

# Evolution of Helping and Harming in Viscous Populations When Group Size Varies

António M. M. Rodrigues<sup>1,\*</sup> and Andy Gardner<sup>1,2</sup>

1. Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom; 2. Balliol College, University of Oxford, Broad Street, Oxford OX1 3BJ, United Kingdom

Submitted July 3, 2012; Accepted January 25, 2013; Electronically published March 14, 2013

Online enhancement: appendix.

**ABSTRACT:** Recent years have seen huge interest in understanding how demographic factors mediate the evolution of social behavior in viscous populations. Here we study the impact of variation in group size on the evolution of helping and harming behavior. Although variation in group size influences the degree of relatedness and the degree of competition between groupmates, we find that these effects often exactly cancel, so as to give no net impact of variation in group size on the evolution of helping and harming. Specifically, (1) obligate helping and harming are never mediated by variation in group size, (2) facultative helping and harming are not mediated by variation in group size when this variation is spatial only, (3) facultative helping and harming are mediated by variation in group size only when this variation is temporal or both spatial and temporal, and (4) when there is an effect of variation in group size, facultative helping is favored in big groups and facultative harming is favored in little groups. Moreover, we find that spatial and temporal heterogeneity in individual fecundity may interact with patch-size heterogeneity to change these predictions, promoting the evolution of harming in big patches and of helping in little patches.

**Keywords:** class structure, conditional behavior, inclusive fitness, kin selection, phenotypic plasticity, reproductive value.

## Introduction

Explaining cooperative behavior has been a central challenge to evolutionary biologists over the past few decades (Maynard Smith and Szathmáry 1995; Hamilton 1996; West et al. 2007a; Bourke 2011). Natural selection favors those individuals who achieve higher relative fitness (Darwin 1859; Fisher 1930; Price 1970). But cooperative behaviors that improve the fitness of others will, all else being equal, decrease the relative fitness of the actor. This problem can be solved if the actor receives a direct fitness benefit as a consequence of her cooperative behavior. For example, the immediate cost of the actor's behavior may

be compensated for later in her life by reciprocation from others (Trivers 1971). Such cooperation is termed “mutually beneficial” (West et al. 2007b). Alternatively, the problem can be solved if the actor receives an indirect fitness benefit, through her genetic relatives benefiting from her cooperative behavior (Hamilton 1963, 1964). Such cooperation is termed “altruistic” (Hamilton 1964; West et al. 2007b).

Altruistic cooperation requires that interacting individuals have genes in common. Hamilton (1964, 1971b) suggested three general mechanisms by which this can occur. First, individuals might be able to identify which of their social partners are genealogically close kin and preferentially direct their altruism to these individuals. Second, a “greenbeard” gene may provide its bearer the ability to directly identify which of their social partners also carry the same gene, independently of their genealogical relationship (Hamilton 1964; Dawkins 1976; Gardner and West 2010). Third, if dispersal is limited, then neighboring individuals are likely to be genealogically close kin, such that even indiscriminate cooperation will tend to occur between individuals who have genes in common. As this third mechanism does not require any discrimination mechanism, it could represent a very general principle that explains the evolution of cooperation at all levels of biological organization (Maynard Smith and Szathmáry 1995; Bourke 2011).

However, limited dispersal can also lead to intensified competition for resources among kin (Hamilton 1964; Queller 1992; West et al. 2002). In the simplest scenario of an infinite island population, the kin-competition effects of limited dispersal exactly cancel its relatedness-enhancing effects, such that population viscosity has no net impact on the evolution of indiscriminate cooperation (Taylor 1992a). This striking result has stimulated the development of a large body of theoretical—and, to a lesser extent, empirical—research examining what additional factors may decouple the relatedness and competition ef-

\* Corresponding author; e-mail: antonio.rodrigues@zoo.ox.ac.uk.

fects of limited dispersal, so that indiscriminate helping may be favored in viscous populations. Some of these factors include population elasticity (Taylor 1992*b*), overlapping generations (Taylor and Irwin 2000; Irwin and Taylor 2001), budding dispersal (Gardner and West 2006; Lehmann et al. 2006*b*; Kümmerli et al. 2009), an organism's life cycle and timing of social behavior (Taylor 1992*a*; Lehmann and Rousset 2010), behaviors mediating patch-extinction probabilities (Lehmann et al. 2006*b*), trans-generational altruism (Lehmann 2007, 2010), dispersal-dependent social behavior (El Mouden and Gardner 2008), empty sites (Alizon and Taylor 2008), sex-biased dispersal (Johnstone and Cant 2008; Gardner 2010), reproductive skew (Johnstone 2008), age structure (Johnstone and Cant 2010), and heterogeneity in resource availability (Rodrigues and Gardner 2012).

All this work has concerned populations in which the size of viable groups is fixed. However, natural populations commonly exhibit some degree of group-size heterogeneity, and the impact of this variation has been of great interest to ecologists and evolutionary biologists (Levins 1968; McPeck and Holt 1992; Begon et al. 2006). This is important to the evolution of kin-selected traits because patch size is likely to have an impact not only on the relatedness among social partners but also on local competition, two important factors that jointly mediate the evolution of social traits in viscous populations.

Here we investigate the impact of heterogeneity in group size on the evolution of indiscriminate helping and harming in viscous populations. Specifically, we consider scenarios in which different groups have different sizes at any given time (spatial heterogeneity) and/or the same group has different sizes at different times (temporal heterogeneity), where group size is extrinsically controlled and not driven by the group's past or present social behavior. We first consider cases where fecundity is equal across all patch types, and we then relax this assumption by considering cases where fecundity depends on patch type. Moreover, we consider scenarios where individuals obligately express their social behavior independently of their patch type and scenarios where individuals may facultatively adjust their social behavior according to their patch type. We determine the impact of this heterogeneity on the genetic relatedness and intensity of competition between group-mates and examine how these interact and affect the evolution of helping and harming behaviors.

## Model and Analysis

### *Model*

We assume an infinite island population (Wright 1931) with two types of patches: big patches with  $n_B$  breeding

sites and little patches with  $n_L$  breeding sites (without loss of generality, we assume  $n_B \geq n_L$ ). We consider that individuals are haploid and asexual, that each individual breeding in a big patch has a very high fecundity  $F_B = f(x, y)$ , and that each individual breeding in a little patch has a very high fecundity  $F_L = \sigma f(x, y)$ , where  $x$  is the individual's investment in the social behavior,  $y$  is the average level of investment among the individual's group-mates, and  $0 < \sigma < \infty$  is the "reproductive factor" describing fecundity differences between individuals breeding in little versus big patches. Note that if  $\sigma = 1$ , individuals breeding in little patches have the same fecundity as those breeding in big patches; if  $\sigma < 1$ , individuals breeding in little patches have less fecundity than those breeding in big patches; and if  $\sigma > 1$ , individuals breeding in little patches have greater fecundity than those breeding in big patches. The average investment strategy in the population is represented by  $z$ . After reproduction, a fraction  $1 - m$  of the offspring stay in their natal patch, while the remaining fraction  $m$  disperse to random patches in the population. After dispersal, patches may undergo changes in their size according to a Markov process: big patches remain big with probability  $\alpha$  and become little with probability  $1 - \alpha$ , whereas little patches remain little with probability  $\beta$  and become big with probability  $1 - \beta$  (see "Patch Dynamics," available online, for details). Subsequently, offspring compete for breeding places, with all nonbreeding individuals perishing. Table 1 provides a summary of model notation.

### *Helping and Harming*

We classify social behaviors according to their impact on fecundity. The impact on the actor's fecundity is  $-C \equiv \partial f(x, y)/\partial x$ , and the impact on the patchmates' fecundity is  $B \equiv \partial f(x, y)/\partial y$ , where, assuming vanishing variation in investment strategies, we evaluate derivatives at  $x = y = z$  (Taylor and Frank 1996). Helping behaviors are those that improve the fecundity of patchmates ( $B > 0$ ), and harming behaviors are those that reduce the fecundity of patchmates ( $B < 0$ ; Lehmann et al. 2006*a*; West and Gardner 2010). We employ the neighbor-modulated approach to kin-selection analysis (Hamilton 1964; Taylor 1996; Taylor and Frank 1996; Frank 1997, 1998; Rousset 2004; Taylor et al. 2007) to determine the direction of natural selection acting on the social trait (see sections "Reproductive Success" through "Reproductive Value" in the appendix, available online, for details). The condition for natural selection to favor the evolution of obligate social behavior is given by

Table 1: Summary of model notation

Symbol	Meaning
B	Big patch
L	Little patch
O	Obligate social behavior
T	Reference to quantities in Taylor's (1992a) model
P	Primary recipient
S	Secondary recipient
$A_x$	Potential for helping in condition $X$
$\alpha$	Probability that a big patch remains big
$\beta$	Probability that a little patch remains little
$c_x$	Class reproductive value in condition $X$
$F_x$	Fecundity of a breeding female in condition $X$
$h_x$	Probability of cophilotry in condition $X$
$m$	Migration rate
$n_x$	Patch size in condition $X$
$\pi_x$	History of density state in condition $X$
$p$	Frequency of big patches in the population
$r_x$	"Other-only" relatedness of an individual in condition $X$
$R_x$	"Whole-group" relatedness of an individual in condition $X$
$\sigma$	Fecundity factor
$\tau$	Temporal coefficient of correlation
$v_x$	Reproductive value of an individual in condition $X$
$v_x$	Relative reproductive value of patch in condition $X$
$x$	Level of helping of a focal actor
$y$	Average level of the neighbor's social behavior

$$c_B[-C + r_{P|B}B - (B - C)r_{S|B}v_{S|B}] - C + r_{P|B}B - (B - C)r_{S|B}v_{S|B} > 0 \tag{2}$$

$$+ c_L[-C + r_{P|L}B - (B - C)r_{S|L}v_{S|L}] > 0, \tag{1} \quad \text{and}$$

$$-C + r_{P|L}B - (B - C)r_{S|L}v_{S|L} > 0 \tag{3}$$

where  $c_B$  and  $c_L$  are the class reproductive values of individuals breeding within big and little patches, respectively (Fisher 1930; Taylor 1990; Grafen 2006);  $r_{P|B}$  and  $r_{P|L}$  are the relatednesses of the actor to her groupmates—"primary recipients"—in big patches and little patches, respectively;  $r_{S|B}$  and  $r_{S|L}$  are the relatednesses of the actor to those offspring competing for breeding sites with the offspring born in her patch—"secondary recipients"—in big patches and little patches, respectively; and  $v_{S|B}$  and  $v_{S|L}$  are the reproductive values of secondary recipients, relative to those of primary recipients, in big patches and little patches, respectively (see "Relatedness" and "Reproductive Value," available online, for details). Note that we are defining actor, primary recipients, and secondary recipients in terms of the three selection pressures that arise from the actor's social behavior, namely, the cost to the actor, the benefit to groupmates, and the kin-competition effect. Also note that a focal individual may experience, either directly or through her offspring, more than one of these selection pressures during her lifetime. The conditions for natural selection to favor the evolution of facultative social behavior are given by

for actors in big patches and little patches, respectively. The left-hand side (LHS) of conditions (1)–(3) can be interpreted as the inclusive fitness effect of the social behavior (Hamilton 1963, 1964, 1970). First, the behavior imposes a cost  $C$  on the actor's own fecundity. Second, it provides a benefit  $B$  to the fecundity of the actor's groupmates, and this benefit is weighted by the relatedness of these primary recipients  $r_p$ . Finally, the behavior leads to an increase of  $B - C$  in the number of offspring born in the patch, and hence it excludes an equal number of secondary recipients who compete with these extra offspring for breeding opportunities. Moreover, this effect is weighted by the relatedness of secondary recipients  $r_s$  and also by the relative reproductive value of secondary recipients  $v_s$ .

Transforming inequalities (1)–(3) into equalities and rearranging into the form  $C/B = A$ , we obtain the "potential for helping"  $A$  (Rodrigues and Gardner 2012; see also Gardner 2010). This represents the valuation that an actor places on the fecundity of the primary recipients, relative to her own fecundity. Thus, if  $A > 0$ , the actor is selectively

avored to help her patchmates, provided that the personal cost to herself is sufficiently small, whereas if  $A < 0$ , the actor is selectively favored to harm her patchmates, provided that the personal cost to herself is sufficiently small. We may define the additive inverse of the potential for helping (i.e.,  $-A$ ) as the “potential for harming” (Rodrigues and Gardner 2012).

