for this case, critical values of the public good V^* and V^{**} (with $V^* < V^{**}$) exist such that the following is true. For $V < V^*$, the game played is a Prisoner's Dilemma; for $V^* < V < V^{**}$, it is a Snowdrift game; for $V > V^{**}$, it is a Harmony Game. In the Snowdrift game, by (12), the partial derivative of p with respect to V equals:

$$\partial p / \partial V = \{-(1-2k)[-c]\} / [(1-2k)V]^2 \quad (\text{A.4})$$

Given that $(1-2k) > 0$ in Case 1b.II, this is larger than zero, it follows that $\partial p / \partial V > 0$, meaning that the proportion of cooperating players in the mixed equilibrium of the Snowdrift game increases as V is increased.

Model 2a: formal analysis of the effect of the number of random attacks.

We now show formally that the four cases described in 2a.I–2a.IV exist. It is clear from (17) and (18) that $(S-P)=(R-T)$ for $A=1$. Furthermore, it is clear from (17) that $(R-T)$ always increases in A . The partial derivative of $(S-P)$ with respect to A equals

$$\partial (S-P) / \partial A = \alpha V(3k-2) / (2^A) \ln(1/2) - 1 / (2^A) (1-\alpha) G \ln(1/2). \quad (\text{A.5})$$

It follows from (A.5) that $\partial (S-P) / \partial A < 0$ only if $k > 2/3$; furthermore α needs to be sufficiently large. With $\partial (S-P) / \partial A < 0$, the type of game may change whether $(S-P)$ and $(R-T)$ are positive or negative for $A=1$. In order for the type of game to change, we assume that $c > \alpha V(1-k) + (1-\alpha)G$, so that $(S-P)$ is negative for A approaching infinity. For $c > 0.5[k\alpha V + (1-\alpha)G]$, we obtain Case 2a.I in Fig. 4. For $c < 0.5[k\alpha V + (1-\alpha)G]$, we obtain Case 2a.II in Fig. 4. For Case 2a.I, given that $(S-P)=(R-T) < 0$ for $A=1$, given that $(S-P) < (R-T)$ for larger A , and given that $(S-P)$ remains negative, there exist some A^* such that for $1 \leq A < A^*$, the game is a Prisoner's Dilemma; for $A > A^*$ the game is a Stag Hunt. For Case 2a.II, given that $(S-P)=(R-T) > 0$ for $A=1$, given that $(S-P) < (R-T)$ for larger A and given that $(S-P)$ eventually becomes negative, there exists some A^* such that for $1 \leq A < A^*$, the game is a Harmony Game; for $A > A^*$ the game is a Stag Hunt. Using (20), we can calculate that in the Stag Hunt:

$$\partial p^* / \partial A = \{ \ln(2)(2k-1)\alpha V[1/(2^A)] [\alpha V k + (1-\alpha)G - 2c] \} / \{ [1-2/(2^A)](2k-1)\alpha V \}^2. \quad (\text{A.6})$$

Given that in Cases 2a.I and 2a.II, it is the case that $k > 2/3$, it follows that $\partial p^* / \partial A < 0$ when $c > 0.5[k\alpha V + (1-\alpha)G]$, and $\partial p^* / \partial A > 0$ when $c < 0.5[k\alpha V + (1-\alpha)G]$. Therefore, the basin of attraction of joint cooperation in the Stag Hunt increases in Case 2a.I, and decreases in Case 2a.II.

From (A.5), it follows that $\partial(S-P)/\partial A > 0$ whenever $k < 2/3$. In this case, when it is additionally valid that $0.5 < k < 2/3$, we have $(S-P) < (R-T)$, and Case 2a.III in Fig. 4 is obtained. When instead $k < 0.5$, by (19), we have $(S-P) > (R-T)$, and Case 2a.IV in Fig. 4 is obtained. When $k > 2/3$, it is only the case that $\partial(S-P)/\partial A > 0$ when α is sufficiently small. As it is necessarily true now that $(S-P) < (R-T)$, this is another instance of Case 2a.III. As our interest is in non-trivial cases where the type of game played changes as A is increased, we focus in Cases 2a.III and 2a.IV on costs of cooperation $c > 0.5[k\alpha V + (1-\alpha)G]$, such that both $(S-P)$ and $(R-T)$ are negative for $A=1$. We furthermore assume for these cases that $c < k\alpha V + (1-\alpha)G$, and $c < \alpha V(1-k) + (1-\alpha)G$, so that as A approaches infinity, both $(S-P)$ and $(R-T)$ are positive.

