J. Math. Biol. (2017) 74:499–529
DOI 10.1007/s00285-016-1017-1

Mathematical Biology

Evolutionary stability in continuous nonlinear public goods games

Chai Molina[1] · **David J. D. Earn**[1]

Received: 3 October 2015 / Revised: 7 February 2016 / Published online: 14 June 2016

Abstract We investigate a type of public goods games played in groups of individuals who choose how much to contribute towards the production of a common good, at a cost to themselves. In these games, the common good is produced based on the sum of contributions from all group members, then equally distributed among them. In applications, the dependence of the common good on the total contribution is often nonlinear (e.g., exhibiting synergy or diminishing returns). To date, most theoretical and experimental studies have addressed scenarios in which the set of possible contributions is discrete. However, in many real-world situations, contributions are continuous (e.g., individuals volunteering their time). The "n-player snowdrift games" that we analyze involve continuously varying contributions. We establish under what conditions populations of contributing (or "cooperating") individuals can evolve and persist. Previous work on snowdrift games, using adaptive dynamics, has found that what we term an "equally cooperative" strategy is *locally* convergently and evolutionarily stable. Using static evolutionary game theory, we find conditions under which this strategy is actually *globally* evolutionarily stable. All these results refer to stability to invasion by a single mutant. We broaden the scope of existing stability results by showing that the equally cooperative strategy is locally stable to potentially large population perturbations, i.e., allowing for the possibility that mutants make up a non-negligible proportion of the population (due, for example, to genetic drift, environmental variability or dispersal).

Keywords Evolutionary game theory · Cooperation · Snowdrift game · Adaptive dynamics

---

✉ Chai Molina
chai.molina@gmail.com

[1] Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada

Mathematics Subject Classification 92Bxx · 92b05 · 91A22 · 92D15 · 00A69

1 Introduction

Public goods games (Hauert et al. 2002; Archetti and Scheuring 2012) arise in a wide variety of biological and social contexts, ranging from microbial evolution (Cordero et al. 2012; Rainey and Rainey 2003), tumor growth (Archetti 2013), the evolution of virulence (Brown et al. 2002) and host manipulation by parasites (Brown 1999), to cooperative nesting and brood care (Leighton 2014; Houston and Davies 1985; Savage et al. 2015; Parker et al. 2002), the evolution of eusociality (Reeve and Hölldobler 2007), fisheries management (Kraak 2011) and family economics (Browning et al. 2014). These are games played among groups of individuals, who may choose to *cooperate* and contribute towards the production or attainment of a common good at a cost to themselves, or to *defect* and contribute nothing. The common good is then distributed among all members of the group (regardless of whether or not they contributed) (Kagel and Roth 1995). This situation is analogous to Hardin's "Tragedy of the Commons" (Hardin 1968), in which the cost of using a common resource is distributed among group members, but the benefit is personal (e.g., intrabrood competition for parental investment, Parker et al. 2002). In both cases, those who act selfishly (by refraining from contribution or by overexploitation), do better than group members who cooperate (either by contributing or by refraining from overexploiting the common resource). Because cooperative ventures are ubiquitous in nature (Dugatkin 1997; Clutton-Brock 2009), much research has been devoted to understanding how cooperation can evolve and persist (Motro 1991; Frank 2010; Boyd and Richerson 2002; Fehr and Gächter 1999); see recent reviews by Gavrilets (2015) and Gokhale and Traulsen (2014).

In experimental economics studies of human behaviour, public goods games are typically set up with a linear relationship between the total cost incurred by group members and the benefit they receive (Fehr and Gächter 2002; Kagel and Roth 1995; Milinski et al. 2002). However, in many biological scenarios, the benefit is a nonlinear function of the total cost (Hauert et al. 2006; Frank 2010; Archetti and Scheuring 2012; Brown 1999), as there may be a threshold (Souza et al. 2009; Bach et al. 2001), a synergistic effect of contributions (Archetti et al. 2015; Archetti and Scheuring 2011; Archetti 2014), diminishing returns (Motro 1991; Motro and Cohen 1989; Poulin 1994), or both synergy and diminishing returns (Archetti et al. 2015; Archetti and Scheuring 2011; Archetti 2014; Brown 1999). Furthermore, in both theoretical and experimental studies of public goods games, contribution levels are typically taken to be discrete: contribution may be an "all or nothing" affair, whereby a group member can either contribute a fixed, nonzero amount of a resource to the public good, or contribute nothing (Motro 1991; Hauert et al. 2006; Archetti et al. 2015; Souza et al. 2009) (usually in studies of the n-player prisoner's dilemma), or, more typically in the economics literature, players are endowed with a number of tokens and decide how many they wish to contribute (Fehr and Gächter 2002; Kagel and Roth 1995). In many real situations, however, individuals can vary their degree of cooperation, often continuously (McGill and Brown 2007; Pulliam et al. 1982; Cornforth et al. 2012;

Killingback and Doebeli 2002; Doebeli et al. 2004; Fujita et al. 2014; Brown and Vincent 2008). Further realism is often added to models by implementing population structure, (Santos et al. 2008; Hauert 2006, 2008), but we will avoid this further complication here.

The differences in evolutionary dynamics between 2- and n- player snowdrift games have been studied in games with binary strategies (Liang et al. 2015; Souza et al. 2009). In public goods games with a continuum of possible contributions, played in unstructured population, studies have investigated how the process by which individual contributions are aggregated affects the possibility of polymorphism (Ito et al. 2015; Sasaki and Okada 2015). Others have investigated how variability in group size (Brännström et al. 2010) and population dynamics (Hauert et al. 2006) affect the evolutionary outcomes.

Recently, interest in the influence of the functional form of the benefit of contribution on evolutionary dynamics of the snowdrift and other public goods games has increased. Most often, the effect of how the benefit depends on collective investment is investigated in the context of binary strategies (cooperate or defect) (Motro 1991; Liang et al. 2015; Archetti and Scheuring 2011), sometimes with the addition of population structure (e.g., Wu et al. 2009). However, Deng and Chu (2011) have investigated how evolutionary dynamics in continuous public goods games are influenced by nonlinearities in how collective investment is translated to the public good, using specific functional forms (linear, step function or sigmoid). While most other studies investigate stability of a homogeneous population against mutations that are close to the resident strategy, Deng and Chu were interested in stability against invasion by *any* strategy (in line with the original definition of evolutionary stability by Maynard Smith and Price 1973). They further considered invasion of populations by non-negligible proportions of invaders, using numerical simulations. Chen et al. (2012) have used simulations to study a similar game played on a spatial lattice using linear cost and two types of sigmoid benefit functions. They found that contributions to the public good are maximized at intermediate values of the steepness and threshold parameters of the sigmoid functions they used.

In this paper, we analyze a class of nonlinear public goods games with continuously varying contributions in unstructured populations and establish under what conditions populations of contributing (or cooperating) individuals can evolve and persist. Examples of public goods games to which our results apply include any in which the dependence of the benefits on the total cost is decelerating or sigmoidal (initially accelerating but eventually decelerating). Most of the specific public goods games considered in the literature fall in this class. Identifying general conditions for the evolution of cooperative strategies and their resistance to invasion is important, because it sheds light on what features of particular biological systems might be responsible for observed evolutionary outcomes. Moreover, since "all models are wrong" (Box 1976) (in the sense that no model can take all aspects of reality into account), general results on cooperation lend credibility to the broader application of qualitative conclusions obtained from highly specialized models of particular biological systems. Lastly, general results such as those obtained here can be useful in situations where exact analytical solution of a mathematical model is difficult or impossible.

In most previous studies on nonlinear public goods games with continuous contributions (e.g., Doebeli et al. 2004; Killingback et al. 2010; Cornforth et al. 2012), the framework of *adaptive dynamics* (Dieckmann and Law 1996; Geritz et al. 1998; Metz et al. 1996; Hofbauer and Sigmund 1998) has typically been used to analyze and determine evolutionary outcomes. The adaptive dynamics framework assumes an infinite population of a particular phenotype (that is, contribution level) and investigates evolutionary stability by considering a single mutant of a different type and determining whether it can invade the resident population. Because the population of residents is infinite, the effect of the mutant on the average fitness of the resident strategy is negligible.

Here, we compare the predictions of adaptive dynamics with those of *static evolutionary game theory* (Maynard Smith 1982; Hofbauer and Sigmund 1998; Nowak 2006) applied to a general class of nonlinear public goods games with continuous contributions. Our analysis still considers the limit of an infinite population, but allows mutants to comprise a *finite proportion* of the population; consequently, mutants can affect the average fitness of the resident population (and of other mutants). Our new analysis extends the predictions of adaptive dynamics on evolutionary and convergent stability (Sect. 3) of a cooperative strategy to biologically plausible scenarios in which genetic drift, migration, and/or environmental variability allow a mutant strategy to make up a significant part of the population (even if it is not selected for when initially rare). Our analysis also generalizes the results of Deng and Chu (2011) (who used Darwinian dynamics, Brown and Vincent 2008).

In Sect. 2, we motivate and construct the class of nonlinear public goods games that we analyze. Section 3 briefly reviews the two frameworks that we use to analyze these games. We present our results in Sect. 4 and proofs in Sects. 5, 6 and 7. Finally, in Sect. 8, we discuss our results and suggest directions for further developments.

2 Class of public goods games

Consider an infinite, well-mixed population of asexual agents. Assume that reproductive fitness is determined by playing a nonlinear public goods game in randomly-assembled groups of $n > 1$ agents. Let $h \geq 0$ be the focal agent's contribution to the public good, and let H denote the mean contribution by the other $n - 1$ agents in the focal agent's group.

Denote the fitness cost and fitness benefit to the focal agent by $c(h, H)$ and $b(h, H)$, respectively. The ***fitness of the focal agent*** is then

$$W(h, H) = b(h, H) - c(h, H), \tag{1}$$

where $b(h, H)$ and $c(h, H)$ are non-negative functions of their arguments.

If the cost of contributing is independent of the other group members' contributions, the focal agent's contribution h can be measured in units of the fitness cost of contribution to the public good. Thus, we henceforth assume (with some abuse of notation)

$$c(h, H) = c(h) = h. \tag{2}$$

The ***total good*** contributed by all members of the group is

$$\eta(h, H) = h + (n-1)H. \tag{3}$$

We assume that the resulting benefit to the focal agent is a continuous function of the total good (i.e., the sum of the individual fitness costs),

$$b(h, H) = f\big(\eta(h, H)\big). \tag{4}$$

Hence, the focal agent's fitness (1) is

$$W(h, H) = f\big(\eta(h, H)\big) - h, \tag{5}$$

which is a continuous function of h and H. Equations (3) and (5) define a large class of public goods games, namely, continuous n-player snowdrift (or hawk-dove) games (Maynard Smith 1982; Doebeli et al. 2004), in which the public good is fitness (see Appendix 3). A particular public goods game in this class is specified by choosing the function $f(\eta)$; see Fig. 1.

Biological intuition suggests that there may be a total contribution threshold, $\eta_{\min} > 0$, below which the marginal benefit of contribution does not outweigh its marginal cost. In that case, $W(h, H)$ decreases for all $h < \eta_{\min} - (n-1)H$. If we define $\eta_{\min} = 0$ in the absence of a range of h over which $W(h, H)$ decreases, then no generality is lost by assuming the existence of a threshold $\eta_{\min} \geq 0$. Below we will see that in the situations we consider the focal agent's fitness has a local minimum if $\eta = \eta_{\min}$; we therefore refer to $\eta_{\min}$ as the ***minimizing total good***.