From inequality (1), the potential for obligate helping is given by

$$A_O = \frac{c_B(r_{P|B} - r_{S|B}v_{S|B}) + c_L(r_{P|L} - r_{S|L}v_{S|L})}{c_B(1 - r_{S|B}v_{S|B}) + c_L(1 - r_{S|L}v_{S|L})}. \quad (4)$$

From inequalities (2) and (3), the potential for facultative helping is given by

$$A_B = \frac{r_{P|B} - r_{S|B}v_{S|B}}{1 - r_{S|B}v_{S|B}} \quad (5)$$

and

$$A_L = \frac{r_{P|L} - r_{S|L}v_{S|L}}{1 - r_{S|L}v_{S|L}} \quad (6)$$

for big patches and little patches, respectively.

### Results and Analysis

In the next two sections, we derive the main results of our model. We begin by considering that the fecundity of individuals is independent of patch size ( $\sigma = 1$ ). We then relax this assumption to consider cases where the individuals’ fecundity depends on the patch they are in ( $0 < \sigma < \infty$ ; see table 2 for a summary of the results).

#### Group-Size Heterogeneity

*Spatial Heterogeneity.* We first consider a population in which group size varies between patches within generations but not within patches between generations (i.e., spatial heterogeneity only). We derive the following results:

*Result 1.* Spatial heterogeneity in group size has no im-

act on the evolution of obligate social behavior ( $A_O = 0$ ), irrespective of parameter values (see “Relatedness” for details). This extends Taylor’s (1992a) result for homogeneous populations to populations with spatial heterogeneity in group size.

*Result 2.* Spatial heterogeneity in group size has no impact on the evolution of facultative social behavior ( $A_B = A_L = 0$ ), irrespective of parameter values (see “Relatedness” for details). This extends result 1 for obligate social behavior to facultative social behavior.

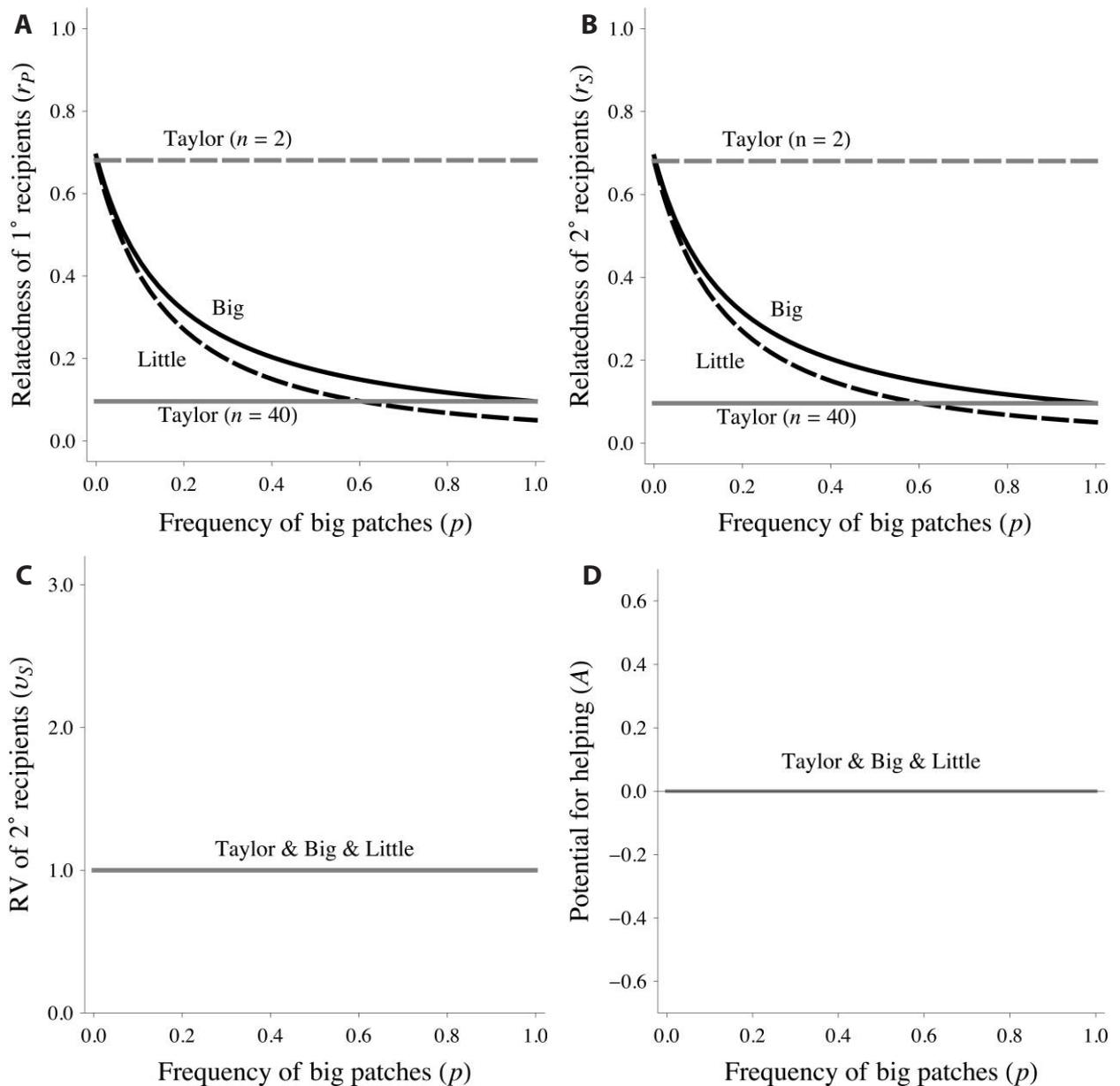
In Taylor’s (1992a) model, the relatedness of primary recipients is given by  $r_{P|T} = h_T R_T$ , where  $h_T$  is the probability that two offspring chosen at random from a patch are both philopatric to that patch after the dispersal phase and  $R_T$  is the relatedness of these two offspring. The relatedness of secondary recipients is also given by  $r_{S|T} = h_T R_T$ . As a consequence, the relatedness of primary recipients is equal to that of secondary recipients,  $r_{P|T} = r_{S|T}$ . Moreover, because Taylor’s (1992a) model is of a homogeneous population, all recipients have the same reproductive value in a neutral population, and hence  $v_{S|T} = 1$ . Since actors place the same value on primary and secondary recipients irrespective of population viscosity ( $r_{P|T} = r_{S|T}v_{S|T}$ ; fig. 1), selection acting on the social behavior is null ( $A_T = 0$ ).

Turning to our model of a spatially heterogeneous population and focusing on facultative helping in big patches, we find that the relatedness of primary recipients is given by  $r_{P|B} = h_B R_B$ , where  $h_B$  is the probability that two offspring chosen at random from a big patch are both philopatric to that patch after the dispersal phase and  $R_B$  is the relatedness of these two offspring. The relatedness of secondary recipients is also given by  $r_{S|B} = h_B R_B$ . As a consequence, the relatedness of primary recipients is equal to that of secondary recipients,  $r_{P|B} = r_{S|B}$ . In addition, the reproductive value of primary recipients is equal to the reproductive value of secondary recipients,  $v_{S|B} = 1$ . Since actors attribute the same value to primary and secondary recipients ( $r_{P|B} = r_{S|B}v_{S|B}$ ; fig. 1), selection acting

**Table 2:** Summary of model results

	Group-size heterogeneity with homogeneous fecundity			Group-size heterogeneity with variable fecundity		
	Spatial	Temporal	Spatial and temporal	Spatial	Temporal	Spatial and temporal
Obligate	Cancel	Cancel	Cancel	–	–	–
Facultative (big patches)	Cancel	Helping	Helping	–	–	+
Facultative (little patches)	Cancel	Harming	Harming	–	–	+

Note: Results for group-size heterogeneity with homogeneous fecundity ( $\sigma = 1$ ) for obligate and facultative behavior and for different group-size heterogeneity assumptions (spatial only, temporal only, and spatial and temporal) and impact of variable fecundity on the results derived for group-size heterogeneity with homogeneous fecundity (minus sign: has no impact; plus sign: has an impact). Obligate trait expression and spatial heterogeneity always lead to the cancellation result. Variable fecundity plays a role in the evolution of social behavior only when populations are spatially and temporally heterogeneous, in which case it can change the predictions of the homogeneous-fecundity model.



**Figure 1:** Impact of spatial heterogeneity in group size on relatedness, reproductive value, and potential for helping. *A*, Relatedness of primary (1°) recipients ( $r_P$ ) as a function of the frequency of big patches ( $p$ ) for big patches ( $r_{P|B}$ ; solid black line), little patches ( $r_{P|L}$ ; dashed black line), and Taylor's (1992a) reference model ( $r_{P|T}$ ; gray lines). The relatedness of primary recipients is higher in big patches, because of higher philopatry, and lower in little patches, because of lower philopatry, than that in the reference model with equivalent group sizes ( $r_{P|B} > r_{P|T}$  and  $r_{P|L} < r_{P|T}$ ). *B*, Relatedness of secondary (2°) recipients ( $r_S$ ) as a function of the frequency of big patches ( $p$ ). The relatedness of secondary recipients is higher in big patches, because of higher philopatry, and lower in little patches, because of lower philopatry, than that in the reference model with equivalent group sizes ( $r_{S|B} > r_{S|T}$  and  $r_{S|L} < r_{S|T}$ ). Importantly, primary and secondary recipients are equally related ( $r_{P|B} = r_{S|B}$  and  $r_{P|L} = r_{S|L}$ ). *C*, The relative reproductive value (RV) of secondary recipients ( $v_S$ ) is constant with respect to the frequency of big patches ( $p$ ) and is equal to that of the reference model ( $v_{S|B} = v_{S|L} = v_{S|T} = 1$ ). *D*, The potential for facultative helping ( $A_B$  and  $A_L$ ) is constant with respect to the frequency of big patches ( $p$ ) and is equal to that of the reference model ( $A_B = A_L = A_T = 0$ ). In all plots, we assume  $n_B = 40$ ,  $n_L = 2$ ,  $m = 0.10$ , and  $\tau = 1.00$ .

on the social behavior is null ( $A_B = 0$ ). The same argument can be made in relation to little patches (namely,  $r_{P|L} = h_L R_L$ ,  $r_{S|L} = h_L R_L$ ,  $r_{P|L} = r_{S|L}$ ,  $v_{S|L} = 1$ ,  $r_{P|L} = r_{S|L} v_{S|L}$ , and  $A_L = 0$ ; fig. 1). Finally, selection acting on obligate social behavior is determined by selection acting on each size-type patch weighted by the respective class reproductive values (see the LHS of inequality [1]). As we have just seen, selection acting on each size-type patch is null ( $A_B = 0$  and  $A_L = 0$ ). Therefore, selection acting on obligate social behavior is also null ( $A_O = 0$ ). Full mathematical details are given in “Cancellation of Obligate Helping and Harming,” available online.

*Temporal Heterogeneity.* We next consider a population in which group size varies within patches between generations but not between patches within generations (i.e., temporal heterogeneity only). In particular, we assume that all patches in the population become big with probability  $p$  or little with probability  $1 - p$  in every generation. We derive the following results:

*Result 3.* Temporal heterogeneity in group size has no impact on the evolution of obligate social behavior ( $A_O = 0$ ), irrespective of parameter values (see “Temporal Heterogeneity” in the appendix, available online, for details). This extends Taylor’s (1992a) cancellation result for homogeneous populations to populations with temporal heterogeneity in group size.

*Result 4.* Temporal heterogeneity in group size does have an impact on the evolution of facultative social behavior ( $A_B \neq 0$  and  $A_L \neq 0$ ). Specifically, selection favors helping in patches with more breeders ( $A_B \geq 0$ ) and harming in patches with fewer breeders ( $A_L \leq 0$ ; see “Temporal Heterogeneity” in the appendix for details).

Why does Taylor’s (1992a) result no longer hold for facultative social behavior in temporally heterogeneous populations? Focusing on helping in big patches only, we find that the relatedness of primary recipients is given by  $r_{P|B} = ph_B R_B + (1 - p)h_L R_L$ , where  $p$  is the probability that the patch was big in the previous generation,  $h_B$  is the probability that two offspring chosen at random from a big patch are both philopatric to that patch after the dispersal phase,  $R_B$  is the relatedness of these two offspring,  $1 - p$  is the probability that the patch was little in the previous generation,  $h_L$  is the probability that two offspring chosen at random from a little patch are both philopatric to that patch after the dispersal phase, and  $R_L$  is the relatedness of these two offspring. The relatedness of secondary recipients is  $r_{S|B} = h_B R_B$ . Note that  $h_B = h_L = (1 - m)^2$  and that because there are more breeders in big patches than in little patches, relatedness is higher among the offspring born in the former and lower among the offspring born in the latter patch type; that is,  $R_B < R_L$ .