It follows that for Case 2a.III, there exist levels A^* and A^{**} , with $A^* < A^{**}$, such that for $A < A^*$, the game is a Prisoner's Dilemma; for $A^* < A < A^{**}$, it is a Stag Hunt; and for $A > A^{**}$, it is a Harmony Game. For Case 2a.IV, there exist levels A^* and A^{**} , with $A^* < A^{**}$, where the only difference with Case 2a.III is now that for $A^* < A < A^{**}$, the players play a Snowdrift game. In Case 2a.III, by the fact that $c > 0.5[k\alpha V + (1-\alpha)G]$, and $k > 0.5$, it follows from (A.6) that $\partial p^*/\partial A < 0$, so that in the Stag Hunt the basin of attraction of joint cooperation increases as the number of attacks is increased.

In Case 2a.IV, using (21), we can calculate the effect of A on the proportion of cooperating players in the mixed equilibrium of the Snowdrift game

$$\partial p/\partial A = \ln(2)(1-2k)\alpha V \left[1/(2^A)\right] \left[2c - \alpha V k - (1-\alpha)G\right] / \left\{ \left[1-2/(2^A)\right] (1-2k)\alpha V \right\}^2. \quad (\text{A.7})$$

Given that $k < 1/2$ in the Snowdrift game, and given that $c > 0.5[k\alpha V + (1-\alpha)G]$, it follows that $\partial p/\partial A > 0$. Put otherwise, a larger number of attacks means a higher proportion of cooperating players in the Snowdrift game.

Model 2b: formal analysis of the effect of the value of the public and the private good

We focus on the effect of the value of the public good, as the effect of the value of the private good is analogous. As can be seen from (17) and (18), $(R-T)$ and $(S-P)$ are equal for $A=1$, and are both linearly increasing in A . We focus on the case where $[1-1/(2^A)](1-\alpha)G < c$, so that for low values of the public good the game is a Prisoner's Dilemma, and eventually switches to a Harmony Game for sufficiently high V . By (19), the difference between the two added payoffs increases in A . Finally, it follows from (19) that $(R-T) > (S-P)$ for $k > 0.5$, and $(R-T) < (S-P)$ for $k < 0.5$. We check that in Case 2b.I, when the game is a Stag Hunt, the basin of attraction of joint cooperation becomes larger as the value of the public good is increased. Using (20),

$$\partial p^*/\partial V = \left\{ \left[1-2/(2^A)\right] (2k-1)\alpha \right\} \left\{ \left[1-1/(2^A)\right] (1-\alpha)G - c \right\} / \left\{ \left[1-2/(2^A)\right] (2k-1)\alpha V \right\}^2 \quad (\text{A.8})$$

This expression is smaller than zero given that $k > 0.5$ and $[1-1/(2^A)](1-\alpha)G < c$, so that indeed in the Stag Hunt the basin of attraction of joint cooperation increases as the value of the public good is increased. Using (21),

$$\partial p/\partial V = \left\{ \left[1-2/(2^A)\right] (1-2k)\alpha \right\} \left\{ c - \left[1-1/(2^A)\right] (1-\alpha)G \right\} / \left\{ \left[1-2/(2^A)\right] (1-2k)\alpha V \right\}^2. \quad (\text{A.9})$$

This expression is larger than zero given that $k < 0.5$ and $[1-1/(2^A)](1-\alpha)G < c$, so that in the Snowdrift game the proportion of cooperating players increases as the value of the public good is increased.