We restrict the class of games we consider slightly by making the biologically sensible assumption that for any level of mean contribution (H) from the non-focal agents, there is a level of focal agent's contribution (h) beyond which its fitness decreases with its contribution; simply put, *the marginal cost of an increase in contribution eventually outweighs its benefit*. In Appendix 1, we show that this is equivalent to the existence of $\eta_{\max} > 0$ such that $f(\eta) - \eta$ decreases for $\eta \geq \eta_{\max}$. Consequently, the focal agent's fitness $W(h, H)$ decreases with its contribution h when $\eta(h, H) > \eta_{\max}$. In the situations we consider the focal agent's fitness has a local maximum if $\eta = \eta_{\max}$; consequently, we refer to $\eta_{\max}$ as the ***maximizing total good***.

For convenience, we define

$$h_{\min}(H) = \eta_{\min} - (n-1)H, \tag{6a}$$

$$h_{\max}(H) = \eta_{\max} - (n-1)H. \tag{6b}$$

For a given mean contribution H by the non-focal agents, $h_{\min}$ and $h_{\max}$ are the levels of contribution required by the focal agent so that the total good is $\eta_{\min} = \eta(h_{\min}(H), H)$ and $\eta_{\max} = \eta(h_{\max}(H), H)$, respectively. Note that $h_{\min}$ and $h_{\max}$ are always well-defined mathematically but they can be negative and hence not biologically meaningful: if the non-focal group members contribute $(n-1)H > \eta_{\min}$ then $h_{\min}(H) < 0$ and if $(n-1)H > \eta_{\max}$ then $h_{\max}(H) < 0$.

If for any mean non-focal agents' contribution H and focal agent's contribution, h, the marginal costs of contributing outweigh the marginal benefits, then $W(h, H)$

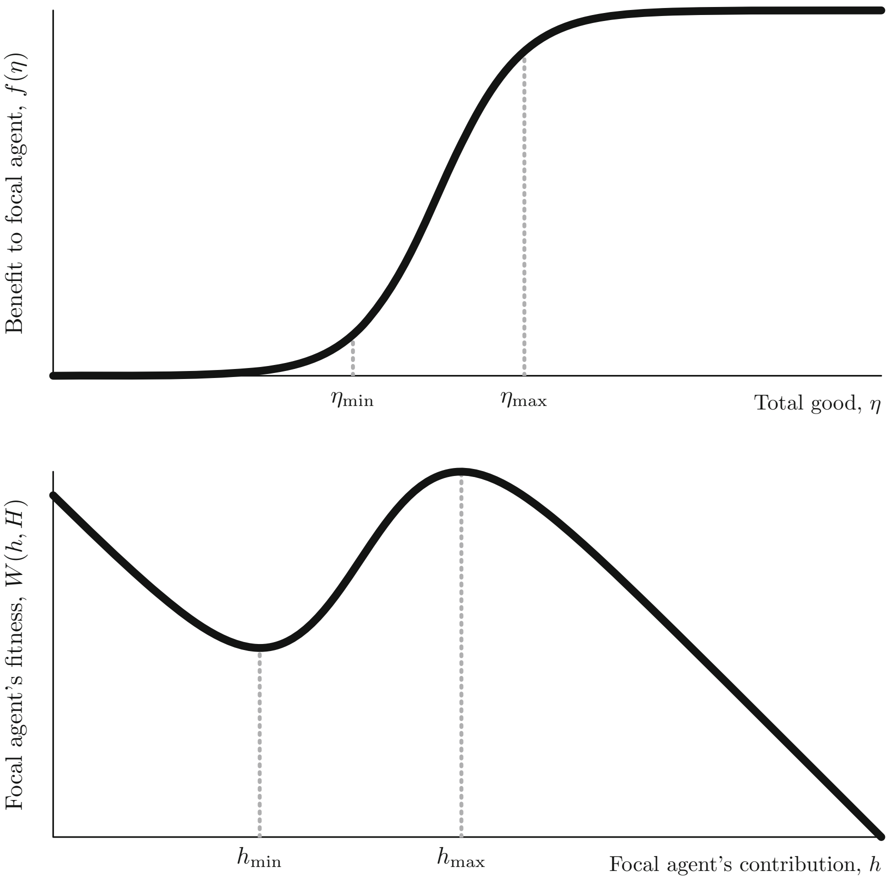

Fig. 1 Sigmoidal benefit to the focal agent, and its corresponding fitness. *Top panel* sigmoidal benefit $f(\eta) = a\big(\beta + \exp(\kappa - b\eta)\big)^{-1} - a\big(\beta + \exp(\kappa)\big)^{-1}$, with $a = 100$, $b = 0.2$, $\beta = 2$, $\kappa = 10$, based on an example from Cornforth et al. (2012) ($\eta_{\min}$ and $\eta_{\max}$ indicated by *dashed lines*). *Bottom panel* The focal agent's fitness $W(h, H)$ (corresponding to the benefit function $f(\eta)$ above) as a function of its contribution h, for fixed mean non-focal agents' contribution $H = 3$ and group size $n = 6$ ($h_{\min}$ and $h_{\max}$ indicated by *dashed lines*)

decreases with h regardless of H. Consequently, the unique evolutionarily stable strategy is not to contribute ($h = H = 0$), and it is convergently stable (Sect. 3). In order to avoid this trivial outcome, we assume further that

$$\eta_{\min} < \eta_{\max}. \tag{7}$$

It then follows that fitness $W(h, H)$ decreases with h if the total good $\eta(h, H) < \eta_{\min}$, i.e., for $0 \leq h < \max\{0, h_{\min}(H)\}$. Equivalently, $f(\eta) - \eta$ decreases for $0 \leq \eta < \eta_{\min}$. Note that the intervals $[0, \max\{0, h_{\min}(H)\})$ and $[0, \eta_{\min})$ are degenerate if the right endpoint is 0.

For simplicity, we assume that a focal agent's fitness $W(h, H)$ increases as a function of its contribution h if the total good $\eta(h, H)$ is between $\eta_{\min}$ and $\eta_{\max}$,

$$\eta_{\min} < \eta(h, H) < \eta_{\max}, \tag{8}$$

or, equivalently, if its contribution h lies in the interval

$$\max\{0, h_{\min}(H)\} < h < \max\{0, h_{\max}(H)\}. \tag{9}$$

Because the fitness benefit $f(\eta(h, H))$ and fitness cost $c(h) = h$ are both increasing functions of h, assuming $W(h, H)$ increases with h means that the benefit of contributing more increases faster than the cost over interval (9); equivalently, $f(\eta) - \eta$ is an increasing function of η for $\eta_{\min} < \eta < \eta_{\max}$.

We can now justify our terminology for $\eta_{\min}$ and $\eta_{\max}$. Our assumptions,

A1 fitness is specified by Eqs. (3) and (5),
A2 $f(\eta)$ is a continuous function defined for $\eta \geq 0$,
A3 $\eta_{\max} > 0$ exists,
A4 if $\eta_{\min}$ exists (which can be assumed without loss of generality) then $0 \leq \eta_{\min} < \eta_{\max}$,
A5 $f(\eta) - \eta$ increases with η when $\eta_{\min} < \eta < \eta_{\max}$, and decreases otherwise,

ensure that for a fixed $H \leq \eta_{\max}/(n-1)$, the focal agent's fitness $W(h, H)$ has a local maximum when the total good $\eta(h, H) = \eta_{\max}$ (i.e., $h = h_{\max}(H) \geq 0$) and a local minimum when the total good $\eta(h, H) = \eta_{\min}$.

Thus, our assumptions describe an n-player snowdrift game with cost $c(h) = h$, and continuous benefit $f(\eta(h, H))$, such that $f(\eta)$ increases in η with slope greater than 1 on a bounded interval, $(\eta_{\min}, \eta_{\max})$, and nowhere else.

3 Analysis frameworks

Two frameworks commonly used in analyzing models such as those in the class described in Sect. 2 are static evolutionary game theory (Maynard Smith 1982; Nowak 2006; Hofbauer and Sigmund 1998) and adaptive dynamics (Dieckmann and Law 1996; Geritz et al. 1998; Metz et al. 1996; Hofbauer and Sigmund 1998). Below, we recall some of the main concepts from these frameworks, as they apply to our analysis. For a general treatment, see the references cited above.

3.1 Static evolutionary game theory

Definition 3.1 (Evolutionary stability) A contribution level $\hat{H} \geq 0$ is (globally) ***evolutionarily stable*** (ES) *iff* a single agent that plays a different strategy cannot invade the population (all strategies different from $\hat{H}$ are selected against) (Maynard Smith and Price 1973).

As different levels of contributions constitute strategies in this game, we also use the term ***evolutionarily stable strategy*** (ESS), when referring to a level of contribution that is ES.

Since evolution by natural selection typically involves mutations that have a small phenotypic effect, the following definition is also biologically relevant:

Definition 3.2 (Local Evolutionary stability) A contribution level $\hat{H} \geq 0$ is ***locally evolutionarily stable*** (locally ES) if a single agent playing a mutant strategy h different from, but sufficiently close to $\hat{H}$ cannot invade the population (h is selected against if $|\hat{H} - h|$ is sufficiently small) (Taylor 1989; Hofbauer and Sigmund 1998).

Definition 3.3 (Local convergent stability) A contribution level $\hat{H} \geq 0$ is ***locally convergently stable*** (locally CS) if, when the resident strategy H is close enough to $\hat{H}$, a mutant playing a strategy between H and $\hat{H}$ can invade the population (h is selected for if $H < h \leq \hat{H}$ or $\hat{H} \leq h < H$) (Christiansen 1991).

3.2 Adaptive dynamics

Adaptive dynamics (Dieckmann and Law 1996; Geritz et al. 1998; Metz et al. 1996) can also be used to gain insight into similar evolutionary problems. In particular, Doebeli et al. (2004) use the adaptive dynamics framework to completely characterize the evolutionary dynamics of the continuous snowdrift game with smooth payoffs. Since the class of models defined in Sect. 2 is a large subclass of realistic snowdrift games, it is interesting to compare the predictions of Doebeli et al. (2004) to our predictions based on static evolutionary game theory. We therefore briefly outline concepts from adaptive dynamics necessary for this comparison.

Following Geritz et al. (1998) and Doebeli et al. (2004), the ***growth rate*** of a rare mutant strategy h in a resident population playing H is

$$s_H(h) = W(h, H) - W(H, H), \tag{10}$$

where $W(x, y)$ is the fitness of a mutant playing x in a population playing y. The ***local fitness gradient*** is then

$$D(H) = \left. \frac{\partial s_H(h)}{\partial h} \right|_{h=H}, \tag{11}$$

and the ***adaptive dynamics*** of H are given by

$$\dot{H} = D(H). \tag{12}$$

An equilibrium of Eq. (12), that is, $\hat{H}$ satisfying $D(\hat{H}) = 0$, is called a ***singular strategy***. A singular strategy that is an attractor of Eq. (12) is convergently stable in the sense of Definition 3.3. A singular strategy H can also be locally evolutionarily stable as in Definition 3.2. The mathematical conditions for these and other possible characteristics of singular strategies are listed in Table 1, following Geritz et al. (1998).

4 Results

Below, we summarize our results on the behaviour of the class of models outlined in Sect. 2, using static evolutionary game theory and adaptive dynamics. These results are proved in Sects. 5, 6 and 7.