Consequently, the relatedness of primary recipients is

greater than the relatedness of secondary recipients; that is,  $r_{P|B} > r_{S|B}$ . In addition, because within each generation all individuals are identical, the reproductive value of primary recipients is equal to the reproductive value of secondary recipients; that is,  $v_{P|B} = v_{S|B}$ . Since actors place more value on primary recipients than on secondary recipients, selection favors helping behavior in big patches ( $A_B > 0$ ; fig. 2). A similar argument can be made in relation to little patches. However, because primary recipients are now less valuable than secondary recipients for actors, selection favors harming behavior in little patches:  $r_{P|L} = ph_B R_B + (1 - p)h_L R_L$ ,  $r_{S|L} = h_L R_L$ ,  $h_B = h_L = (1 - m)^2$ ,  $R_B < R_L$ ,  $v_{S|L} = 1$ , and  $A_L < 0$  (fig. 2; see “Temporal Heterogeneity” in the appendix for details). In other words, when patches increase in size (i.e., after a bottleneck), helping is favored; by contrast, when patches decrease in size, harming is favored.

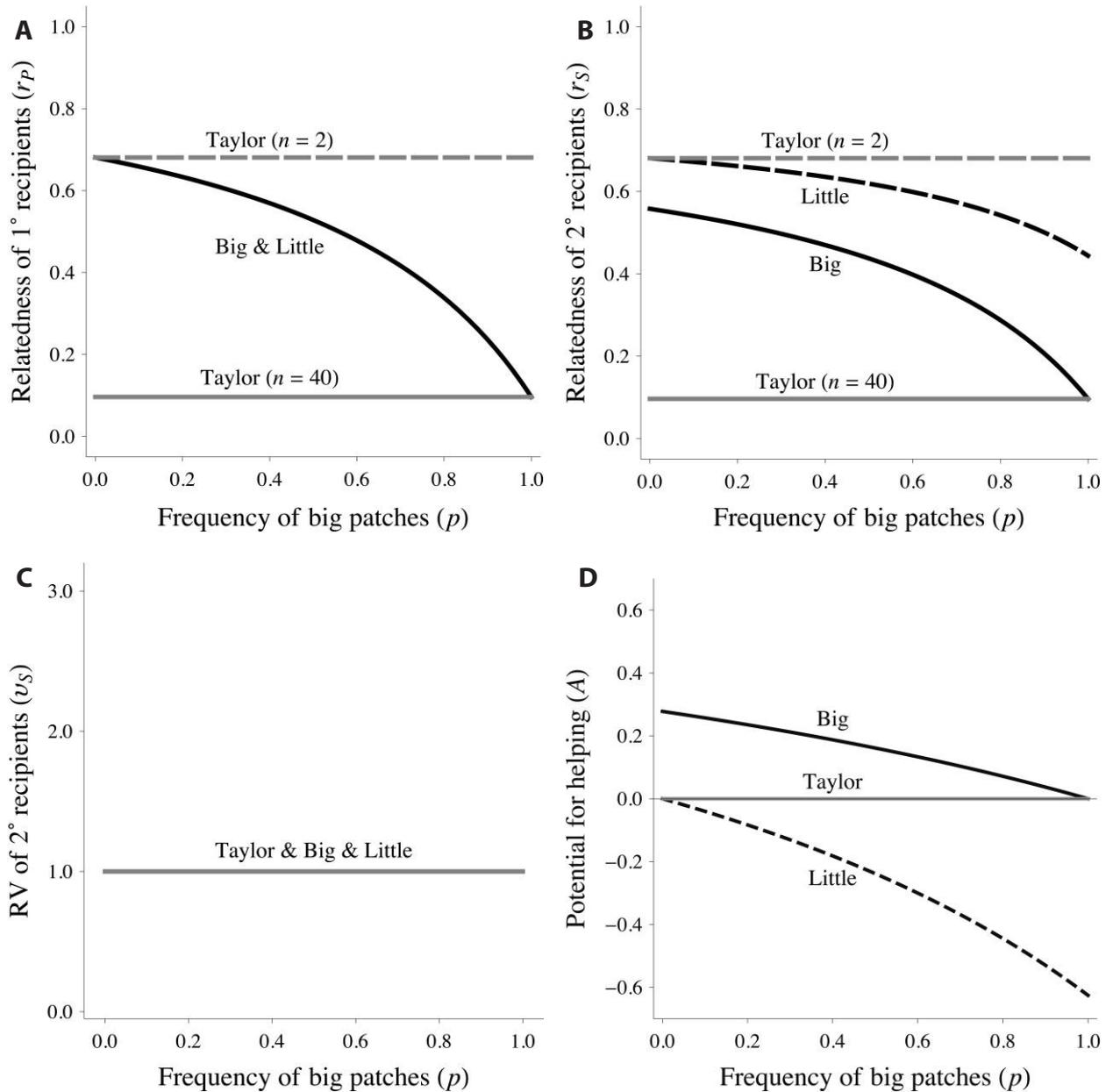
When we turn to obligate social behavior, relatednesses are averaged over patch types, weighted by the corresponding class reproductive values; that is,  $r_{P|O} = c_B r_{P|B} + c_L r_{P|L}$ . Because the relatedness of primary recipients in big patches is equal to that of primary recipients in little patches ( $r_{P|B} = r_{P|L}$ ) and the class reproductive values sum to unity ( $c_B + c_L = 1$ ), the relatedness of primary recipients is given by  $r_{P|O} = r_{P|B} = r_{P|L} = ph_B R_B + (1 - p)h_L R_L$ . The relatedness of secondary recipients is given by  $r_{S|O} = c_B h_B R_B + c_L h_L R_L$ . Because class reproductive values are given by  $c_B = p$  and  $c_L = 1 - p$ , we find that the relatedness of primary recipients is equal to that of secondary recipients ( $r_{P|O} = r_{S|O}$ ). Consequently, selection acting on obligate social behavior is null ( $A_O = 0$ ; see “Temporal Heterogeneity” in the appendix for details).

*Spatial and Temporal Heterogeneity.* Finally, we consider a population characterized by heterogeneity in group size both between patches within generations and within patches between generations (i.e., spatial and temporal heterogeneity). We derive the following results:

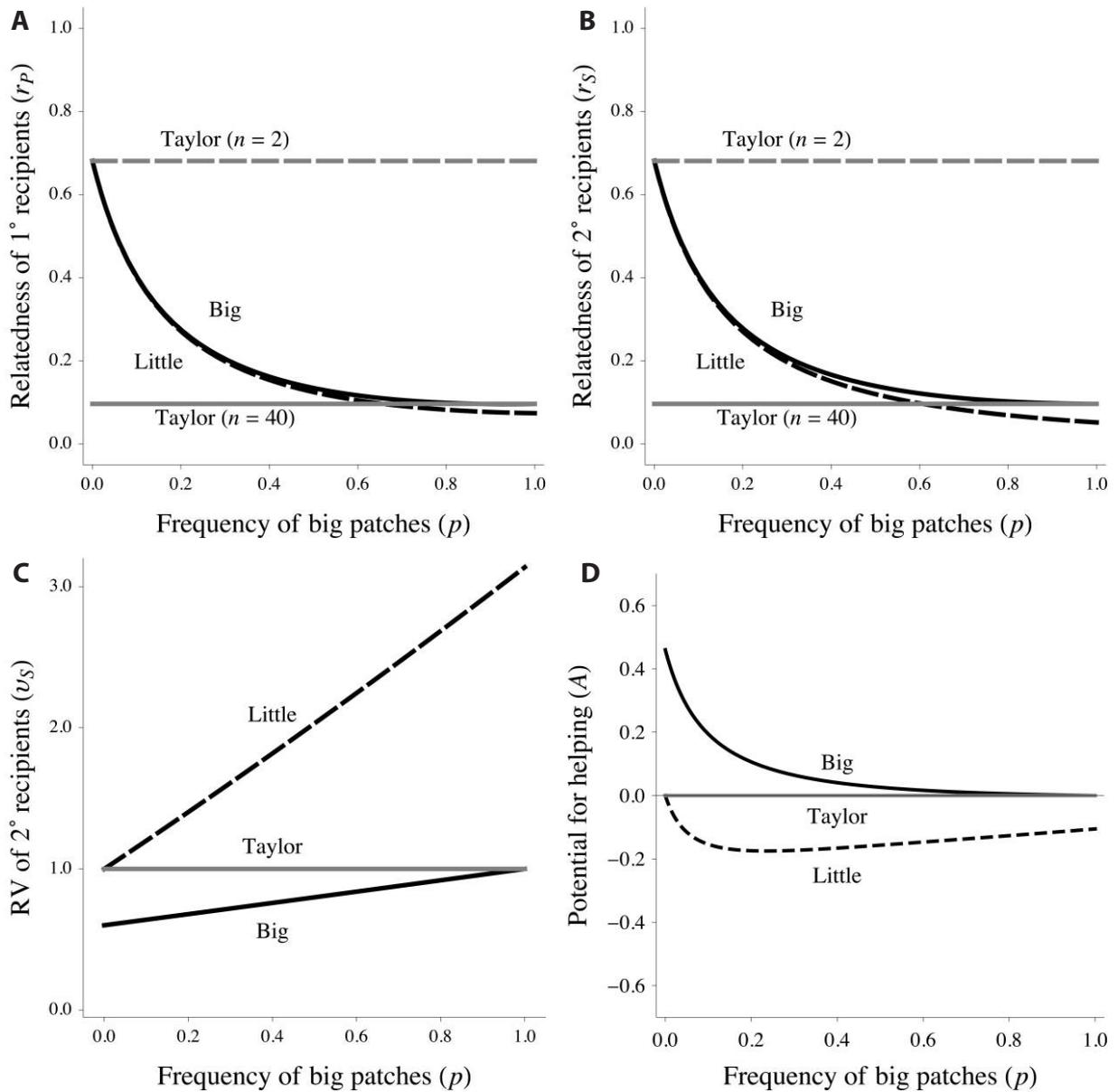
*Result 5.* Spatial and temporal heterogeneity in group size has no impact on the evolution of obligate social behavior ( $A_O = 0$ ), irrespective of parameter values (see “Cancellation of Obligate Helping and Harming” for details). This extends Taylor’s (1992a) cancellation result for homogeneous populations to obligate social behaviors in spatially and temporally heterogeneous populations.

*Result 6.* Spatial and temporal heterogeneity in group size does influence the evolution of facultative social behaviors ( $A_B \neq 0$  and  $A_L \neq 0$ ). Specifically, selection favors helping in patches with more breeders ( $A_B \geq 0$ ) and harming in patches with fewer breeders ( $A_L \leq 0$ ; see “Cancellation of Obligate Helping and Harming” for details).

Why does Taylor’s (1992a) result no longer hold when



**Figure 2:** Impact of temporal heterogeneity in group size on relatedness, reproductive value, and potential for helping. *A*, Relatedness of primary (1°) recipients ( $r_P$ ) as a function of the frequency of big patches ( $p$ ) for big patches ( $r_{P|B}$ ; solid black line), little patches ( $r_{P|L}$ ; dashed black line), and Taylor's (1992a) reference model ( $r_{P|T}$ ; gray lines). The relatedness of primary recipients is higher in big patches, because of higher cophilocy, and lower in little patches, because of lower cophilocy, than that of the reference model with equivalent group sizes ( $r_{P|B} > r_{P|T}$  and  $r_{P|L} < r_{P|T}$ ). *B*, Relatedness of secondary (2°) recipients ( $r_S$ ) as a function of the frequency of big patches ( $p$ ). The relatedness of secondary recipients is higher in big patches, because of higher cophilocy, and lower in little patches, because of lower cophilocy, than that of the reference model with equivalent group sizes ( $r_{S|B} > r_{S|T}$  and  $r_{S|L} < r_{S|T}$ ). Importantly, in big patches primary recipients are more related to the actor than are secondary recipients, while in little patches primary recipients are less related to the actor than are secondary recipients ( $r_{P|B} > r_{S|B}$  and  $r_{P|L} < r_{S|L}$ ). *C*, The relative reproductive value (RV) of secondary recipients ( $v_S$ ) is constant with respect to the frequency of big patches ( $p$ ) and is equal to that of the reference model ( $v_{S|B} = v_{S|L} = v_{S|T} = 1$ ). *D*, Potential for facultative helping ( $A_B$  and  $A_L$ ) as a function of the frequency of big patches ( $p$ ). In big patches there is potential for helping ( $A_B > 0$ ), while in little patches there is potential for harming ( $A_L < 0$ ). In all plots, we assume  $n_B = 40$ ,  $n_L = 2$ , and  $m = 0.10$ .