References

- Archetti, M., Scheuring, I., Hoffman, M., Frederickson, M.E., Pierce, N.E., Yu, D.W., 2011. Economic game theory for mutualism and cooperation. *Ecol. Lett.* 14, 1300–1312.
- Archetti, M., Scheuring, I., 2012. Review: game theory of public goods in one-shot dilemmas without assortment. *J. Theor. Biol.* 299, 9–20.
- Bednekoff, P.A., 2001. Coordination of safe, selfish sentinels based on mutual benefits. *Ann. Zool. Fenn.* 38, 5–14.
- Bergmüller, R., Johnstone, R.A., Russell, A.F., Bshary, R., 2007. Integrating cooperative breeding into theoretical concepts of cooperation. *Behav. Process.* 76, 61–72.
- Bonanni, R., Valsecchi, P., Natoli, E., 2010. Pattern of individual participation and cheating in conflicts between groups of free-ranging dogs. *Anim. Behav.* 79, 957–968.
- Brown, J.L., 1983. Cooperation: a biologist's dilemma. In: Rosenblatt, J.S. (Ed.), *Advanced in the Study of Behavior*. Academic Press, New York, pp. 1–37.
- Brown, J.L., 1987. *Helping and Communal Breeding in Birds*. Princeton University Press, Princeton, NJ.
- Brown, J.S., Vincent, T.L., 2008. Evolution of cooperation with shared costs and benefits. *Proc. R. Soc. B – Biol. Sci.* 275, 1985–1994.
- Callaway, R.G., Booker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B., Aschehoug, E.T., Armas, C., Kikodze, D., Cook, B.J., 2002. Positive interactions among alpine plants increase with stress. *Nature* 417, 844–848.
- Caraco, T., Brown, J.L., 1986. A game between communal breeders: when is food-sharing stable? *J. Theor. Biol.* 118, 379–393.
- Clutton-Brock, T., 2002. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* 296, 69–72.
- Clutton-Brock, T., 2009. Cooperation between non-kin in animal societies. *Nature* 462, 51–57.
- 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.
- Connor, R.C., 1986. Pseudo-reciprocity: investing in mutualism. *Anim. Behav.* 34, 1562–1584.
- Corning, P.A., Szathmáry, E., 2015. "Synergistic selection": a Darwinian frame for the evolution of complexity. *J. Theor. Biol.* 371, 45–58.
- De Jaegher, K., Hoyer, B. Collective action and the common enemy effect. *Defence Peace Econ.*, forthcoming.
- Detto, T., Jennions, M.D., Backwell, P.R.Y., 2010. When and why do territorial coalitions occur? Experimental evidence from a fiddler crab. *Am. Nat.* 175, E119–E125.
- Dionisio, F., Gordo, I., 2006. The tragedy of the commons, the public goods dilemma, and the meaning of rivalry and excludability in evolutionary biology. *Evol. Ecol. Res.* 8, 321–332.
- Doebeli, M., Hauert, C., 2005. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecol. Lett.* 8, 748–766.
- Dugatkin, L.A., 1997. *Cooperation Among Animals: An Evolutionary Perspective*. Oxford University Press, New York.
- Dugatkin, L.A., 2002a. Cooperation in animals: an evolutionary overview. *Biol. Philos.* 17, 459–476.
- Dugatkin, L.A., 2002b. Animal cooperation among unrelated individuals. *Nat.-wissenschaften* 89, 533–541.
- Earl, J., 2011. Political repression: iron fists, velvet gloves and diffuse control. *Annu. Rev. Sociol.* 37, 261–284.
- Elgar, M.A., 1986. House sparrows establish foraging flocks by giving chirrup calls if the resources are divisible. *Anim. Behav.* 34, 169–174.
- Emlen, S.T., 1982. The evolution of helping. I. An ecological constraints model. *Am. Nat.* 119, 29–39.
- Gese, E.M., 2001. Territorial defense by coyotes (*Canis latrans*) in Yellowstone National Park, Wyoming: who, how, where, when, and why. *Can. J. Zool.* 79, pp. 980–987.
- Grinnell, J., Packer, C., Pusey, A.E., 1995. Cooperation in male lions: kinship, reciprocity or mutualism? *Anim. Behav.* 49, 95–105.
- Hammerstein, P., 2003. Why is reciprocity so rare in social animals? – A protestant appeal. In: Hammerstein, P. (Ed.), *Genetic and Cultural Evolution of Cooperation*. The MIT Press, Cambridge, Ma., pp. 83–93.
- Hamilton, W.D., 1964. The genetical evolution of social behavior I. *J. Theor. Biol.* 7, 1–16.
- Hatchwell, B.J., Komdeur, J., 2000. Ecological constraints, life history traits and the evolution of cooperative breeding. *Anim. Behav.* 59, 1079–1086.
- Hauert, C., Michor, F., Nowak, M.A., Doebeli, M., 2006. Synergy and discounting in social dilemmas. *J. Theor. Biol.* 239, 195–202.
- Hirshleifer, J., 1983. From weakest-link to best-shot: the voluntary provision of public goods. *Public Choice* 41, 371–386.
- Krams, I., Berzins, A., Krama, T., Wheatcroft, D., Igaune, K., Rantala, M.J., 2010. The increased risk of predation enhances cooperation. *Proc. R. Soc. B – Biol. Sci.* 277, 513–518.
- Krams, I., Krama, T., Berzins, A., Rantala, M.J., 2010. The risk of predation favors cooperation among breeding prey. *Commun. Integr. Biol.* 3, 243–244.
- Kropotkin, P., 1902. *Mutual Aid*. Heinemann, London.
- Leimar, O., Connor, R.C., 2003. By-product benefits, reciprocity and pseudo-reciprocity in mutualism. In: Hammerstein, P. (Ed.), *Genetic and Cultural Evolution of Cooperation*. MIT Press, Cambridge, MA, pp. 203–222.