Table 1 Local properties of singular strategies in adaptive dynamics, as in (Geritz et al. 1998, Table 1)

Property	Characterization
Local evolutionary stability	$\left.\frac{\partial^2 s_H(h)}{\partial h^2}\right\|_{h=H} < 0$
Convergence stability	$\left.\frac{\partial^2 s_H(h)}{\partial H^2} - \frac{\partial^2 s_H(h)}{\partial h^2}\right\|_{h=H} > 0$
Singular strategy can spread in populations playing similar strategy	$\left.\frac{\partial^2 s_H(h)}{\partial H^2}\right\|_{h=H} > 0$
Mutually-invasible strategies exist near singular point	$\left.\frac{\partial^2 s_H(h)}{\partial H^2} + \frac{\partial^2 s_H(h)}{\partial h^2}\right\|_{h=H} > 0$

Theorem 4.1 (Evolutionary and convergent stability in static theory). *Consider an evolving, infinite, well-mixed population in which fitness is determined by the payoff from a non-linear public goods game played in randomly chosen groups of* $n > 1$ *agents. Suppose that if the total public good contributed is* $\eta \geq 0$, *the benefit to any group member is* $f(\eta)$. *Thus, if* h *is a focal agent's contribution and* H *is the mean non-focal agents' contribution to the public good, the focal agent's fitness is*

$$W(h, H) = f\big(h + (n-1)H\big) - h, \tag{13}$$

assuming the cost of the focal agent's contribution is independent of the other member's contributions and contribution is measured in units of its fitness cost. Suppose further that the benefit function f *has the following properties:*

H1 f *is continuous on* $\eta \geq 0$.
H2 *There exist* $\eta_{\min} \geq 0$ *and* $\eta_{\max} > \eta_{\min}$ *such that* $f(\eta) - \eta$ *increases for* $\eta_{\min} < \eta < \eta_{\max}$ *and decreases for* $\eta < \eta_{\min}$ *and* $\eta > \eta_{\max}$.

Then, writing $H^* = \eta_{\max}/n$,

- *If* $f(\eta_{\max}) \geq \eta_{\max}$ *then the unique ES contribution is* H^*.
- *If* $f(\eta_{\max}) < \eta_{\max}$ *then*

$$f\big(nH^*\big) - f\big((n-1)H^*\big) \begin{array}{lll} > H^* & \Longrightarrow & \textit{two ESSs: } H = 0 \quad \textit{and} \quad H = H^*, \\ \leq H^* & \Longrightarrow & \textit{unique ESS: } H = 0. \end{array} \tag{14}$$

Moreover, all existing ESSs are convergently stable.

Remark 4.2 As shown in the proof of Theorem 4.1, the focal agent's optimal response $h_{\text{opt}}(H)$ (see Sect. 5.1) is a piecewise-linear function of the mean contribution of the non-focal agents (this can also be seen graphically in Fig. 3).

Note that if $f(\eta_{\max}) \geq \eta_{\max}$ then it is worthwhile for the focal agent to contribute the maximizing total good, even if it must do so single-handedly (see Fig. 3, first and second panels). The existence of a nonzero ES level of contribution is not surprising in this case. Condition (14) says that in the non-trivial situation that $f(\eta_{\max}) < \eta_{\max}$, contributing H^* (i.e., an equal share of the maximizing total good $\eta_{\max}$) is an ESS *iff*, when all non-focal agents contribute H^*, the cost of contributing H^* (rather than defecting and contributing nothing) is smaller than the increase in the focal agent's benefit resulting from this contribution.[1]

The corresponding analysis based on adaptive dynamics yields:

Theorem 4.3 (Local evolutionary and convergent stability in adaptive dynamics) *If the hypotheses of Theorem 4.1 are satisfied and, in addition,*

H3 *f is twice-differentiable on $\eta \geq 0$,*

then the adaptive dynamics of H are given by

$$\dot{H} = f'(nH) - 1 \tag{15}$$

and there are two singular points, $H = \eta_{\min}/n$ and $H^* = \eta_{\max}/n$. $H = \eta_{\min}/n$ is a repellor and H^* is an attractor (i.e., convergently stable) and a local ESS.

Figure 2 gives the pairwise invasibility plot (Geritz et al. 1998; Metz et al. 1996) corresponding to the particular choice of f used in Fig. 1 (see captions for details).

Theorems 4.1 and 4.3 rely on the assumption of an infinite population to assert that when a mutant arises, the mean fitness of the resident strategy is unaffected by mutant's behaviour (Nowak and Sigmund 2004; Killingback et al. 2010; Metz et al. 1996). In other words, the average resident does not interact with a mutant. Similarly, it is assumed that the average mutant does not interact with other mutants, so the mean mutant fitness is unaffected by the presence of other mutants (if other mutants exist).

Because real populations are finite, the presence of an invader may well affect the resident's mean fitness, even if the population size is large. For example, in the case of our public goods game, even if there is only one mutant, there are $n-1$ residents in its group, whose fitness is $W(H, \frac{h+(n-2)H}{n-1})$, rather than $W(H, H)$. Thus, in a finite population of size N, even a single invader comprises a nonzero proportion $\epsilon = 1/N > 0$ of the population.

In Theorem 4.4 below, we relax the assumption that mutants do not affect the resident (or mutant) fitnesses. We retain the assumption of an infinite population, but when considering invasion scenarios, we allow individuals playing the mutant strategy to make up a finite proportion, ϵ, of the population. This new analysis is biologically relevant for at least three reasons:

(i) A mutation might not be selected for when present in a single individual, but spread nevertheless by genetic drift (Hartl and Clark 2007). Once present in a non-negligible proportion of the population, the mutation could be selected for. Thus, we use phenotypic theory to address the question of whether an initially non-adaptive mutation that drifts in to become present in a non-negligible proportion of the population can then be selected for.

(ii) A "bud" consisting of multiple mutants can invade a resident population in a dispersal or migration event (e.g., Krupp and Taylor 2015; Gardner and West

---

[1] Condition (14) compares the ***incremental benefit*** of contributing H^* to its cost.

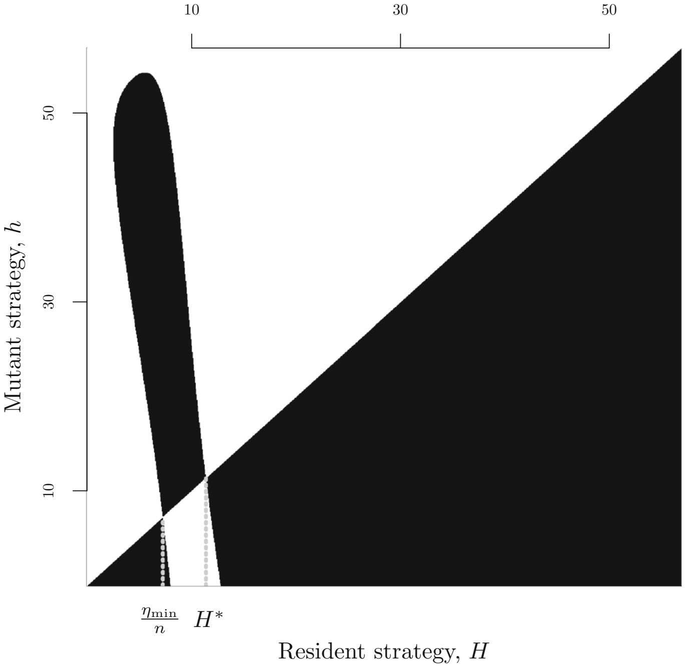

Fig. 2 Pairwise invasibility plot for sigmoidal benefit $f(\eta)$ with parameters as in Fig. 1: areas where a single mutant contributing h (*vertical axis*) can invade an infinite population of agents contributing H (*horizontal axis*) are *shaded*. The singular points $H = \eta_{\min}/n$ and $H = H^*$ are *marked*. The predictions of theorems 4.1 and 4.3 for this choice of benefit function are that $H = 0$ and $H = H^*$ are the only ESSs and are convergently stable (because $f(\eta_{\max}) < \eta_{\max}$ and $f\big(nH^*\big) - f\big((n-1)H^*\big) > H^*$), and that $\eta_{\min}/n$ is a repellor. The *vertical lines* at $H = H^*$ and $H = 0$ are unshaded, implying that these are ESS contributions (because no mutant can invade). Near H^*, resident strategies $H \neq H^*$ can be invaded by mutants playing strategies h that are closer to H^*, so H^* is convergently stable, and similarly, so is $H = 0$. $\eta_{\min}/n$ is a repellor, since resident strategies H near $\eta_{\min}/n$ can be invaded by mutants playing h that is farther away from $\eta_{\min}/n$ than H

2006; Kümmerli et al. 2009; Clutton-Brock 2002; Pfeiffer and Bonhoeffer 2003 and references therein), in which case the invading mutants may comprise a non-negligible proportion of the population.

(iii) A mutant may under normal conditions be selected against when rare, but due to an environmental disturbance (either anthropogenic or natural), conditions may temporarily change to allow the mutant to spread (similarly to disturbances of ecological communities, Buma and Wessman 2011; Dale et al. 2001). When the environmental conditions return to normal, the proportion of mutants in the population may have already become non-negligible. In such cases, the invasion analysis must account for more than a single mutant.

For simplicity, we state Theorem 4.4 with the restriction that $f(\eta_{\max}) < \eta_{\max}$. As noted above, the existence of a nonzero ESS level of contribution when $f(\eta_{\max}) \geq \eta_{\max}$ is trivial. Theorem 4.4 then extends the results of Theorem 4.1 on evolutionary and convergent stability to scenarios where the invading mutants comprise a proportion

$\epsilon > 0$ of the population. The result we obtain is weaker, however, in that H^* is no longer guaranteed to be globally ES; it is resistant to invasion by nearby strategies only.

Theorem 4.4 (Local evolutionary and convergent stability in static theory with a finite proportion of mutants). *Suppose the hypotheses of Theorem 4.1 are satisfied and write* $H^* = \eta_{\max}/n$. *If* $f(\eta_{\max}) < \eta_{\max}$ *then:*

R1 (*Local ES*) *If* $h \neq H^*$ *and a proportion* ϵ *of mutants playing* h *arises in a population playing* H^* *then, if* h *is sufficiently close to* H^*, *the mean fitness of a mutant* $\overline{\mathrm{W}}_{\mathrm{m}}(h)$ *is smaller than the mean fitness of a resident* $\overline{\mathrm{W}}_{\mathrm{r}}(h)$, *for any proportion* $\epsilon > 0$ (*i.e., the mutants are selected* <u>*against*</u>).

R2 (*Local CS*) *If* H *is sufficiently close to* H^*, h *is between* H *and* H^*, *and a proportion* ϵ *of mutants playing* h *arises in a population playing* H, *then the mean fitness of a mutant* $\overline{\mathrm{W}}_{\mathrm{m}}(h)$ *is greater than the mean fitness of a resident* $\overline{\mathrm{W}}_{\mathrm{r}}(h)$, *for any proportion* $\epsilon > 0$ (*i.e., the mutants are selected* <u>*for*</u>).

5 Proof of Theorem 4.1

Without loss of generality, we can assume that

$$f(0) = 0. \tag{16}$$

To see this, note that the dynamics of the game are not changed by adding a constant to the fitness function $W(h, H)$. If $f(0) \neq 0$ then we would analyze $\tilde{f}(\eta) = f(\eta) - f(0)$ (so $\tilde{f}(0) = 0$) and $\tilde{W}(h, H) = W(h, H) - f(0) = \tilde{f}(\eta(h, H)) - h$.

The structure of our proof of Theorem 4.1 is as follows: In Sect. 5.1, we find the optimal response for the focal agent as a function of the non-focal agents' mean contribution, H. Then, in Sect. 5.2, we use the focal agent's optimal response to show that the only possible ESSs are either not contributing ($H = 0$), or contributing an equal share of the maximizing total good, $H^* = \eta_{\max}/n > 0$. Lastly, in Sect. 5.3, we show that these ESSs are also convergently stable.