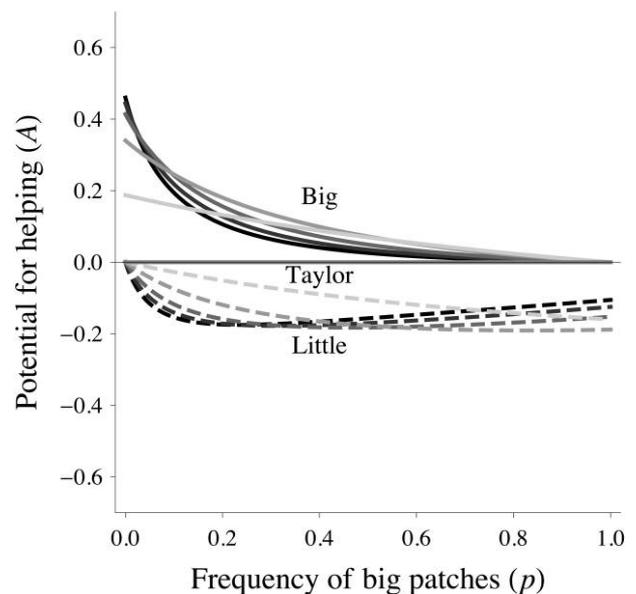


**Figure 3:** Impact of spatial and temporal heterogeneity in group size on relatedness, reproductive value, and potential for helping. *A*, Relatedness of primary (1°) recipients ( $r_P$ ) as a function of the frequency of big patches ( $p$ ) for big patches ( $r_{P|B}$ ; solid black line), little patches ( $r_{P|L}$ ; dashed black line), and Taylor's (1992a) reference model ( $r_{P|T}$ ; gray lines). The relatedness of primary recipients is higher in big patches, because of higher cophipatry, and lower in little patches, because of lower cophipatry, than that of the reference model with equivalent group sizes ( $r_{P|B} > r_{P|T}$  and  $r_{P|L} < r_{P|T}$ ). *B*, Relatedness of secondary (2°) recipients ( $r_S$ ) as a function of the frequency of big patches ( $p$ ). The relatedness of secondary recipients is higher in big patches, because of higher cophipatry, and lower in little patches, because of lower cophipatry, than that of the reference model with equivalent group sizes ( $r_{S|B} > r_{S|T}$  and  $r_{S|L} < r_{S|T}$ ). Importantly, in big patches primary recipients are less related to the actor than are secondary recipients, while in little patches primary recipients are more related to the actor than are secondary recipients ( $r_{P|B} < r_{S|B}$  and  $r_{P|L} > r_{S|L}$ ). *C*, Relative reproductive value (RV) of secondary recipients ( $v_S$ ) as a function of the frequency of big patches ( $p$ ). In big patches the relative reproductive value of secondary recipients is less than or equal to that of the reference model, and in little patches the relative reproductive value is greater than or equal to that of the reference model ( $v_{S|B} \leq v_{S|T}$  and  $v_{S|L} \geq v_{S|T}$ ). *D*, Potential for facultative helping ( $A_B$  and  $A_L$ ) as a function of the frequency of big patches ( $p$ ). In big patches there is potential for helping ( $A_B > 0$ ), while in little patches there is potential for harming ( $A_L < 0$ ). By numerical analysis, these results were confirmed to hold over the entire range of possible parameter state space. In all plots, numerical examples are given for  $n_b = 40$ ,  $n_l = 2$ ,  $m = 0.1$ , and  $\tau = 1/2$ .

we consider facultative helping and harming in populations that are both spatially and temporally heterogeneous with respect to group size? Focusing our attention on facultative helping in big patches, we find that the potential for this to be favored is given by equation (5); that is,  $A_B = (r_{P|B} - r_{S|B}v_{S|B})/(1 - r_{S|B}v_{S|B})$ . If patch quality is heterogeneous in space and time, then the actor's big patch may have been either big or little in the previous generation, and so her relatedness to primary recipients  $r_{P|B}$  is a weighted average of the product of the probability of cophilopatry  $h_B$  and whole-group relatedness  $R_B$  for big patches and the product of the probability of cophilopatry  $h_L$  and whole-group relatedness  $R_L$  for little patches (see "Relatedness" for details). In contrast, her relatedness  $r_{S|B}$  to the secondary recipients is the product of the probability of cophilopatry  $h_B$  and whole-group relatedness  $R_B$  for her big patch (see "Relatedness" for details). Moreover, the reproductive value of her secondary recipients—that is, adults in her patch in the next generation—need not be equal to her own reproductive value. While she has the reproductive value of an individual in a big patch, theirs is a weighted average of the reproductive values for individuals in big and little patches. Hence,  $v_{S|B} < 1$  (see "Reproductive Value" for details).

Consequently, the actor may place different values on her primary and secondary recipients:  $r_{P|B} \neq r_{S|B}v_{S|B}$  (fig. 3; see "Cancellation of Obligate Helping and Harming" for details). Hence, the potential for facultative helping in big patches may be nonzero,  $A_B \neq 0$ . The same is true for facultative helping in little patches:  $r_{P|L} \neq r_{S|L}$  and  $v_{S|L} > 1$ , so  $r_{P|L} \neq r_{S|L}v_{S|L}$  and  $A_L \neq 0$  (fig. 3; see "Cancellation of Obligate Helping and Harming" for details). We find that there is potential for helping in big patches and potential for harming in little patches ( $A_B > 0$  and  $A_L < 0$ ; fig. 4).

Although actors in big patches place different values on their primary and secondary recipients and this is also true of actors in little patches, we find that, on average over both patch types, actors place equal value on their primary and secondary recipients. This owes to the properties of the ecological dynamics that determines how patch size changes across generations. Specifically, this is characterized by a time-homogeneous Markov chain, such that the backward processes that determine the value of primary recipients are identical to the forward processes that determine the value of secondary recipients (see "Cancellation of Obligate Helping and Harming" for details). Hence, while there is nonzero potential for facultative helping and harming in big and little patches ( $A_B \neq 0$  and  $A_L \neq 0$ ), there is zero potential for obligate helping and harming ( $A_O = 0$ ) in populations characterized by both spatial and temporal heterogeneity.



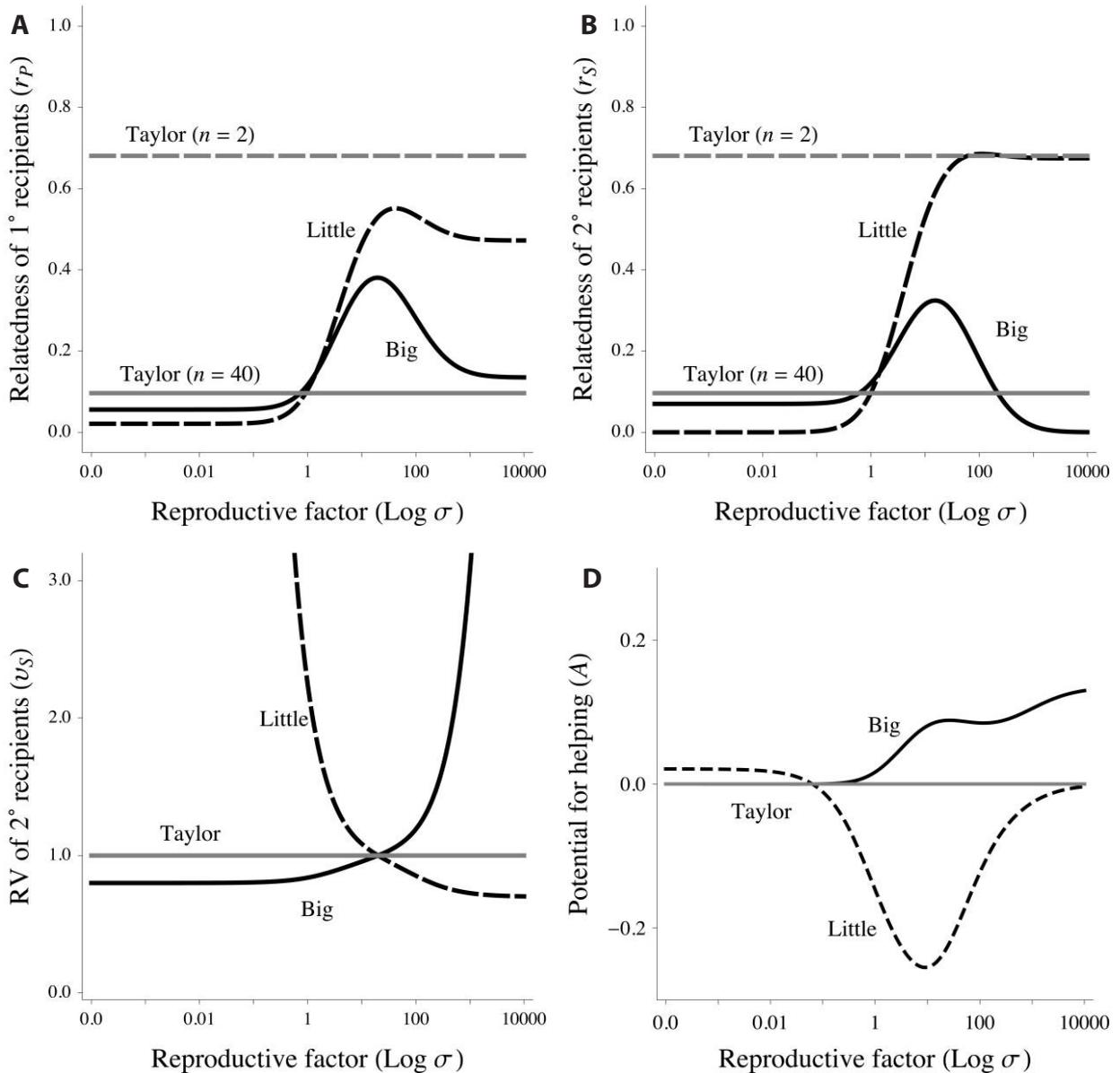
**Figure 4:** Impact of spatial and temporal heterogeneity in group size on the potential for helping: potential for facultative helping ( $A_B$  and  $A_L$ ) as a function of the frequency of high-quality patches ( $p$ ), for varying sizes of big patches ( $n_B = 40, 30, 20, 10$ , and  $4$ ). Darker lines represent larger patch sizes, while lighter lines represent smaller patch sizes. There is always potential for helping in big patches, while there is always potential for harming in little patches. Parameter values:  $n_L = 2$ ,  $m = 0.1$ , and  $\tau = 1/2$ .

#### Group-Size Heterogeneity with Variable Fecundity

Here we consider that a group's size may affect its constituents' fecundity ( $0 < \sigma < \infty$ ). This enables us to study how the interaction between patch-size heterogeneity and variable fecundity mediates the evolution of obligate and facultative helping and harming. For example, several studies have shown that increasing a group's size may decrease the fitness of its constituent members, for example, because of infectious diseases (e.g., Rifkin et al. 2012). By contrast, other studies have shown that increasing a group's size may increase the fitness of its constituent members, for example, because of a decrease in predation rates or more efficient foraging (e.g., Hamilton 1971a). We derive the following results:

**Result 7.** We find that introducing patch productivity heterogeneity has no impact on the results derived for patch-size heterogeneity when (1) individuals express obligate behavior and/or (2) there is spatial heterogeneity only or (3) there is temporal heterogeneity only.

**Result 8.** We find that introducing patch productivity heterogeneity does have an impact on the results derived for patch-size heterogeneity when individuals express facultative behavior and there is both spatial and temporal heterogeneity.