- Leimar, O., Hammerstein, P., 2010. Cooperation for direct fitness benefits. *Philos. Trans. R. Soc. B* 365, 2619–2626.
- Lehmann, L., Keller, L., 2006. The evolution of cooperation and altruism – a general framework and a classification of models. *J. Evol. Biol.* 19, 1365–1376.
- Maynard Smith, J., 1983. Game theory and the evolution of cooperation. In: Bendall, D.S. (Ed.), *Evolution from Molecules to Men*. Cambridge University Press, Cambridge.
- Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18.
- Mesterton-Gibbons, M., 2001. *An Introduction to Game-Theoretic Modelling*, 2nd ed. American Mathematical Society.
- Mesterton-Gibbons, M., Dugatkin, L.A., 1992. Cooperation among unrelated individuals: evolutionary factors. *Q. Rev. Biol.* 67, 267–281.
- Mesterton-Gibbons, M., Dugatkin, L.A., 1997. Cooperation and the prisoner's dilemma: towards testable models of mutualism versus reciprocity. *Anim. Behav.* 54, 551–557.
- Nowak, M.A., 2006. Five rules for the evolution of cooperation. *Science* 314, 1560–1563.
- Nunn, C.L., Lewis, R.J., 2001. Cooperation and collective action in animal behaviour. In: Noë, R., van Hooff, J.A.R.A.M., Hammerstein, P. (Eds.), *Economics in Nature*. Cambridge University Press, Cambridge, pp. 42–66.
- Packer, C., Rutten, L., 1988. The evolution of cooperative hunting. *Am. Nat.* 132, 159–198.
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., Bonhoeffer, S., 2005. Evolution of cooperation by generalized reciprocity. *Proc. R. Soc. B – Biol. Sci.* 272, 1115–1120.
- Port, M., Kappeler, P.M., Johnstone, R.A., 2011. Communal defense of territories and the evolution of sociality. *Am. Nat.* 178, 787–800.
- Ray, D., Baland, J.-M., Dagnelie, O., 2007. Inequality and inefficiency in joint projects. *Econ. J.* 117, 922–935.
- Roberts, G., 2005. Cooperation through interdependence. *Anim. Behav.* 70, 901–908.
- Rubenstein, D.I., Nunez, C.N., 2009. Sociality and reproductive skew in horses and zebras. In: Hager, R., Jones, C.B. (Eds.), *Reproductive Skew in Vertebrates: Proximate and Ultimate Courses*. Cambridge University Press, Cambridge, pp. 196–226.
- Russell, B.W., 1959. *Common Sense and Nuclear Warfare*. George Allen and Unwin, London.
- Sachs, J.L., Mueller, U.G., Wilcox, T.P., Bull, J.J., 2004. The evolution of cooperation. *Q. Rev. Biol.* 79, 135–160.
- Sandoval, L., Wilson, D.R., 2012. Local predation pressure predicts the strength of mobbing responses in tropical birds. *Curr. Zool.* 58, 781–790.
- Scheel, D., Packer, C., 1991. Group hunting behavior of lions: a search for cooperation. *Anim. Behav.* 41, 697–709.
- Sigmund, K., 2010. *The Calculus of Selfishness*. Princeton University Press, Princeton, NJ.
- Skyrms, B., 2004. *The Stag Hunt and Evolution of Social Structure*. Cambridge University Press, Cambridge, MA.
- Smaldino, P.E., Schank, J.C., McElreath, R., 2013. Increased cost of cooperation helps cooperators in the long run. *Am. Nat.* 181, 451–463.
- Spieler, M., 2003. Risk of predation affects aggregation size: a study with tadpoles of *Phrynomantis microps* (Anura: Microhylidae). *Anim. Behav.* 65, 179–184.
- Stander, P., 1992. Cooperative hunting in lions: the role of the individual. *Behav. Ecol. Sociobiol.* 29, 445–454.
- Stevens, J.R., Gilby, I.C., 2004. A conceptual framework for nonkin food sharing: timing and currency of benefits. *Anim. Behav.* 67, 603–614.
- Strassman, J.E., Zhu, Y., Queller, D.C., 2000. Altruism and social cheating in social amoeba. *Dictyostelium discoideum*. *Nature* 408, 965–967.
- Sugden, R., 1986. *The Economics of Rights, Cooperation and Welfare*. Blackwell, Oxford.
- Taylor, P.D., Jonker, L.B., 1978. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* 40, 145–156.
- Trivers, R.L., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Tucker, A., 1950. A two-person dilemma. In: Rasmussen, E. (Ed.), *Reading in Games and Information*. Blackwell, Oxford, pp. 7–8.
- Weibull, J., 1995. *Evolutionary Game Theory*. MIT Press, Cambridge, MA.
- West, S.A., Pen, I., Griffin, A.S., 2002. Cooperation and competition between relatives. *Science* 296, 72–75.
- West Eberhard, M.J., 1975. The evolution of social behavior by kin selection. *Q. Rev. Biol.* 50, 1–35.
- Wilson, D.S., 1980. *The Natural Selection of Populations and Communities*. Benjamin/Cummings, Menlo Park.