5.1 Optimal response for focal agent

For a given mean contribution from the non-focal agents, H, what must the focal agent contribute in order to maximize its fitness? This is the focal agent's ***optimal response*** to H, which we denote $h_{\mathrm{opt}}(H)$. The optimal response in the class of games we are considering is given in Lemma 5.1 and plotted in Fig. 3.

Lemma 5.1 (Best response lemma) *Under the conditions of Theorem 4.1*

- *if* $f(\eta_{\max}) > \eta_{\max}$, *then*

$$h_{\mathrm{opt}}(H) = \max\left\{0, h_{\max}(H)\right\} = \begin{cases} h_{\max}(H) & 0 \leq H < \frac{\eta_{\max}}{n-1}, \\ 0 & \frac{\eta_{\max}}{n-1} \leq H, \end{cases} \tag{17a}$$

where $h_{\max}(H)$ *is defined in Eq.* (6b). *Note that the interval* $[0, \eta_{\max}/(n-1))$ *is never empty, because* $\eta_{\max} > 0$.

- *if* $f(\eta_{\max}) \leq \eta_{\max}$, *then,*

$$h_{\mathrm{opt}}(H) = \begin{cases} 0 & 0 \leq H < H^0, \\ 0 \quad \mathit{or} \quad h_{\max}(H^0) & H = H^0, \\ h_{\max}(H) & H^0 < H < \frac{\eta_{\max}}{n-1}, \\ 0 & \frac{\eta_{\max}}{n-1} \leq H, \end{cases} \tag{17b}$$

where H^0 *is the unique solution of*

$$f\big((n-1)H\big) - (n-1)H = f(\eta_{\max}) - \eta_{\max} \tag{18}$$

such that $0 \leq H^0 < \frac{\eta_{\min}}{n-1}$.

In Eq. (17b), *note that the first interval* $(0 \leq H < H^0)$ *is empty if* $H^0 = 0$, *and that for* $H = H^0$, $h_{\mathrm{opt}}(H)$ *is 2-valued.*

Proof Observe that depending on the mean non-focal agents' contribution H, $W(h, H)$ behaves in one of the following ways:

1. If the non-focal agents' contribution is no less than the maximizing total good ($(n-1)H \geq \eta_{\max}$), then the focal agent's fitness $W(h, H)$ decreases with its contribution, h. The optimal contribution for the focal agent is then $h_{\mathrm{opt}}(H) = 0$.
2. If the non-focal agents' contribution is greater than or equal to the minimizing total good and smaller than the maximizing total good ($\eta_{\min} \leq (n-1)H < \eta_{\max}$), then the focal agent's fitness $W(h, H)$ increases for $0 \leq h \leq h_{\max}(H)$ and decreases for $h > h_{\max}(H)$, so the focal agent's optimal contribution is $h_{\mathrm{opt}}(H) = \eta_{\max} - (n-1)H$.
3. If the non-focal agents' contribution is lower than the minimizing total good ($(n-1)H < \eta_{\min}$), fitness decreases for $0 \leq h \leq h_{\min}(H)$, increases for $\max\{0, h_{\min}(H)\} < h \leq h_{\max}(H)$ and decreases again for $h > h_{\max}(H)$.
   Thus, two levels of contribution locally maximize the focal agent's fitness: $h = \eta_{\max} - (n-1)H$ and $h = 0$, for which the focal agent's fitness is $W(\eta_{\max} - (n-1)H, H) = f(\eta_{\max}) - \eta_{\max} + (n-1)H$ and $W(0, H) = f((n-1)H)$, respectively. These two local fitness maxima are the candidates for the global fitness maximum, that is, the focal agent's optimal response $h_{\mathrm{opt}}(H)$. Let

$$\begin{aligned} \Delta W(H) &= W(\eta_{\max} - (n-1)H, H) - W(0, H) \\ &= f(\eta_{\max}) - \eta_{\max} - [f((n-1)H) - (n-1)H] \end{aligned} \tag{19}$$

be the difference between the focal agent's two local fitness maxima. Note that since $f(\eta) - \eta$ decreases with η on $[0, \eta_{\min}]$, $\Delta W(H)$ is an increasing function of H on $\left[0, \frac{\eta_{\min}}{n-1}\right)$.

Because $f(\eta) - \eta$ increases with η on $[\eta_{\min}, \eta_{\max}]$, it follows that

$$\Delta W\left(\frac{\eta_{\min}}{n-1}\right) = f(\eta_{\max}) - \eta_{\max} - [f(\eta_{\min}) - \eta_{\min}] > 0. \tag{20}$$

Thus, since $\Delta W(H)$ is continuous, it follows that for large enough values of $H < \eta_{\min}/(n-1)$, $\Delta W(H) > 0$, implying that the focal agent maximizes fitness by contributing $h_{\text{opt}}(H) = h_{\max}(H) > 0$.
At the other extreme end of the interval $0 \leq H < \eta_{\min}/(n-1)$, we have:

$$\Delta W(0) = f(\eta_{\max}) - \eta_{\max}. \tag{21}$$

There are three possible cases:

(i) If $\Delta W(0) > 0$, then for all $0 \leq H < \eta_{\min}/(n-1)$, the focal agent's optimal response is $h_{\text{opt}}(H) = h_{\max}(H) > 0$.
The condition $\Delta W(0) > 0$ has a simple biological interpretation: If $\Delta W(0) > 0$, then the benefit to the focal agent when the total good is equal to the total maximizing good ($\eta = \eta_{\max}$) is so large that—even if the non-focal group members contribute nothing—the focal agent gains by single-handedly contributing the total maximizing good ($h = \eta_{\max}$). It is then sensible that if the non-focal agents have collectively contributed less than the minimizing total good $(n-1)H < \eta_{\min}$ (or indeed, less than the maximizing total good, $\eta_{\max}$), then the focal agent still does best by ensuring that the total maximizing good is attained ($\eta = \eta_{\max}$) (in fact, if the non-focal agents contribute $0 < (n-1)H < \eta_{\max}$, the focal agent's fitness must be higher than when $H = 0$, since it is now required to contribute less to obtain the same benefit).

(ii) If $\Delta W(0) < 0$, then because $\Delta W(H)$ is continuous, increasing, and $\Delta W\left(\frac{\eta_{\min}}{n-1}\right) > 0$, it follows that there is a unique solution to

$$\Delta W(H^0) = 0, \quad 0 < H^0 < \frac{\eta_{\min}}{n-1}. \tag{22}$$

Moreover, $\Delta W(H) < 0$ for $H < H^0$ and $\Delta W(H) > 0$ for $H > H^0$.
Thus, the optimal response for the focal agent is $h_{\text{opt}}(H) = 0$ if $H < H^0$, and $h_{\text{opt}}(H) = h_{\max}(H) > 0$ if $H^0 < H < \frac{\eta_{\min}}{n-1}$. If $H = H^0$, the focal agent can maximize its fitness by contributing either $h = 0$ or $h = \eta_{\max} - (n-1)H^0 > 0$ (because $W(0, H^0) = W(\eta_{\max} - (n-1)H^0, H^0)$).
H^0 is the mean non-focal agents' contribution for which the focal agent obtains the same fitness either by contributing nothing ($h = 0$), or by completing the difference between the maximizing total good and the collective contribution of the non-focal agents (so that $\eta = \eta_{\max}$).

(iii) If $\Delta W(0) = 0$, then $\Delta W(H) > 0$ for all $H > 0$, and $h_{\text{opt}}(H) = h_{\max}(H) > 0$ for $0 < H < \eta_{\min}$. For $H = 0$, the focal agent can maximize its fitness by contributing either $h = 0$ or $h = \eta_{\max} > 0$ (again, $W(0, 0) = W(\eta_{\max}, 0)$).

□

5.2 Evolutionarily stable contribution levels

To interpret Definition 3.1 for evolutionary stability mathematically, suppose that the entire population consists of agents playing H. The focal agent's fitness is given by Eq. (5) where h can be an alternative strategy $h \neq H$. The fitness of the $n-1$ non-focal individuals in the focal agent's group is

$$W\left(H, \frac{h+(n-2)H}{n-1}\right) = f\big(\eta(h,H)\big) - H. \tag{23}$$

Thus, the focal agent's fitness is larger than that of the non-focal individuals in its group *iff* $h < H$. However, since the population is infinitely large, an infinite number of non-focal individuals are in homogeneous groups in which everyone contributes H, so their fitness is $W(H, H)$. Thus, the mean fitness of a non-focal individual remains $W(H, H)$. Then, $H = H^*$ is an ESS *iff* when all non-focal agents play H^*, if the focal agent plays an alternative strategy $h \neq H^*$, its fitness is lower than the resident strategy H^*, or: $W(h, H^*) < W(H^*, H^*)$ for all $h \neq H^*$.

Thus, $H^* \geq 0$ is an ESS if and only if it is the unique optimal response to itself. Explicitly, H^* must satisfy

$$h_{\mathrm{opt}}(H^*) = H^*, \tag{24}$$

and $h_{\mathrm{opt}}(H^*)$ must be univalued. To see that a point at which $h_{\mathrm{opt}}(H)$ is multi-valued cannot be an ESS, suppose that $h_{\mathrm{opt}}(H^*)$ takes the two values H^* and H'; then if the non-focal agents play H^*, the focal agent can play H' without decreasing its fitness. Geometrically, solutions of Eq. (24) are intersections (in the H–h plane) of the curve $h = h_{\mathrm{opt}}(H)$ with the line $h = H$.

Note that while solutions of Eq. (24) at which h_{opt} is not univalued are not ESSs, they are still technically Nash Equilibria (Fudenberg and Tirole 1991).

We separate the discussion into the following cases:

1. $\underline{f(\eta_{\max}) > \eta_{\max}}$:
   In this case, $h_{\mathrm{opt}}(H)$ is given by Eq. (17a) (see Fig. 3, top panel). Solving Eq. (24) yields a unique ESS,

   $$H^* = \frac{\eta_{\max}}{n}. \tag{25}$$

   Note that when $f(\eta_{\max}) > \eta_{\max}$, it is beneficial for the focal agent to ensure the maximizing total good is attained even if it must do so single-handedly (so there is no "tragedy of the commons" in this scenario). Thus, it is biologically sensible that at the ESS all group members contribute equally towards the maximizing total good.
2. $\underline{f(\eta_{\max}) = \eta_{\max}}$:
   In this case, $h_{\mathrm{opt}}(H)$ is given by Eq. (17b) with $H^0 = 0$, that is,

$$h_{\mathrm{opt}}(H) = \begin{cases} 0 \quad \text{or} \quad \eta_{\max} & H = 0, \\ h_{\max}(H) & 0 < H < \frac{\eta_{\max}}{n-1}, \\ 0 & H \geq \frac{\eta_{\max}}{n-1}. \end{cases} \tag{26}$$