**Figure 5:** Impact of spatial and temporal heterogeneity in group size and variable fecundity on relatedness, reproductive value, and potential for helping. *A*, Relatedness of primary (1°) recipients ( $r_p$ ) as a function of the reproductive factor ( $\sigma$ ) for big patches ( $r_{p|B}$ ; solid black line), little patches ( $r_{p|L}$ ; dashed black line), and Taylor's (1992a) reference model ( $r_{p|T}$ ; gray lines). In big patches, for lower values of the reproductive factor, the relatedness of primary recipients is less than that of the reference model ( $r_{p|B} < r_{p|T}$ ), while for higher values of the reproductive factor, the relatedness of primary recipients is greater than that of the reference model ( $r_{p|B} > r_{p|T}$ ). In little patches, the relatedness of primary recipients is always less than that of the reference model ( $r_{p|L} < r_{p|T}$ ). *B*, Relatedness of secondary (2°) recipients ( $r_s$ ) as a function of the reproductive factor ( $\sigma$ ). In big patches, for lower values of the reproductive factor, the relatedness of secondary recipients is less than that of the reference model ( $r_{s|B} < r_{s|T}$ ), while for higher values of the reproductive factor, the relatedness of secondary recipients is greater than that of the reference model ( $r_{s|B} > r_{s|T}$ ). In little patches, the relatedness of secondary recipients is less than that of the reference model for a wide range of the reproductive factor ( $r_{s|L} < r_{s|T}$ ). Importantly, in big patches and for lower values of the reproductive factor, primary recipients are less related to the actor than are secondary recipients ( $r_{p|B} < r_{s|B}$ ), while for higher values of the reproductive factor, primary recipients are more related to the actor than are secondary recipients ( $r_{p|B} > r_{s|B}$ ); by contrast, in little patches, primary recipients are more related to the actor than are secondary recipients for smaller values of the reproductive factor ( $r_{p|L} > r_{s|L}$ ), while for higher values of the reproductive factor, primary recipients are less related to the actor than are secondary recipients ( $r_{p|L} < r_{s|L}$ ). *C*, Relative reproductive value (RV) of secondary recipients ( $v_s$ ) as a function of the reproductive factor ( $\sigma$ ). In big patches, the relative reproductive value of secondary recipients is less than that of the reference model for smaller values of the reproductive factor ( $v_{s|B} < v_{s|T}$ ), while it is greater than that of the reference model for higher values of the reproductive factor ( $v_{s|B} > v_{s|T}$ ). In little patches, the relative

Why does variable fecundity have no impact on the results obtained for the evolution of helping and harming in a group-size-heterogeneous population with homogeneous fecundity when we assume (1) obligate behavior and/or (2) group-size spatial heterogeneity only or (3) group-size temporal heterogeneity only? We find that the cancellation result obtained for a group-size-heterogeneous population (i.e., our results 1–3 and 5) continues to hold, for the same reason, irrespective of variation in fecundity. This is because, while variable fecundity alters both the relatedness coefficient and the reproductive value of each recipient, the relationship between the overall value (i.e., life-for-life relatedness) of primary and secondary recipients remains unaltered, and therefore the cancellation result continues to hold.

Why does variable fecundity influence the consequences of group-size heterogeneity in the model with both spatial and temporal heterogeneity? We find that if breeders in little patches have low fecundity (low  $\sigma$ ), then facultative helping can be favored. This is because secondary recipients are usually unrelated immigrants, and therefore the relatedness of secondary recipients is close to 0, making the kin-competition effect negligible. By contrast, facultative harming is favored in big patches. This is because secondary recipients are usually philopatric to the patch, which intensifies the kin-competition effect, favoring harming behavior. By contrast, if breeders in little patches have high fecundity (high  $\sigma$ ), then we recover the same qualitative result derived for patch-size heterogeneity with homogeneous fecundity, namely, that helping is favored in big patches and harming is favored in little patches (fig. 5).

### Discussion

Population viscosity has been proposed as a general mechanism that promotes interactions between closely related individuals and may therefore favor the evolution of kin-selected traits such as altruistic cooperation (Hamilton 1964, 1971*b*). However, it is now well recognized that population viscosity also promotes competition between kin. Here we have shown that Taylor's (1992*a*) seminal result, that the relatedness and competition effects of viscosity exactly cancel in homogeneous populations, extends to populations where groups vary in size because of extrinsic

factors. Specifically, population viscosity has no impact on the evolution of obligate helping or harming in populations characterized by spatial and/or temporal heterogeneity in group size. Moreover, if individuals can facultatively adjust their behavior conditionally on their group's size, then there is no selection for either helping or harming when there is only spatial heterogeneity in group size. However, we have shown that facultative helping is favored in big groups and that facultative harming is favored in little groups when there is either only temporal or both temporal and spatial heterogeneity in group size. More generally, we have shown that Taylor's (1992*a*) cancellation result for obligate behavior is robust across a wide range of ecologically realistic scenarios, and we have clarified why this cancellation occurs. Specifically, the proper weights of each selection pressure are the life-for-life relatedness coefficients, which are given in terms of recipients' relatednesses and reproductive values, both of which emerge from the specific demographic dynamics of the population (see "Cancellation of Obligate Helping and Harming" for details).

Here we considered that differences in patch size may manifest themselves as differences in the number of breeders that the patches are able to support. Rodrigues and Gardner (2012) also studied how population heterogeneity in patch quality mediates the evolution of social behavior when differences in patch quality are manifested as differences in the fecundity of the individuals breeding in the patches. In line with our results here, Rodrigues and Gardner (2012) found that spatial and/or temporal heterogeneity has no impact on the evolution of obligate helping or harming. However, in contrast to our results, Rodrigues and Gardner (2012) found that both spatial and temporal heterogeneity were required in order for facultative helping and harming to be favored. Thus, an alteration in a particular assumption as to how patch quality translates into improved productivity—that is, more breeders having the same fecundity rather than the same number of breeders having greater fecundity—leads to a less stringent condition for population viscosity to promote the evolution of social behavior. Furthermore, we have explained why variable fecundity influences the results of patch-size heterogeneity and homogeneous fecundity when the spatial and temporal treatment is considered, but not when the temporal treatment alone is considered. This clarifies the

---

reproductive value is greater than that of the reference model for lower values of the reproductive factor ( $v_{S|T} < v_{S|L}$ ), while it is less than that of the reference model for higher values of the reproductive factor ( $v_{S|T} > v_{S|L}$ ). *D*, Potential for facultative helping ( $A_B$  and  $A_L$ ) as a function of the reproductive factor ( $\sigma$ ). In big patches there is potential for harming for lower values of the reproductive factor ( $A_B < 0$ ), while there is potential for helping for higher values of the reproductive factor ( $A_B > 0$ ). In little patches there is potential for helping for lower values of the reproductive factor ( $A_L > 0$ ), while there is potential for harming for higher values of the reproductive factor ( $A_L < 0$ ). In all plots, numerical examples are given for  $n_b = 40$ ,  $n_l = 2$ ,  $m = 0.1$ ,  $p = 0.6$ , and  $\tau = 1/2$ .

importance of understanding how ecological and genetic factors affect life-for-life coefficients of relatedness.

Grafen (2007) also studied the evolution of obligate helping in a population characterized by heterogeneity in group size and limited dispersal. However, his model assumed global competition for reproductive resources and hence the absence of kin competition. Consequently, he found that obligate helping readily evolved in the context of his model. Here we have considered that limited dispersal determines both the genetic structure of the population and the scale of competition. This emphasizes the importance of demography for mediating the costs and benefits of social behavior as well as the genetic relatedness of social partners. Lehmann et al. (2006b) studied the evolution of social traits in a metapopulation characterized by patch extinctions, and we recover a similar scenario by considering that there are no breeders in little patches ( $n_l = 0$ ). In this special case of our model, it is meaningless to distinguish obligate from facultative behavior, as all breeders are in big patches, and consequently there is no avenue for helping or harming to be favored in this scenario. The Lehmann et al. model considered that transitions between occupied and empty patches are governed by social interaction, allowing helping behaviors to be favored. This contrasts with our results, which do not require this mechanism for helping and harming to be favored. Our results also show that environmental heterogeneity favors plastic social responses to local environments and therefore support the idea that plasticity and varying ecological factors can play a role in promoting social behavior in viscous populations (e.g., Kelly 1992; Queller 1992). More generally, future work should explore how extrinsic factors (e.g., climate change) and intrinsic factors (e.g., social behavior) interact to drive patch heterogeneity and mediate the evolution of helping and harming.

Our model also relates to the “bottleneck” hypothesis in the evolution of multicellularity, which suggests that a unicellular bottleneck, followed by a growth cycle in the absence of migration, has been fundamental for the evolution of multicellular organisms (Dawkins 1982; Maynard Smith 1988; Maynard Smith and Szathmáry 1995; Grosberg and Strathmann 1998, 2007; Bourke 2011). These characteristics of the life cycle already assume preadaptations, which can be social traits themselves. Our model predicts that facultative helping is favored among group-mates after the group has passed through a bottleneck and has grown to full size, and it also predicts that facultative harming is favored before the group has grown to full size (see result 4). Therefore, less stringent bottlenecks may have been important to produce preadaptations before the full onset of highly complex cooperative societies (e.g., multicellular organisms or eusocial insects). Although some empirical work has been done in this area (Brock-

hurst 2007; Brockhurst et al. 2007), the kin-competition effects of population viscosity in the context of bottlenecks have so far been neglected. Experimental protocols that test the effect of population viscosity in social evolution are already available (Griffin et al. 2004; Kümmerli et al. 2009), and these could be extended to incorporate the effects of spatial and temporal patch-size heterogeneity.

Our results show that heterogeneity in patch size and individual fecundity affects the value of each selective force in populations characterized by some degree of viscosity. This happens because (1) it changes the genetic structure of the population and consequently the relatedness of each recipient and (2) it changes the reproductive value of each recipient. Therefore, heterogeneity in patch size and individual fecundity is likely to mediate the evolution of social traits other than those considered here. For example, sex ratio (i.e., the fraction of offspring that are male) is a classic social trait (Hamilton 1967; Charnov 1982; West 2009). Selection pressures acting on sex ratio—including those relating to kin competition—depend on both the relatedness and the reproductive value of the different recipients (males vs. females). Thus, extending our framework to study sex ratio evolution is a promising line of future research.

#### Acknowledgments

We thank A. Grafen and S. West for helpful discussion. A.M.M.R. is supported by the PhD Program in Computational Biology of the Instituto Gulbenkian de Ciência and funded by the Fundação para a Ciência e a Tecnologia (SFRH/BD/33851/2009), and A.G. is supported by research fellowships from Balliol College and the Royal Society.

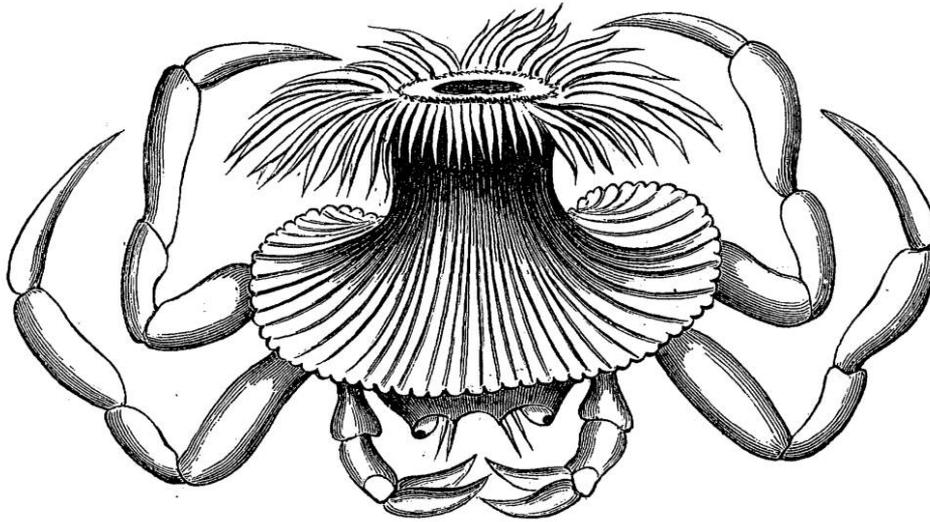
#### Literature Cited

- Alizon, S., and P. D. Taylor. 2008. Empty sites can promote altruistic behavior. *Evolution* 62:1335–1344.
- Begon, M., J. L. Harper, and C. R. Townsend. 2006. *Ecology: individuals, populations and communities*. Blackwell, Oxford.
- Bourke, A. F. G. 2011. *Principles of social evolution*. Oxford University Press, Oxford.
- Brockhurst, M. A. 2007. Population bottlenecks promote cooperation in bacterial biofilms. *PLoS ONE* 2(7):e634.
- Brockhurst, M. A., A. Buckling, and A. Gardner. 2007. Cooperation peaks at intermediate disturbance. *Current Biology* 17:761–765.
- Charnov, E. L. 1982. *The theory of sex allocation*. Princeton University Press, Princeton, NJ.
- Darwin, C. R. 1859. *On the origin of species by means of natural selection*. J. Murray, London.
- Dawkins, R. 1976. *The selfish gene*. Oxford University Press, Oxford.

- . 1982. *The extended phenotype*. Oxford University Press, Oxford.
- El Mouden, C., and A. Gardner. 2008. Nice natives and mean migrants: the evolution of dispersal-dependent social behaviour in viscous populations. *Journal of Evolutionary Biology* 21:1480–1491.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon, Oxford.
- Frank, S. A. 1997. Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. *Journal of Theoretical Biology* 189:307–316.
- . 1998. *Foundations of social evolution*. Princeton University Press, Princeton, NJ.
- Gardner, A. 2010. Sex-biased dispersal of adults mediates the evolution of altruism among juveniles. *Journal of Theoretical Biology* 262:339–345.
- Gardner, A., and S. A. West. 2006. Demography, altruism, and the benefits of budding. *Journal of Evolutionary Biology* 19:1707–1716.
- . 2010. Greenbeards. *Evolution* 64:25–38.
- Grafen, A. 2006. A theory of Fisher's reproductive value. *Journal of Mathematical Biology* 53:15–60.
- . 2007. Detecting kin selection at work using inclusive fitness. *Proceedings of the Royal Society B: Biological Sciences* 274:713–719.
- Griffin, A. S., S. A. West, and A. Buckling. 2004. Cooperation and competition in pathogenic bacteria. *Nature* 430:1024–1027.
- Grosberg, R. K., and R. R. Strathmann. 1998. One cell, two cell, red cell, blue cell: the persistence of a unicellular stage in multicellular life histories. *Trends in Ecology and Evolution* 13:112–116.
- . 2007. The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics* 38:621–654.
- Hamilton, W. D. 1963. The evolution of altruistic behavior. *American Naturalist* 97:354–356.
- . 1964. The genetical evolution of social behaviour. I, II. *Journal of Theoretical Biology* 7:1–16, 17–52.
- . 1967. Extraordinary sex ratios. *Science* 156:477–488.
- . 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* 228:1218–1220.
- . 1971a. Geometry for the selfish herd. *Journal of Theoretical Biology* 31:295–311.
- . 1971b. Selection of selfish and altruistic behavior in some extreme models. Pages 57–91 in J. F. Eisenberg and W. S. Wilson, eds. *Man and beast: comparative social behavior*. Smithsonian Institution, Washington, DC.
- . 1996. *Narrow roads of gene land*. Vol. 1. *Evolution of social behaviour*. W. H. Freeman Spektrum, Oxford.
- Irwin, A., and P. D. Taylor. 2001. The evolution of altruism in a stepping-stone population with overlapping generations. *Theoretical Population Biology* 60:315–325.
- Johnstone, R. A. 2008. Kin selection, local competition, and reproductive skew. *Evolution* 62:2592–2599.
- Johnstone, R. A., and M. A. Cant. 2008. Sex differences in dispersal and the evolution of helping and harming. *American Naturalist* 172:318–330.
- . 2010. The evolution of menopause in cetaceans and humans: the role of demography. *Proceedings of the Royal Society B: Biological Sciences* 277:3765–3771.
- Kelly, J. K. 1992. Kin selection in density regulated populations. *Journal of Theoretical Biology* 157:447–461.
- Kümmerli, R., A. Gardner, S. A. West, and A. S. Griffin. 2009. Limited dispersal, budding dispersal, and cooperation: an experimental study. *Evolution* 63:939–949.
- Lehmann, L. 2007. The evolution of trans-generational altruism: kin selection meets niche construction. *Journal of Evolutionary Biology* 20:181–189.
- . 2010. Space-time relatedness and Hamilton's rule for long-lasting behaviors in viscous populations. *American Naturalist* 175:136–143.
- Lehmann, L., K. Bargum, and M. Reuter. 2006a. An evolutionary analysis of the relationship between spite and altruism. *Journal of Evolutionary Biology* 19:1507–1516.
- Lehmann, L., N. Perrin, and F. Rousset. 2006b. Population demography and the evolution of helping behaviors. *Evolution* 60:1137–1151.
- Lehmann, L., and F. Rousset. 2010. How life history and demography promote or inhibit the evolution of helping behaviours. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:1–19.
- Levins, R. 1968. *Evolution in changing environments: some theoretical explorations*. Princeton University Press, Princeton, NJ.
- Maynard Smith, J. 1988. Evolutionary progress and levels of selection. Pages 219–230 in M. H. Nitecki, ed. *Evolutionary progress*. University of Chicago Press, Chicago.
- Maynard Smith, J., and E. Szathmáry. 1995. *The major transitions in evolution*. Oxford University Press, Oxford.
- McPeck, M. A., and D. H. Holt. 1992. The evolution of dispersal in spatially and temporally varying environments. *American Naturalist* 140:1010–1027.
- Price, G. R. 1970. Selection and covariance. *Nature* 227:520–521.
- Queller, D. C. 1992. Does population viscosity promote kin selection? *Trends in Ecology and Evolution* 7:322–324.
- Rifkin, J. L., C. L. Nunn, and L. S. Garamszegi. 2012. Do animals living in larger groups experience greater parasitism? a meta-analysis. *American Naturalist* 180:70–82.
- Rodrigues, A. M. M., and A. Gardner. 2012. Evolution of helping and harming in heterogeneous populations. *Evolution* 66:2065–2079.
- Rousset, F. 2004. *Genetic structure and selection in subdivided populations*. Princeton University Press, Princeton, NJ.
- Taylor, P. D. 1990. Allele-frequency change in a class-structured population. *American Naturalist* 135:95–106.
- . 1992a. Altruism in viscous populations: an inclusive fitness model. *Evolutionary Ecology* 6:352–356.
- . 1992b. Inclusive fitness in a homogeneous environment. *Proceedings of the Royal Society B: Biological Sciences* 249:299–302.
- . 1996. Inclusive fitness arguments in genetic models of behaviour. *Journal of Mathematical Biology* 34:654–674.
- Taylor, P. D., and S. A. Frank. 1996. How to make a kin selection model. *Journal of Theoretical Biology* 180:27–37.
- Taylor, P. D., and A. J. Irwin. 2000. Overlapping generations can promote altruistic behavior. *Evolution* 54:1135–1141.
- Taylor, P. D., G. Wild, and A. Gardner. 2007. Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of Evolutionary Biology* 20:296–304.
- Trivers, R. 1971. Evolution of reciprocal altruism. *Quarterly Review of Biology* 46:35–57.

- West, S. A. 2009. Sex allocation. Princeton University Press, Princeton, NJ.
- West, S. A., and A. Gardner. 2010. Altruism, spite, and greenbeards. *Science* 327:1341–1344.
- West, S. A., A. S. Griffin, and A. Gardner. 2007*a*. Evolutionary explanations for cooperation. *Current Biology* 17:661–672.
- . 2007*b*. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* 20:415–432.
- West, S. A., I. Pen, and A. S. Griffin. 2002. Cooperation and competition between relatives. *Science* 296:72–75.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Associate Editor: Sean H. Rice  
Editor: Mark A. McPeck



*CancriSOCia expansa* Stimpson. From “Review of Synopsis of the Polyps and Corals of the North Pacific Exploring Expedition, under Commodore C. Ringgold and Capt. John Rodgers, U. S. N., from 1853 to 1856. Collected by Dr. Wm. Stimpson, Naturalist to the Expedition,” by A. E. Verrill (*American Naturalist*, 1870, 4:488–491).

# Appendix from A. M. M. Rodrigues and A. Gardner, “Evolution of Helping and Harming in Viscous Populations When Group Size Varies”

(Am. Nat., vol. 181, no. 5, p. 609)

## Patch Dynamics

We use the transition probabilities in patch quality, given in the main text, to define a transition matrix,

$$\mathbf{P} = \begin{pmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{pmatrix}. \quad (\text{A1})$$

At equilibrium, the proportion of big patches is given by

$$p = \frac{1 - \beta}{2 - (\alpha + \beta)}, \quad (\text{A2})$$

with  $\alpha < 1$  and  $\beta < 1$ . We define a random variable  $X_t$  that characterizes a patch's quality ( $B = 1, L = 0$ ) in any given generation  $t$ . The coefficient of temporal correlation between any two generations is given by  $\tau = \text{Cov}(X_t, X_{t+1}) / (\text{Var}(X_t) \text{Var}(X_{t+1}))^{1/2}$ . In terms of model parameters, this is

$$\tau = \alpha - (1 - \beta) \quad (\text{A3})$$

at ecological equilibrium. We find that the temporal correlation  $\tau$  and the frequency of big patches  $p$  cannot be varied totally independently of each other when the temporal correlation is negative ( $\tau < 0$ ). For example, when  $\tau = -1$ , the frequency of big patches is constrained to be  $p = 1/2$ .

## Reproductive Success

We follow the life cycle described in the main text; however, we make the additional assumptions that a mother in a big patch has a large number  $F_B = f(x_B, y_B)$  of offspring, while a mother in a little patch has a large number  $F_L = \sigma f(x_L, y_L)$ , where  $0 < \sigma < \infty$  is the reproductive factor. This recovers the model of Rodrigues and Gardner (2012) as a special case and hence facilitates connections between the two models. An offspring of a focal breeding mother in a big patch remains in her natal patch with probability  $1 - m$  and subsequently wins a breeding site with probability

$$E_{B \rightarrow B}^* = \frac{f(x_B, y_B)}{(1 - m)f(x_B, Y_B)n_B + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}, \quad (\text{A4})$$

where  $Y_B$  is the average level of the social behavior in the focal patch and  $z_B$  and  $z_L$  are the average levels of the social behavior across the population in big and little patches, respectively. Conversely, with probability  $m$  the offspring migrates to a random patch in the population. With probability  $p$  she reaches a big patch, and she subsequently wins a breeding site with probability

$$E_{B \rightarrow B}^\circ = \frac{f(x_B, y_B)}{(1 - m)f(z_B, z_B)n_B + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}. \quad (\text{A5})$$

With probability  $1 - p$  she reaches a little patch, and she subsequently wins a breeding site with probability

$$E_{B \rightarrow L}^\circ = \frac{f(x_B, y_B)}{(1 - m)\sigma f(z_L, z_L)n_L + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}. \quad (\text{A6})$$

We derive the same expressions for a mother in a little patch. We may then define a matrix

$$\mathbf{E} = \begin{pmatrix} E_{B \rightarrow B}^* & E_{L \rightarrow B} \\ E_{B \rightarrow L}^\circ & E_{L \rightarrow L}^\circ \end{pmatrix} = \begin{pmatrix} (1 - m)E_{B \rightarrow B}^* + mpE_{B \rightarrow B}^\circ & mpE_{L \rightarrow B}^\circ \\ m(1 - p)E_{B \rightarrow L}^\circ & (1 - m)E_{L \rightarrow L}^\circ + m(1 - p)E_{L \rightarrow L}^\circ \end{pmatrix}. \quad (\text{A7})$$

We define a diagonal matrix  $\mathbf{n}$  where the first element of the diagonal is  $n_B$  and the second is  $n_L$ . The matrix that gives the reproductive success of a focal individual for each patch type, in relation to the patch type where their offspring are going to reproduce, is then given by

$$\mathbf{w} = \begin{pmatrix} w_{B \rightarrow B} & w_{L \rightarrow B} \\ w_{B \rightarrow L} & w_{L \rightarrow L} \end{pmatrix} = \mathbf{nPE}. \quad (\text{A8})$$

Thus, the reproductive success of a focal individual is given by

$$W_B = w_{B \rightarrow B} + \frac{v_B}{v_L} w_{B \rightarrow L} \quad (\text{A9})$$

and

$$W_L = w_{L \rightarrow L} + \frac{v_L}{v_B} w_{L \rightarrow B} \quad (\text{A10})$$

for big and little patches, respectively. The successful offspring must be weighted by their reproductive values (Fisher 1930; Taylor 1990; Grafen 2006). The reproductive value of offspring that reproduce in big patches is  $v_B$ , while the reproductive value of offspring that reproduce in little patches is  $v_L$ . The reproductive success of a random focal individual in the population is therefore given by

$$W = c_B W_B + c_L W_L, \quad (\text{A11})$$

where  $c_B$  and  $c_L$  are the class reproductive values of big and little patches, respectively. Class reproductive values are the product of the frequency of individuals in a given class and their individual reproductive values:  $c_B = u_B v_B$  and  $c_L = u_L v_L$  for big and little patches, respectively. Class reproductive values are normalized such that  $c_B + c_L = 1$  (Taylor 1990, 1996; Taylor et al. 2007).