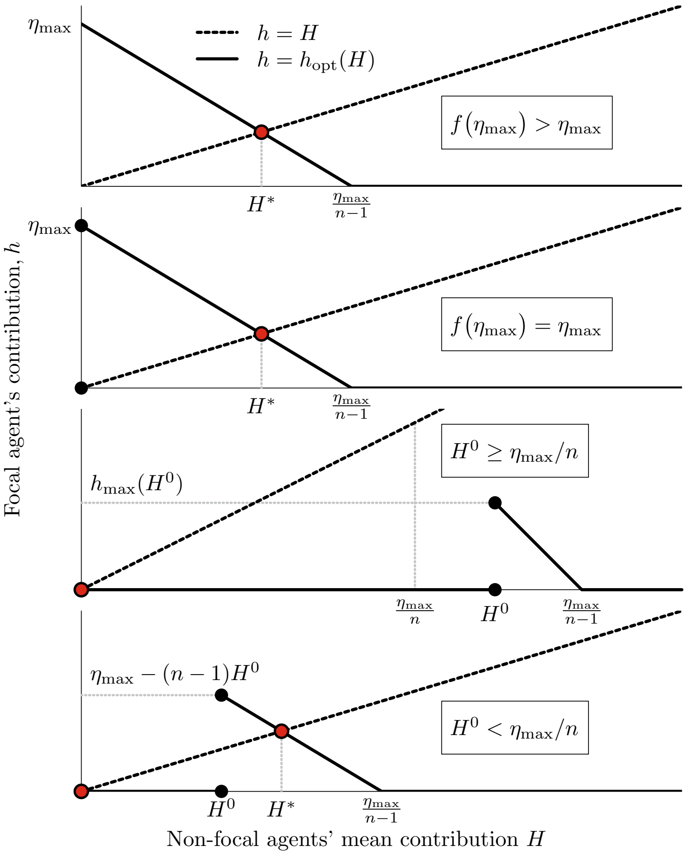

Fig. 3 ESSs (*red dots*) are solutions of $H = h_{\text{opt}}(H)$ [Eq. (24)] at which $h_{\text{opt}}(H)$ is univalued. $H = h_{\text{opt}}(H)$ implies $h = H$ is a best response to itself, and $h_{\text{opt}}(H)$ being univalued ensures that no other invading strategy matches the residents' fitness. The four panels (*top to bottom*), depict the intersections of $h = H$ (*dashed black line*) and $h = h_{\text{opt}}(H)$ (*solid black line*) in the h-H plane in the four possible cases: $f(\eta_{\max}) > \eta_{\max}$, $f(\eta_{\max}) = \eta_{\max}$, $H^0 \geq \eta_{\max}/n$, and $H^0 < \eta_{\max}/n$. For values of H (the mean non-focal agents' help) where the focal agent's best response $h_{\text{opt}}(H)$ is two-valued, its values are indicated by *black dots*

Thus, $H = 0$ is not an ESS, because if the non-focal agents contribute $H = 0$, the focal agent's fitness if it contributes $h = \eta_{\max} > 0$ is identical to its fitness if it does not contribute ($h = 0$).

If $H \geq \frac{\eta_{\max}}{n-1}$, then $h_{\text{opt}}(H) = 0 < H$ and so H cannot be an ESS. Lastly, if $0 < H < \frac{\eta_{\max}}{n-1}$ then solving Eq. (24) again yields a unique ESS given by Eq. (25) (see Fig. 3, second panel).

The biological interpretation of the ESS $H^* = \eta_{\max}/n$ is similar to the previous case ($f(\eta_{\max}) > \eta_{\max}$). The only change is that now, $H = 0$ (no contribution) is a Nash Equilibrium (whereas it was not when $f(\eta_{\max}) > \eta_{\max}$), because when the non-focal agents do not contribute, the focal agent's fitness can be maximized either by contributing $\eta_{\max}$ or not contributing. By contrast, $H^* = \eta_{\max}/n$ is still

ES, because when all group members contribute H^*, then the focal agent's fitness will decrease if it contributes $h \neq H^*$.

3. $\underline{f(\eta_{\max}) < \eta_{\max}}$: In this case, h_{opt} is given by Eq. (17b), so $H = 0$ is ES.

   For $0 < H < H^0$ and $H \geq \frac{\eta_{\max}}{n-1}$, $h_{\text{opt}}(H) = 0 < H$, so Eq. (24) is not satisfied, and H cannot be ES. Likewise, $H = H^0$ cannot be an ESS, since $h_{\text{opt}}(H^0)$ is not univalued.

   However, depending on the relationship between $\eta_{\max}$, n and H^0, there may or may not be another ESS in the range $H^0 < H < \frac{\eta_{\max}}{n-1}$:

   (a) $\underline{H^0 \geq \eta_{\max}/n}$:

   When $H^0 \geq \eta_{\max}/n$, there is no additional (nonzero) ESS in the range $H^0 < H < \frac{\eta_{\max}}{n-1}$, because

$$h_{\text{opt}}(H) < \eta_{\max} - (n-1)H^0 < \frac{\eta_{\max}}{n} < H, \tag{27}$$

   (recall that $h_{\text{opt}}(H)$ decreases linearly with H in this range; see Fig. 3, third panel). Thus, for this sub-case, $H^* = 0$ is the unique ESS.

   The condition $H^0 \geq \eta_{\max}/n$ is equivalent to $\Delta W(\eta_{\max}/n) \leq 0$, or

$$\begin{aligned} W\left(\frac{\eta_{\max}}{n}, \frac{\eta_{\max}}{n}\right) &= f(\eta_{\max}) - \frac{\eta_{\max}}{n} \\ &\leq f\left(\frac{n-1}{n}\eta_{\max}\right) = W\left(0, \frac{\eta_{\max}}{n}\right). \end{aligned} \tag{28}$$

   Condition (28) simply states that if all agents contribute equally towards achieving the maximizing total good, the focal agent does no worse by withdrawing its contribution (i.e., switching to $h = 0$).

   (b) $\underline{H^0 < \eta_{\max}/n}$:

   To see that in this case, there is a second (nonzero) ES level of contribution H, we seek a solution of Eq. (24) in the range $H^0 < H < \frac{\eta_{\max}}{n-1}$. In this range, $h_{\text{opt}}(H) = \eta_{\max} - (n-1)H$, so $H^* = \eta_{\max}/n$ solves Eq. (24) ($H^0 < H^* = \eta_{\max}/n$ by our assumption for this sub-case, and $H^* = \eta_{\max}/n < \eta_{\max}/(n-1)$ trivially, so $H^* \in \left(H^0, \frac{\eta_{\max}}{n-1}\right)$ as required).

   Thus, there are in this case two ESS contribution levels: $H = 0$ and $H = \frac{\eta_{\max}}{n}$ (see Fig. 3, bottom panel).

   To understand why there is an additional (non-zero) ES level of contribution when $H^0 < \eta_{\max}/n$, note that by definition, $H^0 < \frac{\eta_{\min}}{n-1}$ [see Eq. (22)], so there are two possibilities:

   (i) If $\eta_{\max}/n < \eta_{\min}/(n-1)$ then $H^0 < \eta_{\max}/n$ *iff* $\Delta W\left(\frac{\eta_{\max}}{n}\right) > 0$, or equivalently,

$$f(\eta_{\max}) - f\left(\frac{n-1}{n}\eta_{\max}\right) > \frac{\eta_{\max}}{n}, \tag{29}$$

   which is the converse of condition (28).

   The biological intuition for this case is that if $\eta_{\max}/n < \eta_{\min}/(n-1)$, then when all non-focal agents contribute $H^* = \eta_{\max}/n$, their total contribution is less than the minimizing total good ($(n-1)H^* < \eta_{\min}$). Consequently, $W(h, H^*)$ decreases for $0 \leq h \leq \eta_{\min} - (n-1)H^*$, then increases for $\eta_{\min} - (n-1)H^* < h < \eta_{\max} - (n-1)H^*$ and decreases again for $h \geq \eta_{\max} - (n-1)H^*$. Thus, the two candidates for the best response for the focal agent are $h = 0$ and $h = \eta_{\max} - (n-1)H^* = H^*$, and H^* is an ESS *iff* $W(H^*, H^*) > W(0, H^*)$ (that is, $\Delta W(H^*) > 0$). Note that condition (29) stipulates that the mean slope of f on the interval $\left[\frac{n-1}{n}\eta_{\max}, \eta_{\max}\right]$ is larger than 1.

(ii) If $\eta_{\min}/(n-1) \leq \eta_{\max}/n$, then $H^0 < \eta_{\max}/n$ is satisfied (because $H^0 < \eta_{\min}/n$). This is because if $\eta_{\min}/(n-1) \leq \eta_{\max}/n$, then when $H = H^* = \eta_{\max}/n$, the total non-focal agents' contribution exceeds the minimizing total good ($(n-1)H > \eta_{\min}$), so $W(h, H^*)$ is unimodal and has a unique global maximum (i.e., in the range $h \geq 0$) at

$$h = \eta_{\max} - (n-1)H^* = \frac{\eta_{\max}}{n} = H^*. \tag{30}$$

While the condition $\eta_{\min}/(n-1) \leq \eta_{\max}/n$ seems weaker than condition (29), note that $W(0, H^*) < W(H^*, H^*)$, so condition (29) must hold in this case as well.

We conclude that if $f(\eta_{\max}) < \eta_{\max}$, then $H = 0$ is an ESS, and additionally, $H^* = \eta_{\max}/n$ is an ESS *iff* condition (29) holds.

Also, note that for a fixed benefit function, $f(\eta)$, for sufficiently large group size n, $\eta_{\min}/(n-1) \leq \eta_{\max}/n$ *must* hold. Thus, all else being equal, larger groups are more likely to have a nonzero ESS contribution.

5.3 Convergent stability of the ESSs

Section 5.2 showed that unless both $f(\eta_{\max}) < \eta_{\max}$ and $f(\eta_{\max}) \leq f(\frac{n-1}{n}\eta_{\max}) + \eta_{\max}/n$, the contribution level $H^* = \eta_{\max}/n$ is ES. We now show that when $H^* = \eta_{\max}/n$ is an ESS, it is also convergently stable, that is:

Suppose that all members of the population contribute H, and that a mutant playing $h \neq H$ enters the population. Recalling that we assume an infinite population, the mean resident fitness is unaffected by the mutant, and is simply $W(H, H)$. Thus, we wish to show that if H is sufficiently close to H^*, then for any invading strategy h that is between H and H^*, $W(h, H) > W(H, H)$.

Suppose that the resident strategy is $H = H^* - \mu$, where $\mu > 0$, and that the mutant strategy satisfies $H^* - \mu = H < h < H^*$. If $\mu < (\eta_{\max} - \eta_{\min})/n$ then $\eta_{\min} < nH < h + (n-1)H\eta_{\max}$, so

$$\begin{aligned} W(h, H) - W(H, H) &= [f(h + (n-1)H) - h] - [f(nH)) - H] \\ &= [f(h + (n-1)H) - (h + (n-1)H)] \\ &\quad - [f(nH)) - nH] > 0 \end{aligned} \tag{31}$$

because $f(\eta) - \eta$ is increasing on $[\eta_{\min}, \eta_{\max}]$.

Now suppose that the resident strategy is $H = H^* + \mu$, where $\mu > 0$. Because $\eta_{\max} < h + (n-1)H < nH$ and $f(\eta) - \eta$ decreases for $\eta > \eta_{\max}$, we have

$$\begin{aligned} W(h,H) - W(H,H) &= [f\,(h + (n-1)\,H) - (h + (n-1)H)] \\ &\quad - [f\,(nH)) - nH] > 0, \end{aligned} \tag{32}$$

for *any* $\mu > 0$.

It follows that if all members of the group use a strategy H sufficiently near the non-zero equilibrium, H^*, then the fitness of a mutant strategy between H and H^* is larger than the mean resident fitness, so H^* is convergently stable.

We also saw in Sect. 5.2, that if $f(\eta_{\max}) < \eta_{\max}$ then $H = 0$ is ES. To see that it is also convergently stable, note that if $0 < H < \eta_{\min}/n$, then $\eta(h,H) \le \eta(H,H) < \eta_{\min}$ for all for $0 < h \le H$. Since $f(\eta) - \eta$ decreases with η for $\eta < \eta_{\min}$, and $\eta(h,H)$ increases with h, it follows that $W(h,H) = f(\eta(h,H)) - \eta(h,H) + (n-1)H$ decreases with h, which implies $H = 0$ is convergently stable.