## Hamilton's Rule and Potential for Helping and Harming

We assume that the social behavior is under the control of a single locus. Drawing a gene at random from this locus from a focal individual, we denote its genic value by  $g$ . Assuming vanishingly small genetic variation in the population, the direction of selection acting on the social behavior is given by

$$\frac{dW}{dg} = c_B \frac{dW_B}{dg} + c_L \frac{dW_L}{dg}. \quad (\text{A12})$$

If  $dW/dg > 0$ , then selection favors an increase in the social behavior, and if  $dW/dg < 0$ , then selection favors a decrease in the social behavior. In the right-hand side of equation (A12), the derivative in the first term can be expanded as follows:

$$\frac{dW_B}{dg} = \frac{\partial W_B}{\partial x_B} \frac{dx_B}{dg_B} \frac{dg_B}{dg} + \frac{\partial W_B}{\partial y_B} \frac{dy_B}{dg'_B} \frac{dg'_B}{dg} + \frac{\partial W_B}{\partial Y_B} \frac{dY_B}{dG_B} \frac{dG_B}{dg}. \quad (\text{A13})$$

All the derivatives are evaluated as  $x_B = y_B = Y_B = z_B = z_L$  (Taylor and Frank 1996; Frank 1998). The associations between reproductive success and phenotypes are the effects of the mutant phenotype on the vital rates of the different recipients. The associations between phenotypes and genotype  $\gamma$  constitute the genotype-phenotype mapping, and these can be set to unity:  $dx_B/dg_B = dy_B/dg'_B = Y_B/dG_B = \gamma_B = 1$ . The associations between genotypes and genic values are the coefficients of consanguinity of the different recipients in relation to the actor. We normalize these coefficients of consanguinity, which give the coefficients of relatedness of the different recipients (Bulmer 1994). These are derived in the next section.

## Relatedness

We derive the kin-selection coefficients of relatedness, assuming a neutral population (Wright 1969; Taylor 1992; Rousset 2004). In a focal patch after dispersal, two offspring chosen at random are philopatric to the patch with probability  $h$ .

With probability  $1/n$  they are siblings and have relatedness 1. With probability  $1 - (1/n)$  they are not siblings and they have relatedness  $r$ . Hence, relatedness among two offspring chosen at random in a focal patch is given by

$$Q_B = h_B \left[ \frac{1}{n_B} + \left( 1 - \frac{1}{n_B} \right) r_B \right] \quad (\text{A14})$$

and

$$Q_L = h_L \left[ \frac{1}{n_L} + \left( 1 - \frac{1}{n_L} \right) r_L \right] \quad (\text{A15})$$

for big and little patches, respectively. A focal big patch was big in the previous generation with probability  $\pi_B$ , in which case relatedness among patchmates is  $Q_B$ . The same patch was little in the previous generation with probability  $1 - \pi_B$ , in which case relatedness among patchmates is  $Q_L$ . A focal little patch was little in the previous generation with probability  $\pi_L$ , in which case relatedness among patchmates is  $Q_L$ . The same patch was big with probability  $1 - \pi_L$ , in which case relatedness among patchmates is  $Q_B$ . This means that relatedness among two patchmates in the next generation is given by

$$\begin{aligned} r'_B &= \pi_B Q_B + (1 - \pi_B) Q_L, \\ r'_L &= (1 - \pi_L) Q_B + \pi_L Q_L. \end{aligned} \quad (\text{A16})$$

At equilibrium,  $r'_B = r_B$  and  $r'_L = r_L$ , and we can solve equation system (A16) to obtain the ‘‘other-only’’ kin-selection coefficients of relatedness (Pepper 2000). These are given by

$$r_B = \frac{h_L n_B (1 - \pi_B) + h_B [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]}{n_B [n_L - (n_L - 1) h_L \pi_L] - h_B (n_B - 1) [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]} \quad (\text{A17})$$

and

$$r_L = \frac{h_L n_B \pi_L + h_B [n_L (1 - \pi_L) + h_L (n_B - 1) (1 - \pi_B - \pi_L)]}{n_B [n_L - (n_L - 1) h_L \pi_L] - h_B (n_B - 1) [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]}, \quad (\text{A18})$$

for big and little patches, respectively. The ‘‘whole-group’’ coefficients of relatedness (Pepper 2000) are given by  $R = 1/n + [1 - (1/n)]r$ . Relatedness of a focal actor in relation to a primary recipient is equal to the relatedness among patchmates. Hence,  $r_{P|B} = r_B$  and  $r_{P|L} = r_L$  for big and little patches, respectively. Relatedness of an actor in relation to a secondary recipient depends on the probability of cophlopatry and the ‘‘whole-group’’ relatedness. This is given by  $r_{S|B} = h_B R_B$  and  $r_{S|L} = h_L R_L$  for big and little patches, respectively.

In spatially heterogeneous populations, patches do not undergo changes in size; hence,  $\pi_B = 1$  and  $\pi_L = 1$ . This means that relatedness of primary recipients in big patches is  $r_{P|B} = h_B R_B$ . Likewise, relatedness of primary recipients in little patches is  $r_{P|L} = h_L R_L$ .

## Reproductive Value

The reproductive values of individuals are given by the left eigenvector of matrix  $\mathbf{w}$  (Taylor 1990; Grafen 2006). The individual reproductive value is the compound value of the number of offspring times their reproductive value. Multiplying individual reproductive value by group size gives the patch reproductive values:  $v_B = n_B v_{B_i}$  and  $v_L = n_L v_{L_i}$  for big and little patches, respectively. Relative patch reproductive values are given by

$$v_{S|B} = \frac{\alpha v_B + (1 - \alpha) v_L}{v_B} \quad (\text{A19})$$

and

$$v_{S|L} = \frac{\beta v_L + (1 - \beta) v_B}{v_L}, \quad (\text{A20})$$

for big and little patches, respectively.

## Cancellation of Obligate Helping and Harming

Taylor’s (1992a) cancellation result occurs because, in his model, the relatedness of primary recipients is equal to that of

secondary recipients,  $r_{P|T} = r_{S|T}$ . This occurs because the probability of cophylopatry used in the derivation of relatedness,  $h_T$ , is equal to the scale of competition:  $h_T = a_T$  (Frank 1998). This identity still holds for fluctuating environments:  $h_B = a_B$  and  $h_L = a_L$ . We find additional symmetries that lead to the cancellation result for obligate social behavior. We can write the different quantities in Hamilton's rule using this notation, such that  $\pi_B = P(X_{t-1} = B|X_t = B)$ ,  $1 - \pi_B = P(X_{t-1} = L|X_t = B)$ ,  $\pi_L = P(X_{t-1} = L|X_t = L)$ ,  $1 - \pi_L = P(X_{t-1} = B|X_t = L)$ ,  $\alpha = P(X_{t+1} = B|X_t = B)$ ,  $1 - \alpha = P(X_{t+1} = L|X_t = B)$ ,  $\beta = P(X_{t+1} = L|X_t = L)$ , and  $1 - \beta = P(X_{t+1} = B|X_t = L)$ . As a result, we can write the life-for-life relatedness (Hamilton 1972; Bulmer 1994) as

$$\begin{aligned} \rho_P &= p_B(P(X_{t+1} = B|X_t = B)h_B R_B v_B + P(X_{t+1} = L|X_t = B)h_L R_L v_B) \\ &\quad + p_L(P(X_{t+1} = B|X_t = L)h_B R_B v_L + P(X_{t+1} = L|X_t = L)h_L R_L v_L), \end{aligned} \quad (A21)$$

$$\begin{aligned} \rho_S &= p_B(P(X_{t+1} = B|X_t = B)h_B R_B v_B + P(X_{t+1} = L|X_t = B)h_B R_B v_L) \\ &\quad + p_L(P(X_{t+1} = L|X_t = L)h_L R_L v_L + P(X_{t+1} = B|X_t = L)h_L R_L v_B). \end{aligned} \quad (A22)$$

We can prove that the life-for-life relatedness coefficients cancel each other in a pairwise way. For example, first note that at equilibrium,  $\forall t$ :  $P(X_t = B) = p_B = p$  and  $P(X_t = L) = p_L = 1 - p$ ; second, from Bayes's theorem we find that  $P(X_{t-1} = L|X_t = B) = P(X_t = B|X_{t-1} = L)p_L/p_B$ ; third, because the patch dynamics is described by a time-homogeneous Markov chain, we have  $P(X_t = B|X_{t-1} = L) = P(X_{t+1} = B|X_t = L)$ ; as a consequence, we find that  $p_B P(X_{t-1} = L|X_t = B) = p_L P(X_{t+1} = B|X_t = L)$ . We can generalize this result for all pairwise terms in equations (A14) and (A15). As a result, the value of primary recipients is equal to the value of secondary recipients,  $\rho_P = \rho_S$ , and therefore obligate social behavior is never favored.

## Temporal Heterogeneity

We assume that all patches become big with probability  $p$  or become little with probability  $1 - p$ , independently of the present state of the population (cf. Iizuka 2001; Rousset 2004, ch. 10). As a consequence, at any given generation, the probability that patches are in the big state is  $p$ , while the probability that patches are in the little state is  $1 - p$ . Hence, over time, the expected number of individuals in the big state is  $u_B = pn_B$ , while the expected number of individuals in the little state is  $u_L = (1 - p)n_L$ . Given that the population is in the big state, the probability that it remains big is  $p$ , in which case the fecundity of a focal individual in a neutral population is 1. With probability  $1 - p$ , the population changes to the little state, in which case the fecundity of a focal individual in a neutral population is  $n_L/n_B$ . Hence, at equilibrium, the reproductive value of a focal individual in a big patch is  $v_B = pv_B + (1 - p)(n_L/n_B)v_L$ . This means that  $v_B = (n_L/n_B)v_L$ . We can normalize the reproductive values such that  $v_B = 1/n_B$  and  $v_L = 1/n_L$ . Note that the probability that the population was in the big state in a previous generation is independent of the present state of the population, such that  $\pi_B = 1 - \pi_L = p$  and  $1 - \pi_B = \pi_L = 1 - p$ . As a consequence, relatedness coefficients obey the following identities:  $r_B = phR_B + (1 - p)hR_L$  and  $r_L = phR_B + (1 - p)hR_L$ . Hence,  $r_B = r_L$  and  $r_{P|B} = r_{P|L}$ .

## Literature Cited Only in the Appendix

- Bulmer, M. G. 1994. Theoretical evolutionary ecology. Sinauer, Sunderland, MA.
- Hamilton, W. D. 1972. Altruism and related phenomena, mainly in social insects. Annual Review of Ecology and Systematics 3:193–232.
- Iizuka, M. 2001. The effective size of fluctuating populations. Theoretical Population Biology 59:281–286.
- Pepper, J. W. 2000. Relatedness in trait group models of social evolution. Journal of Theoretical Biology 206:355–368.
- Wright, S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.

# Appendix from A. M. M. Rodrigues and A. Gardner, “Evolution of Helping and Harming in Viscous Populations When Group Size Varies”

(Am. Nat., vol. 181, no. 5, p. 609)

## Patch Dynamics

We use the transition probabilities in patch quality, given in the main text, to define a transition matrix,

$$\mathbf{P} = \begin{pmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{pmatrix}. \quad (\text{A1})$$

At equilibrium, the proportion of big patches is given by

$$p = \frac{1 - \beta}{2 - (\alpha + \beta)}, \quad (\text{A2})$$

with  $\alpha < 1$  and  $\beta < 1$ . We define a random variable  $X_t$  that characterizes a patch's quality ( $B = 1, L = 0$ ) in any given generation  $t$ . The coefficient of temporal correlation between any two generations is given by  $\tau = \text{Cov}(X_t, X_{t+1})/(\text{Var}(X_t) \text{Var}(X_{t+1}))^{1/2}$ . In terms of model parameters, this is

$$\tau = \alpha - (1 - \beta) \quad (\text{A3})$$

at ecological equilibrium. We find that the temporal correlation  $\tau$  and the frequency of big patches  $p$  cannot be varied totally independently of each other when the temporal correlation is negative ( $\tau < 0$ ). For example, when  $\tau = -1$ , the frequency of big patches is constrained to be  $p = 1/2$ .