6 Proof of Theorem 4.3

Following Geritz et al. (1998) and Doebeli et al. (2004), the growth rate of a rare mutant strategy h in a resident population playing H is

$$s_H(h) = W(h,H) - W(H,H) = f(h + (n-1)H) - f(nH) + H - h. \tag{33}$$

The local fitness gradient is then

$$D(H) = \left.\frac{\partial s_H(h)}{\partial h}\right|_{h=H} = f'(nH) - 1, \tag{34}$$

and the adaptive dynamics of H are given by

$$\dot{H} = D(H) = f'(nH) - 1. \tag{35}$$

Singular strategies satisfy $f'(nH) = 1$. Since by our assumptions, $f(\eta) - \eta$ increases when $\eta_{\min} < \eta < \eta_{\max}$ and decreases otherwise, the two singular strategies are $H = \eta_{\min}/n$ and $H = \eta_{\max}/n = H^*$. Because $\frac{d}{dH}D(H) = nf''(nH)$ and $f''(\eta_{\min}) > 0$, it follows that $H = \eta_{\min}/n$ is a repellor.

As for the singular strategy $H = H^*$, using Table 1 (adapted from Geritz et al. 1998) and letting

$$\begin{aligned} a &= \left.\frac{\partial^2 s_H(h)}{\partial^2 H}\right|_{h=H=H^*} = (n-1)^2 f''(nH^*) - n^2 f''(nH^*) \\ &= (1-2n) f''(\eta_{\max}) > 0, \end{aligned} \tag{36}$$

$$b = \left.\frac{\partial^2 s_H(h)}{\partial^2 h}\right|_{h=H=H^*} = f''(nH^*) = f''(\eta_{\max}) < 0, \tag{37}$$

(since $f''(\eta_{\max}) < 0$ and $n \geq 1$) we see that H^* is a convergently stable local-ESS (because $b < 0$ and $a > b$). Though these are the only properties necessary for theorem 4.3, for completeness, we also note that since $a > 0$, H^* can invade a homogeneous population playing a sufficiently similar strategy $H \neq H^*$. Lastly, if $n > 1$, then mutually-invasible strategies exist near H^* since $a + b > 0$ (however, dimorphic populations will tend to disappear as the population converges towards the ESS H^*, see Geritz et al. 1998, p. 42).

7 Proof of Theorem 4.4

7.1 Local ES

Proof of R1 (Local ES) Consider an infinite population playing H^* invaded by a proportion $\epsilon > 0$ of mutants playing $h \neq H^*$. We wish to compare the mean fitnesses of a resident playing H^* and a mutant playing h.

To obtain the mean mutant fitness, first note that the payoff to a mutant in a group with a total of k mutants is

$$W_{\mathrm{m},k}(h) = W\left(h, \frac{(k-1)h + (n-k)H^*}{n-1}\right). \tag{38}$$

We now calculate the proportion of mutants that are in a group containing k mutants. Choose an agent at random from the population by first choosing a group at random and then choosing an agent at random from within that group. Let I be an indicator for whether the chosen agent is a mutant ($I = 1$ if the chosen agent is a mutant, and $I = 0$ otherwise). Let M be the number of mutants in the chosen group. We use Bayes' Theorem (Ross 2010) to find $P(M = k|I = 1)$, that is, the probability that a chosen mutant is in a group containing k mutants:

$$P(M = k|I = 1) = \frac{P(M = k)P(I = 1|M = k)}{P(I = 1)}. \tag{39}$$

Because the population is assumed infinite, the probability that a randomly chosen group contains k mutants is binomially distributed with parameters n and ϵ,

$$P(M = k) = \binom{n}{k}\epsilon^k(1-\epsilon)^{n-k}. \tag{40}$$

The probability of drawing a mutant at random from a group containing k mutants is $P(I = 1|M = k) = k/n$. The probability that an individual chosen at random from the whole population is a mutant is $P(I = 1) = \epsilon$. Thus,

$$\begin{aligned} P(M = k|I = 1) &= \frac{\binom{n}{k}\epsilon^k(1-\epsilon)^{n-k}k/n}{\epsilon} \\ &= \binom{n-1}{k-1}\epsilon^{k-1}(1-\epsilon)^{n-k}. \end{aligned} \tag{41}$$

that is, the remaining number of mutants in the group is distributed binomially with parameters $n-1$ and ϵ.

It follows that the mean payoff for a mutant is:

$$\begin{aligned}\overline{W}_{\mathrm{m}}(h) &= \sum_{k=1}^{n} P(M=k|I=1)W_{\mathrm{m},k}(h)\\ &= \sum_{k=1}^{n}\binom{n-1}{k-1}\epsilon^{k-1}(1-\epsilon)^{n-k}W_{\mathrm{m},k}(h)\\ &= \sum_{k=0}^{n-1}\binom{n-1}{k}\epsilon^{k}(1-\epsilon)^{n-1-k}W_{\mathrm{m},k+1}(h)\end{aligned}$$

Similarly, the probability that a randomly chosen resident's group contains k mutants is

$$P(M=k|I=0)=\binom{n-1}{k}\epsilon^{k}(1-\epsilon)^{n-1-k}, \tag{42}$$

and the payoff to a resident in a group containing k mutants is (Eq. (5))

$$W_{\mathrm{r},k}=W\left(H^*,\frac{kh+(n-1-k)H^*}{n-1}\right)=f\left(kh+(n-k)H^*\right)-H^*. \tag{43}$$

So, the mean payoff to an agent playing the resident strategy H^* is

$$\begin{aligned}\overline{W}_{\mathrm{r}}(h) &= \sum_{k=0}^{n-1} P(M=k|I=0)W_{\mathrm{r},k}(h)\\ &= \sum_{k=0}^{n-1}\binom{n-1}{k}\epsilon^{k}(1-\epsilon)^{n-1-k}W_{\mathrm{r},k}(h).\end{aligned}$$

The difference between the mean fitnesses of the mutant and resident strategies is then

$$\begin{aligned}\delta\overline{W}(h) &= \overline{W}_{\mathrm{m}}(h)-\overline{W}_{\mathrm{r}}(h)\\ &= \sum_{k=0}^{n-1}\binom{n-1}{k}\epsilon^{k}(1-\epsilon)^{n-1-k}\Big[W_{\mathrm{m},k+1}(h)-W_{\mathrm{r},k}(h)\Big].\end{aligned} \tag{44}$$

Denoting the total contribution in a group containing k mutants and $n-k$ residents by

$$\eta_k = kh+(n-k)H^*, \tag{45}$$

and noting that $\eta_{k+1} - \eta_k = h - H^*$, we have

$$\begin{aligned} W_{\mathrm{m},k+1}(h) - W_{\mathrm{r},k}(h) &= \big[f(\eta_{k+1}) - h\big] - \big[f(\eta_k) - H^*\big] \\ &= \big[f(\eta_{k+1}) - \eta_{k+1}\big] - \big[f(\eta_k) - \eta_{k+1}\big]. \end{aligned} \tag{46}$$

If $\frac{\eta_{\min}}{n} < h < H^*$ then, for all $0 \le k \le n-1$,

$$\eta_{\min} < nh = \eta_n \le \eta_{k+1} < \eta_k \le \eta_0 = nH^* = \eta_{\max}, \tag{47}$$

so because $f(\eta) - \eta$ is increasing for $\eta_{\min} < \eta < \eta_{\max}$,

$$W_{\mathrm{m},k+1}(h) - W_{\mathrm{r},k}(h) = \big[f(\eta_{k+1}) - \eta_{k+1}\big] - \big[f(\eta_k) - \eta_{k+1}\big] < 0, \tag{48}$$

that is, each term in the sum (44) above is negative, implying $\delta\overline{W}(h) < 0$.

Similarly, if $h > H^*$ then, for all $0 \le k \le n-1$,

$$\eta_{\max} = nH^* < \eta_k < \eta_{k+1}, \tag{49}$$

and since $f(\eta) - \eta$ is decreasing for $\eta > \eta_{\max}$, inequality (48) holds again, implying $\delta\overline{W}(h) < 0$.

Thus, a mutant strategy sufficiently close to the equilibrium H^* cannot invade, regardless of its initial proportion in the population, ϵ. □

7.2 Local CS

Proof of R2 (Local CS) Similar to the derivation of Eq. (44) in Sect. 7.1, the mean fitness difference between a mutant contributing h and a resident contributing H is

$$\begin{aligned} \delta\overline{W}(h) &= \sum_{k=0}^{n-1} \binom{n-1}{k} \epsilon^k (1-\epsilon)^{n-1-k} \Bigg[W\left(h, \frac{kh + (n-1-k)H}{n-1}\right) \\ &\qquad - W\left(H, \frac{kh + (n-1-k)H}{n-1}\right)\Bigg] \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} \epsilon^k (1-\epsilon)^{n-1-k} \Big\{ \big[f\big((k+1)h + (n-1-k)H\big) - h\big] \\ &\qquad - \big[f\big(kh + (n-k)H\big) - H\big]\Big\} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} \epsilon^k (1-\epsilon)^{n-1-k} \\ &\quad \times \Big\{ \big[f\big((k+1)h + (n-1-k)H\big) - \big((k+1)h + (n-1-k)H\big)\big] \\ &\qquad - \big[f\big(kh + (n-k)H\big) - \big(kh + (n-k)H\big)\big]\Big\}. \end{aligned} \tag{50}$$

If $\eta_{\min}/n < H < h < H^*$, then $\eta_{\min} < kh + (n-k)H < \eta_{\max}$ for all $0 \le k \le n$, and because $f(\eta) - \eta$ is increasing on $[\eta_{\min}, \eta_{\max}]$, each term in the sum in the last line of equation (50) is positive, so $\delta\overline{W}(h) > 0$. Conversely, if $H^* < h < H$, then because $\eta_{\max} < kh + (n-k)H$ for all $0 \le k \le n$ and $f(\eta) - \eta$ decreases for $\eta > \eta_{\max}$, again, $\delta\overline{W}(h) > 0$. Thus, the ESS H^* is convergently stable. □

8 Discussion

We have analyzed the general class of public goods games described in Sect. 2 continuous n-player snowdrift games, Doebeli et al. 2004 using the two frameworks summarized in Sect. 3 [static evolutionary game theory (Maynard Smith 1982; Hofbauer and Sigmund 1998; Nowak 2006) and adaptive dynamics (Dieckmann and Law 1996; Geritz et al. 1998; Metz et al. 1996; Hofbauer and Sigmund 1998)]. Our results are expressed in three theorems stated in Sect. 4 and proved in Sects. 5, 6 and 7.

With the standard static theory, we identified two candidate evolutionarily stable strategies (ESSs): either contributing nothing ($H = 0$, “defection”) or contributing an equal share of the maximizing total good ($H = H^* = \eta_{\max}/n$, “cooperation”). Defection is an ESS unless the benefit of contributing to the public good is so large that it is worth doing so even if no-one else contributes. Cooperation is an ESS unless the cost of contributing the maximizing total good *single-handedly* outweighs its benefit and the incremental cost of contributing an equal share also exceeds its benefit (condition (14)). When they exist, each ESS is resistant to invasion by a mutant that contributes *any* other amount (globally evolutionary stable) and is selected for in populations of individuals contributing nearly the ESS level (locally convergently stable).

Our conclusions do not depend on the form of the benefit function $f(h)$ beyond the biologically sensible hypotheses H1 and H2, and are hence applicable to a wide range of biological, social and economic problems. Moreover, the conditions we find under which cooperation is inherently evolutionarily stable are independent of any external mechanism such as population structure (Reeve and Hölldobler 2007; Barker et al. 2012; Krupp and Taylor 2015), kin selection (Cornforth et al. 2012; Krupp et al. 2008; Krupp and Taylor 2015; Clutton-Brock 2002), reciprocity (Killingback and Doebeli 2002; Nowak and Sigmund 1998; Trivers 1971) or partner selection (McNamara et al. 2008).

With the adaptive dynamics framework, we found only one possible ESS, which is to “cooperate” by contributing an equal share of the maximizing total good, as in the static theory. Unlike the static theory, the standard formulation of adaptive dynamics requires a smooth fitness function (hypothesis H3) and the typical definition of evolutionary stability in the adaptive dynamics literature is *local* (e.g., Geritz et al. 1998), so the conditions for stability can be weaker, which is what our analysis revealed for the public goods games that we considered: cooperation is *locally* evolutionary and convergently stable *no matter what*. Note also that the notion of global ES is more relevant than the local ES in cases when the deviation of mutant strategies from the resident one are not small, e.g., in the case of flexible decision-making (rather than genetically predetermined behaviour) (Ito et al. 2015).

The reason that our adaptive dynamics analysis did not detect contributing nothing ("defection") as an ESS is an artifact of the analysis method's focus on evolutionarily singular points (see Appendix 2). Numerical simulations based on adaptive dynamics (e.g., those presented in Doebeli et al. 2004 and Killingback et al. 2010) are not subject to this constraint.

Compared with the static theory, the adaptive dynamics framework has the advantage of being able to describe evolutionary *dynamics*, even far from ESSs. However, studying the dynamics is possible only if a particular fitness function is adopted. In this paper, we focused on a general setting, without restricting attention to a particular benefit function $f(\eta)$, so that the inferences we make are as broad as possible.

With the adaptive dynamics framework, it is possible to describe and investigate the evolution of a dimorphic population (as opposed to a single mutant in an otherwise uniform resident population) but, again, only if a specific fitness function is chosen. Our third theorem (Theorem 4.4), based on static theory, successfully considers manifestly dimorphic populations in order to broaden the scope of the stability results to include potential effects of invasion by a significant proportion of mutants (which is applicable to a number of biological situations). We find that cooperation is locally evolutionarily and convergently stable in a much stronger sense than typically considered: when it is stable to invasion by a single mutant, H^* is actually selected for no matter how large a proportion of the population is playing the mutant strategy (if the mutants play a strategy that is sufficiently similar to the residents' strategy).

Throughout this paper, we have retained the standard assumption that the underlying population is infinite. An infinite population size is often justified in the adaptive dynamics literature on the grounds that small populations are unlikely to persist due to demographic stochasticity (Metz et al. 1996, §2.1). Of course, evolutionary stability predictions might differ in finite populations (Nowak 2006), a possibility that we will explore in further work.

Other possible complications that are not accounted for in our present analyses are the effects of a structured (i.e., not well-mixed) population (Hauert and Doebeli 2004; Boyd and Richerson 2002), relatedness among some or all agents in the population (Krupp et al. 2008; Cornforth et al. 2012), assymetry or variability among individuals, due to differences in age, sex, resources, abilities or costs (Frank 2010; Lotem et al. 1999; Houston and Davies 1985; McNamara et al. 1999), as well as inter- and intra-group competition affecting the division of resources (Reeve and Hölldobler 2007; Barker et al. 2012). Furthermore, in the class of games we have studied, individual agents choose their level of contribution independently and without knowledge of other agents' contributions. However, it is possible that agents choose their contributions in sequence, or negotiate their levels of contributions (Wright and Cuthill 1990; Markman et al. 1995; McNamara et al. 1999).

Acknowledgements We are grateful to Sigal Balshine, Pat Barclay, Ben Bolker, Jonathan Dushoff, Paul Higgs, Rufus Johnstone and Danny Krupp for valuable discussions and comments.

Compliance with ethical standards

Funding We were supported by NSERC (DE), the Ontario Trillium Foundation (CM) and the department of Mathematics and Statistics at McMaster University (CM).

Conflict of interest The authors declare that they have no conflicts of interest.

Appendix 1: Motivation for assumption A3 (existence of $\eta_{\max}$)

In this appendix, we motivate assumption A3 by showing that if the focal agent's fitness is defined by Eq. (5) and f is continuous, the following two statements are equivalent:

S1 For any fixed non-focal agent's mean contribution H there exists $h^\dagger(H) \geq 0$ such that the focal agent's fitness $W(h, H)$ decreases with its contribution h for all $h > h^\dagger(H)$.

S2 There exists $\eta_{\max} \geq 0$ such that $f(\eta) - \eta$ decreases with η for all $\eta > \eta_{\max}$.

To gain some intuition, we first suppose f is a differentiable function of η (in Appendix 1.1), and then give the general proof (in Appendix 1.2)

Appendix 1.1: Proof for differentiable f

Suppose that f is differentiable. Then, by the chain rule and Eq. (3), S1 implies that

$$\begin{aligned}\frac{\partial W}{\partial h} &= \left(f'(\eta) \frac{\partial \eta}{\partial h} \right)\Big|_{\eta=\eta(h,H)} - 1 = f'(\eta)\big|_{\eta=\eta(h,H)} - 1 \\ &= \frac{\mathrm{d}}{\mathrm{d}\eta}\Big(f(\eta) - \eta \Big)\Big|_{\eta=\eta(h,H)} . \end{aligned} \tag{51}$$

Consequently, if $W(h, H)$ decreases with h for $h > h^\dagger(H)$ then $f(\eta) - \eta$ decreases with η for all $\eta > \eta(h^\dagger(H), H)$. Letting

$$\eta_{\max} = \min_{H \geq 0} \eta(h^\dagger(H), H), \tag{52}$$

$f(\eta) - \eta$ decreases for $\eta > \eta_{\max}$. Thus, S1 implies S2.

Conversely, if there exists $\eta_{\max} \geq 0$ such that $f(\eta) - \eta$ decreases for $\eta > \eta_{\max}$, then letting $h^\dagger(H) = \eta_{\max} - (n-1)H$, we see that $\eta(h, H) > \eta_{\max}$ *iff* $h > h^\dagger(H)$. It then follows from Eq. (51) that $W(h, H)$ decreases with h for $h > h^\dagger(H)$, so S2 implies S1.

Appendix 1.2: Proof for general f

Suppose S1 holds. Noting that

$$\begin{aligned} h > h^{\dagger}(H) &\iff h + (n-1)H > h^{\dagger}(H) + (n-1)H \\ &\iff \eta(h, H) > \eta^{\dagger}(H), \end{aligned} \tag{53}$$

where

$$\eta^{\dagger}(H) = \eta\big(h^{\dagger}(H), H\big), \tag{54}$$

and rewriting Eq. (5) as

$$\begin{aligned} W(h, H) &= f\big(h + (n-1)\,H\big) - h \\ &= f\big(\eta\,(h, H)\big) - \big[\eta\big(h, H\big) - (n-1)H\big] \\ &= f(\eta) - \eta + (n-1)H, \end{aligned} \tag{55}$$

we see that S1 is equivalent to the assumption that for fixed H there exists $\eta^{\dagger}(H) \geq 0$ such that $f(\eta) - \eta$ decreases for $\eta > \eta^{\dagger}(H)$. We define $\eta_{\max}$ to be the minimal such (non-negative) total good. Because η can vary independently of H, it follows that $f(\eta) - \eta$ decreases for all $\eta > \eta_{\max}$, so S1 implies S2.

Conversely, if S2 is true, then Eqs. (53) and (55) imply that $W(h, H)$ decreases for all $h > h^{\dagger}(H) = \eta_{\max} - (n-1)H$. Thus, S2 implies S1.

Appendix 2: Boundary ESSs need not be singular strategies

In adaptive dynamics, evolutionarily singular points are singled out as candidate ESSs (e.g., Geritz et al. 1998; Doebeli et al. 2004). These are points at which there is no directional selection, since the fitness gradient $D(H)$ vanishes.

However, when the evolving variable H is restricted to an interval (in our case $H \geq 0$), it is not necessary for the fitness gradient to vanish at an endpoint of this interval in order for it to be ES: as we have seen in Theorem 4.1, for the class of models defined in Sect. 2, the endpoint $H = 0$ is globally evolutionarily stable whenever $f(\eta_{\max}) < \eta_{\max}$, but the fitness gradient is negative in a right-hand neighbourhood of the endpoint $H = 0$ (including at $H = 0$). In fact, it is $D(H)$ being negative near $H = 0$ that ensures that $H = 0$ is both locally convergently and evolutionarily stable.

The source of this issue is that the restriction to the biologically meaningful interval $H \geq 0$ is not built into the dynamical model Eq. (12), in that solutions of Eq. (12) do not necessarily remain in this interval (because the fitness gradient at the left endpoint $H = 0$ points outside the interval, into $H < 0$).

Note also that this cannot be easily fixed by artificially setting $D(0) = 0$, because doing so will insert a discontinuity into the fitness gradient, and adaptive dynamics assumes that the fitness gradient is at least continuous, in order to ensure the existence of solutions of Eq. (12) (see Hirsch et al. 2013) and in order to perform the local analysis leading to Table 1.

We conclude that when using adaptive dynamics to model a trait that is restricted to an interval (for biological reasons), points on the boundary of this interval may be ES, despite not being singular points. More care is thus required to examine the dynamics of such models near boundary points.

Appendix 3: The assumption that contribution is measured in units of fitness cost, $c(h) = h$

In this appendix, we comment on the biological interpretation of our assumption that the contribution of the focal agent is measured in units of the fitness cost it incurs, $c\left(h, H\right) = h$ (Eq. (2)).

Suppose, as before, that the population is engaged in an n-player public goods game, and let $h_1, \ldots, h_n$ be the contributions of all the members of the focal agent's group, including the focal agent (for example, if the index of the focal agent is $i = 1$, then $h = h_1$).

Thus, substituting

$$\eta(h, H) = \sum_{i=1}^{n} h_i \tag{56}$$

in Eq. (4), we have

$$b\left(h, H\right) = f\big(\eta(h, H)\big) = f\left(\sum_{i=1}^{n} h_i\right). \tag{57}$$

However, we relax our assumption in Eq. (2) and instead only assume that

$$c(h, H) = c(h), \tag{58}$$

so that the fitness cost incurred by the focal agent is independent of the contributions of the other members in its group.

The focal agent's fitness is then

$$W(h, H) = f\left(\sum_{i=1}^{n} h_i\right) - c(h). \tag{59}$$

For $1 \leq i \leq n$, let $C_i = c(h_i)$, and $C = c(h)$ be the costs incurred by the n members of the focal agent's group, and the focal agent (respectively). Suppose that the cost function is one-to-one, so that there exists a left-inverse function $k(\cdot)$ satisfying $k\big(c(h)\big) = h$ and $k\big(c(h_i)\big) = h_i$ for all $1 \leq i \leq n$. Then, Eq. (59) becomes

$$W(h, H) = f\left(\sum_{i=1}^{n} k(C_i)\right) - C. \tag{60}$$

The benefit to the focal agent, $f\left(\sum_{i=1}^{n} k(C_i)\right)$ is then generally not a function of the sum of the group members' fitness costs, $\sum_{i=1}^{n} C_i$.

By assuming that contributions to the public good are expressed in units of fitness cost (i.e., $c(h) = h$, as in Eq. (2)), we implicitly assumed that fitness itself is the public good. Expressed in more biological terms, we are assuming that reproductive costs are effectively transferable: each individual in a group obtains a fitness benefit $f(\eta)$ regardless of how the associated costs (which sum to η) are distributed among the group members; for example, the fitness benefit is the same if the focal agent contributes the entire cost ($h = \eta$), or if the cost is distributed equally among group members ($h_i = \eta/n$ for each i).

References

Archetti M (2013) Evolutionary game theory of growth factor production: implications for tumour heterogeneity and resistance to therapies. Br J Cancer 109(4):1056–1062

Archetti M (2014) Evolutionary dynamics of the Warburg effect: glycolysis as a collective action problem among cancer cells. J Theor Biol 341:1–8

Archetti M, Ferraro DA, Christofori G (2015) Heterogeneity for IGF-II production maintained by public goods dynamics in neuroendocrine pancreatic cancer. Proc Natl Acad Sci 112(6):1833–1838

Archetti M, Scheuring I (2011) Coexistence of cooperation and defection in public goods games. Evolution 65(4):1140–1148

Archetti M, Scheuring I (2012) Review: game theory of public goods in one-shot social dilemmas without assortment. J Theor Biol 299:9–20

Bach LA, Bentzen S, Alsner J, Christiansen FB (2001) An evolutionary-game model of tumour-cell interactions: possible relevance to gene therapy. Euro J Cancer 37(16):2116–2120

Barker JL, Barclay P, Reeve HK (2012) Within-group competition reduces cooperation and payoffs in human groups. Behav Ecol 23(4):735–741

Box GE (1976) Science and statistics. J Am Stat Assoc 71(356):791–799

Boyd R, Richerson PJ (2002) Group beneficial norms can spread rapidly in a structured population. J Theor Biol 215(3):287–296

Brännström Å, Gross T, Blasius B, Dieckmann U (2010) Consequences of fluctuating group size for the evolution of cooperation. J Math Biol 63(2):263–281. Available from doi: 10.1007/s00285-010-0367-3

Brown JS, Vincent TL (2008) Evolution of cooperation with shared costs and benefits. Proc R Soc B 275(1646):1985–1994. Available from doi:10.1098/rspb.2007.1685

Brown S (1999) Cooperation and conflict in host-manipulating parasites. Proc R Soc Lond B Biol Sci 266(1431):1899–1904

Brown SP, Hochberg ME, Grenfell BT (2002) Does multiple infection select for raised virulence? Trends Microbiol 10(9):401–405

Browning M, Chiappori PA, Weiss Y (2014) Economics of the family. Cambridge University Press, Cambridge

Buma B, Wessman C (2011) Disturbance interactions can impact resilience mechanisms of forests. Ecosphere 2(5):art64

Chen X, Szolnoki A, Perc M, Wang L 2012 Impact of generalized benefit functions on the evolution of cooperation in spatial public goods games with continuous strategies. Phys Rev E 85(6). Available from doi:10.1103/PhysRevE.85.066133

Christiansen FB (1991) On conditions for evolutionary stability for a continuously varying character. Am Nat 138(1):37–50

Clutton-Brock T (2002) Breeding together: kin selection and mutualism in cooperative vertebrates. Science 296(5565):69–72

Clutton-Brock T (2009) Cooperation between non-kin in animal societies. Nature 462(7269):51–57

Cordero OX, Ventouras LA, DeLong EF, Polz MF (2012) Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. Proc Natl Acad Sci 109(49):20059–20064

Cornforth DM, Sumpter DJ, Brown SP, Brännström Å (2012) Synergy and group size in microbial cooperation. Am Nat 180(3):296

Dale VH, Joyce LA, McNulty S, Neilson RP, Ayres MP, Flannigan MD et al (2001) Climate change and forest disturbances: climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. BioScience 51(9):723–734

Deng K, Chu T (2011) Adaptive evolution of cooperation through Darwinian dynamics in public goods games. PLoS One 6(10):e25496. Available from doi:10.1371/journal.pone.0025496

Dieckmann U, Law R (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. J Math Biol 34(5–6):579–612

Doebeli M, Hauert C, Killingback T (2004) The evolutionary origin of cooperators and defectors. Science 306(5697):859–862

Dugatkin LA (1997) Cooperation among animals: an evolutionary perspective: an evolutionary perspective. Oxford University Press, USA

Fehr E, Gächter S (1999) Cooperation and punishment in public goods experiments. Institute for empirical research in economics working paper (10)

Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415(6868):137–140

Frank SA (2010) A general model of the public goods dilemma. J Evol Biol 23(6):1245–1250

Fudenberg D, Tirole J (1991) Game Theory. MIT Press, Cambridge

Fujita H, Aoki S, Kawaguchi M (2014) Evolutionary dynamics of nitrogen fixation in the Legume-Rhizobia symbiosis. PLoS One 9(4):e93670. Available from doi:10.1371/journal.pone.0093670

Gardner A, West S (2006) Demography, altruism, and the benefits of budding. J Evol Biol. 19(5):1707–1716

Gavrilets S (2015) Collective action problem in heterogeneous groups. Philos Trans R Soc B 370(1683):20150016

Geritz SA, Mesze G, Metz JA et al (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evol Ecol 12(1):35–57

Gokhale CS, Traulsen A (2014) Evolutionary multiplayer games. Dyn Games Appl 4(4):468–488. Available from doi:10.1007/s13235-014-0106-2

Hardin G (1968) The tragedy of the commons. Science 162(3859):1243–1248

Hartl DL, Clark AG (2007) Principles of population genetics. Sinauer Associates, Sunderland

Hauert C (2006) Cooperation, collectives formation and specialization. Adv Complex Syst 9(04):315–335

Hauert C (2008) Evolution from cellular to social scales. In: Skjeltorp AT, Belushkin AV (eds) Evolutionary dynamics. Springer, Berlin, pp 11–44

Hauert C, Doebeli M (2004) Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428(6983):643–646

Hauert C, Holmes M, Doebeli M (2006) Evolutionary games and population dynamics: maintenance of cooperation in public goods games. Proc R Soc B 273(1600):2565–2571. Available from doi: 10.1098/rspb.2006.3600

Hauert C, Michor F, Nowak MA, Doebeli M (2006) Synergy and discounting of cooperation in social dilemmas. J Theor Biol 239(2):195–202

Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Replicator dynamics for optional public good games. J Theor Biol 218(2):187–194

Hirsch MW, Smale S, Devaney RL (2013) Differential equations, dynamical systems, and an introduction to chaos, 3rd edn. Academic Press, Waltham

Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge

Houston AI, Davies NB (1985) The evolution of cooperation and life history in the Dunnock, Prunella modularis. In: Sibly RM, Smith RH (eds) Behavioural ecology: ecological consequences of adaptive behaviour. Blackwell Scientific Publications, Oxford, pp 471–487

Ito K, Ohtsuki H, Yamauchi A (2015) Relationship between aggregation of rewards and the possibility of polymorphism in continuous snowdrift games. J Theor Biol 372:47–53. Available from doi:10.1016/j.jtbi.2015.02.015

Kagel JH, Roth AE (1995) The handbook of experimental economics. Princeton University Press, Princeton

Killingback T, Doebeli M (2002) The continuous prisoners dilemma and the evolution of cooperation through reciprocal altruism with variable investment. Am Nat 160(4):421–438

Killingback T, Doebeli M, Hauert C (2010) Diversity of cooperation in the tragedy of the commons. Biol Theory 5:3–6

Kraak S (2011) Exploring the public goods game model to overcome the tragedy of the commons in fisheries management. Fish Fish 12(1):18–33

Krupp DB, Debruine LM, Barclay P (2008) A cue of kinship promotes cooperation for the public good. Evol Human Behav 29(1):49–55

Krupp D, Taylor PD (2015) Social evolution in the shadow of asymmetrical relatedness. Proc R Soc Lond B Biol Sci 282(1807):20150142

Kümmerli R, Gardner A, West SA, Griffin AS (2009) Limited dispersal, budding dispersal, and cooperation: an experimental study. Evolution 63(4):939–949

Leighton GM (2014) Sex and individual differences in cooperative nest construction of sociable weavers Philetairus socius. J Ornithol 155(4):927–935

Liang H, Cao M, Wang X (2015) Analysis and shifting of stochastically stable equilibria for evolutionary snowdrift games. Syst Control Lett 85:16–22. Available from doi:10.1016/j.sysconle.2015.08.004

Lotem A, Fishman MA, Stone L (1999) Evolution of cooperation between individuals. Nature 400(6741):226–227

Markman S, Yom-Tov Y, Wright J (1995) Male parental care in the orange-tufted sunbird: behavioural adjustments in provisioning and nest guarding effort. Anim Behav 50(3):655–669

Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge

Maynard Smith J, Price G (1973) lhe logic of animal conflict. Nature 246:15

McGill BJ, Brown JS (2007) Evolutionary game theory and adaptive dynamics of continuous traits. Ann Rev Ecol Syst 38(1):403–435. Available from doi:10.1146/annurev.ecolsys.36.091704.175517

McNamara JM, Barta Z, Fromhage L, Houston AI (2008) The coevolution of choosiness and cooperation. Nature 451(7175):189–192. Available from doi:10.1038/nature06455

McNamara JM, Gasson CE, Houston AI (1999) Incorporating rules for responding into evolutionary games. Nature 401(6751):368–371

Metz JA, Geritz SA, Meszéna G, Jacobs FJ, Van Heerwaarden JS et al (1996) Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. Stoch Spat Struct Dyn Syst 45:183–231

Milinski M, Semmann D, Krambeck HJ (2002) Reputation helps solve the tragedy of the commons. Nature 415(6870):424–426

Motro U (1991) Co-operation and defection: playing the field and the ESS. J Theor Biol 151(2):145–154

Motro U, Cohen D (1989) A note on vigilance behavior and stability against recognizable social parasites. J Theor Biol 136(1):21–25

Nowak MA (2006) Evolutionary dynamics: exploring the equations of life. Harvard University Press, Massachusetts

Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. Nature 393(6685):573–577

Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. Science 303(5659):793–799

Parker GA, Royle NJ, Hartley IR (2002) Intrafamilial conflict and parental investment: a synthesis. Philos Trans R Soc Lond B Biol Sci 357(1419):295–307

Pfeiffer T, Bonhoeffer S (2003) An evolutionary scenario for the transition to undifferentiated multicellularity. Proc Natl Acad Sci 100(3):1095–1098

Poulin R (1994) The evolution of parasite manipulation of host behaviour: a theoretical analysis. Parasitology 109(S1):S109–S118

Pulliam HR, Pyke GH, Caraco T (1982) The scanning behavior of juncos: a game-theoretical approach. J Theor Biol 95(1):89–103

Rainey PB, Rainey K (2003) Evolution of cooperation and conflict in experimental bacterial populations. Nature 425(6953):72–74

Reeve HK, Hölldobler B (2007) The emergence of a superorganism through intergroup competition. Proc Natl Acad Sci 104(23):9736–9740

Ross SM (2010) A First Course in Probability. Pearson Prentice Hall, Upper Saddle River

Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. Nature 454(7201):213–216. Available from doi:10.1038/nature06940

Sasaki T, Okada I (2015) Cheating is evolutionarily assimilated with cooperation in the continuous snowdrift game. Biosystems 131:51–59. Available from doi:10.1016/j.biosystems.2015.04.002

Savage JL, Russell AF, Johnstone RA (2015) Maternal allocation in cooperative breeders: should mothers match or compensate for expected helper contributions? Anim Behav 102:189–197

Souza MO, Pacheco JM, Santos FC (2009) Evolution of cooperation under N-person snowdrift games. J Theor Biol 260(4):581–588

Taylor PD (1989) Evolutionary stability in one-parameter models under weak selection. Theor Popul Biol 36(2):125–143

Trivers RL (1971) The evolution of reciprocal altruism. Q Rev Biol 246(1):35–57

Wright J, Cuthill I (1990) Biparental care: short-term manipulation of partner contribution and brood size in the starling, *Sturnus vulgaris*. Behav Ecol 1(2):116–124

Wu T, Fu F, Wang L (2009) Partner selections in public goods games with constant group size. Phys Rev E 80(2). Available from doi:10.1103/PhysRevE.80.026121