## Reproductive Success

We follow the life cycle described in the main text; however, we make the additional assumptions that a mother in a big patch has a large number  $F_B = f(x_B, y_B)$  of offspring, while a mother in a little patch has a large number  $F_L = \sigma f(x_L, y_L)$ , where  $0 < \sigma < \infty$  is the reproductive factor. This recovers the model of Rodrigues and Gardner (2012) as a special case and hence facilitates connections between the two models. An offspring of a focal breeding mother in a big patch remains in her natal patch with probability  $1 - m$  and subsequently wins a breeding site with probability

$$E_{B \rightarrow B}^* = \frac{f(x_B, y_B)}{(1 - m)f(x_B, Y_B)n_B + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}, \quad (\text{A4})$$

where  $Y_B$  is the average level of the social behavior in the focal patch and  $z_B$  and  $z_L$  are the average levels of the social behavior across the population in big and little patches, respectively. Conversely, with probability  $m$  the offspring migrates to a random patch in the population. With probability  $p$  she reaches a big patch, and she subsequently wins a breeding site with probability

$$E_{B \rightarrow B}^\circ = \frac{f(x_B, y_B)}{(1 - m)f(z_B, z_B)n_B + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}. \quad (\text{A5})$$

With probability  $1 - p$  she reaches a little patch, and she subsequently wins a breeding site with probability

$$E_{B \rightarrow L}^\circ = \frac{f(x_B, y_B)}{(1 - m)\sigma f(z_L, z_L)n_L + m[pf(z_B, z_B)n_B + (1 - p)\sigma f(z_L, z_L)n_L]}. \quad (\text{A6})$$

We derive the same expressions for a mother in a little patch. We may then define a matrix

$$\mathbf{E} = \begin{pmatrix} E_{B \rightarrow B}^* & E_{L \rightarrow B} \\ E_{B \rightarrow L}^\circ & E_{L \rightarrow L}^\circ \end{pmatrix} = \begin{pmatrix} (1 - m)E_{B \rightarrow B}^* + mpE_{B \rightarrow B}^\circ & mpE_{L \rightarrow B}^\circ \\ m(1 - p)E_{B \rightarrow L}^\circ & (1 - m)E_{L \rightarrow L}^\circ + m(1 - p)E_{L \rightarrow L}^\circ \end{pmatrix}. \quad (\text{A7})$$

We define a diagonal matrix  $\mathbf{n}$  where the first element of the diagonal is  $n_B$  and the second is  $n_L$ . The matrix that gives the reproductive success of a focal individual for each patch type, in relation to the patch type where their offspring are going to reproduce, is then given by

$$\mathbf{w} = \begin{pmatrix} w_{B \rightarrow B} & w_{L \rightarrow B} \\ w_{B \rightarrow L} & w_{L \rightarrow L} \end{pmatrix} = \mathbf{nPE}. \quad (\text{A8})$$

Thus, the reproductive success of a focal individual is given by

$$W_B = w_{B \rightarrow B} + \frac{v_B}{v_L} w_{B \rightarrow L} \quad (\text{A9})$$

and

$$W_L = w_{L \rightarrow L} + \frac{v_L}{v_B} w_{L \rightarrow B} \quad (\text{A10})$$

for big and little patches, respectively. The successful offspring must be weighted by their reproductive values (Fisher 1930; Taylor 1990; Grafen 2006). The reproductive value of offspring that reproduce in big patches is  $v_B$ , while the reproductive value of offspring that reproduce in little patches is  $v_L$ . The reproductive success of a random focal individual in the population is therefore given by

$$W = c_B W_B + c_L W_L, \quad (\text{A11})$$

where  $c_B$  and  $c_L$  are the class reproductive values of big and little patches, respectively. Class reproductive values are the product of the frequency of individuals in a given class and their individual reproductive values:  $c_B = u_B v_B$  and  $c_L = u_L v_L$  for big and little patches, respectively. Class reproductive values are normalized such that  $c_B + c_L = 1$  (Taylor 1990, 1996; Taylor et al. 2007).

## Hamilton's Rule and Potential for Helping and Harming

We assume that the social behavior is under the control of a single locus. Drawing a gene at random from this locus from a focal individual, we denote its genic value by  $g$ . Assuming vanishingly small genetic variation in the population, the direction of selection acting on the social behavior is given by

$$\frac{dW}{dg} = c_B \frac{dW_B}{dg} + c_L \frac{dW_L}{dg}. \quad (\text{A12})$$

If  $dW/dg > 0$ , then selection favors an increase in the social behavior, and if  $dW/dg < 0$ , then selection favors a decrease in the social behavior. In the right-hand side of equation (A12), the derivative in the first term can be expanded as follows:

$$\frac{dW_B}{dg} = \frac{\partial W_B}{\partial x_B} \frac{dx_B}{dg_B} \frac{dg_B}{dg} + \frac{\partial W_B}{\partial y_B} \frac{dy_B}{dg'_B} \frac{dg'_B}{dg} + \frac{\partial W_B}{\partial Y_B} \frac{dY_B}{dG_B} \frac{dG_B}{dg}. \quad (\text{A13})$$

All the derivatives are evaluated as  $x_B = y_B = Y_B = z_B = z_L$  (Taylor and Frank 1996; Frank 1998). The associations between reproductive success and phenotypes are the effects of the mutant phenotype on the vital rates of the different recipients. The associations between phenotypes and genotype  $\gamma$  constitute the genotype-phenotype mapping, and these can be set to unity:  $dx_B/dg_B = dy_B/dg'_B = Y_B/dG_B = \gamma_B = 1$ . The associations between genotypes and genic values are the coefficients of consanguinity of the different recipients in relation to the actor. We normalize these coefficients of consanguinity, which give the coefficients of relatedness of the different recipients (Bulmer 1994). These are derived in the next section.

## Relatedness

We derive the kin-selection coefficients of relatedness, assuming a neutral population (Wright 1969; Taylor 1992; Rousset 2004). In a focal patch after dispersal, two offspring chosen at random are philopatric to the patch with probability  $h$ .

With probability  $1/n$  they are siblings and have relatedness 1. With probability  $1 - (1/n)$  they are not siblings and they have relatedness  $r$ . Hence, relatedness among two offspring chosen at random in a focal patch is given by

$$Q_B = h_B \left[ \frac{1}{n_B} + \left( 1 - \frac{1}{n_B} \right) r_B \right] \quad (\text{A14})$$

and

$$Q_L = h_L \left[ \frac{1}{n_L} + \left( 1 - \frac{1}{n_L} \right) r_L \right] \quad (\text{A15})$$

for big and little patches, respectively. A focal big patch was big in the previous generation with probability  $\pi_B$ , in which case relatedness among patchmates is  $Q_B$ . The same patch was little in the previous generation with probability  $1 - \pi_B$ , in which case relatedness among patchmates is  $Q_L$ . A focal little patch was little in the previous generation with probability  $\pi_L$ , in which case relatedness among patchmates is  $Q_L$ . The same patch was big with probability  $1 - \pi_L$ , in which case relatedness among patchmates is  $Q_B$ . This means that relatedness among two patchmates in the next generation is given by

$$r'_B = \pi_B Q_B + (1 - \pi_B) Q_L, \quad (\text{A16})$$

$$r'_L = (1 - \pi_L) Q_B + \pi_L Q_L.$$

At equilibrium,  $r'_B = r_B$  and  $r'_L = r_L$ , and we can solve equation system (A16) to obtain the ‘‘other-only’’ kin-selection coefficients of relatedness (Pepper 2000). These are given by

$$r_B = \frac{h_L n_B (1 - \pi_B) + h_B [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]}{n_B [n_L - (n_L - 1) h_L \pi_L] - h_B (n_B - 1) [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]} \quad (\text{A17})$$

and

$$r_L = \frac{h_L n_B \pi_L + h_B [n_L (1 - \pi_L) + h_L (n_B - 1) (1 - \pi_B - \pi_L)]}{n_B [n_L - (n_L - 1) h_L \pi_L] - h_B (n_B - 1) [n_L \pi_B + h_L (n_L - 1) (1 - \pi_B - \pi_L)]}, \quad (\text{A18})$$

for big and little patches, respectively. The ‘‘whole-group’’ coefficients of relatedness (Pepper 2000) are given by  $R = 1/n + [1 - (1/n)]r$ . Relatedness of a focal actor in relation to a primary recipient is equal to the relatedness among patchmates. Hence,  $r_{P|B} = r_B$  and  $r_{P|L} = r_L$  for big and little patches, respectively. Relatedness of an actor in relation to a secondary recipient depends on the probability of cophlopatry and the ‘‘whole-group’’ relatedness. This is given by  $r_{S|B} = h_B R_B$  and  $r_{S|L} = h_L R_L$  for big and little patches, respectively.

In spatially heterogeneous populations, patches do not undergo changes in size; hence,  $\pi_B = 1$  and  $\pi_L = 1$ . This means that relatedness of primary recipients in big patches is  $r_{P|B} = h_B R_B$ . Likewise, relatedness of primary recipients in little patches is  $r_{P|L} = h_L R_L$ .

## Reproductive Value

The reproductive values of individuals are given by the left eigenvector of matrix  $\mathbf{w}$  (Taylor 1990; Grafen 2006). The individual reproductive value is the compound value of the number of offspring times their reproductive value. Multiplying individual reproductive value by group size gives the patch reproductive values:  $v_B = n_B v_{B|B}$  and  $v_L = n_L v_{L|L}$  for big and little patches, respectively. Relative patch reproductive values are given by

$$v_{S|B} = \frac{\alpha v_B + (1 - \alpha) v_L}{v_B} \quad (\text{A19})$$

and

$$v_{S|L} = \frac{\beta v_L + (1 - \beta) v_B}{v_L}, \quad (\text{A20})$$

for big and little patches, respectively.

## Cancellation of Obligate Helping and Harming

Taylor’s (1992a) cancellation result occurs because, in his model, the relatedness of primary recipients is equal to that of

secondary recipients,  $r_{P|T} = r_{S|T}$ . This occurs because the probability of cophylopatry used in the derivation of relatedness,  $h_T$ , is equal to the scale of competition:  $h_T = a_T$  (Frank 1998). This identity still holds for fluctuating environments:  $h_B = a_B$  and  $h_L = a_L$ . We find additional symmetries that lead to the cancellation result for obligate social behavior. We can write the different quantities in Hamilton's rule using this notation, such that  $\pi_B = P(X_{t-1} = B|X_t = B)$ ,  $1 - \pi_B = P(X_{t-1} = L|X_t = B)$ ,  $\pi_L = P(X_{t-1} = L|X_t = L)$ ,  $1 - \pi_L = P(X_{t-1} = B|X_t = L)$ ,  $\alpha = P(X_{t+1} = B|X_t = B)$ ,  $1 - \alpha = P(X_{t+1} = L|X_t = B)$ ,  $\beta = P(X_{t+1} = L|X_t = L)$ , and  $1 - \beta = P(X_{t+1} = B|X_t = L)$ . As a result, we can write the life-for-life relatedness (Hamilton 1972; Bulmer 1994) as

$$\begin{aligned} \rho_P &= p_B(P(X_{t+1} = B|X_t = B)h_B R_B v_B + P(X_{t+1} = L|X_t = B)h_L R_L v_B) \\ &\quad + p_L(P(X_{t+1} = B|X_t = L)h_B R_B v_L + P(X_{t+1} = L|X_t = L)h_L R_L v_L), \end{aligned} \quad (A21)$$

$$\begin{aligned} \rho_S &= p_B(P(X_{t+1} = B|X_t = B)h_B R_B v_B + P(X_{t+1} = L|X_t = B)h_B R_B v_L) \\ &\quad + p_L(P(X_{t+1} = L|X_t = L)h_L R_L v_L + P(X_{t+1} = B|X_t = L)h_L R_L v_B). \end{aligned} \quad (A22)$$

We can prove that the life-for-life relatedness coefficients cancel each other in a pairwise way. For example, first note that at equilibrium,  $\forall t$ :  $P(X_t = B) = p_B = p$  and  $P(X_t = L) = p_L = 1 - p$ ; second, from Bayes's theorem we find that  $P(X_{t-1} = L|X_t = B) = P(X_t = B|X_{t-1} = L)p_L/p_B$ ; third, because the patch dynamics is described by a time-homogeneous Markov chain, we have  $P(X_t = B|X_{t-1} = L) = P(X_{t+1} = B|X_t = L)$ ; as a consequence, we find that  $p_B P(X_{t-1} = L|X_t = B) = p_L P(X_{t+1} = B|X_t = L)$ . We can generalize this result for all pairwise terms in equations (A14) and (A15). As a result, the value of primary recipients is equal to the value of secondary recipients,  $\rho_P = \rho_S$ , and therefore obligate social behavior is never favored.

## Temporal Heterogeneity

We assume that all patches become big with probability  $p$  or become little with probability  $1 - p$ , independently of the present state of the population (cf. Iizuka 2001; Rousset 2004, ch. 10). As a consequence, at any given generation, the probability that patches are in the big state is  $p$ , while the probability that patches are in the little state is  $1 - p$ . Hence, over time, the expected number of individuals in the big state is  $u_B = pn_B$ , while the expected number of individuals in the little state is  $u_L = (1 - p)n_L$ . Given that the population is in the big state, the probability that it remains big is  $p$ , in which case the fecundity of a focal individual in a neutral population is 1. With probability  $1 - p$ , the population changes to the little state, in which case the fecundity of a focal individual in a neutral population is  $n_L/n_B$ . Hence, at equilibrium, the reproductive value of a focal individual in a big patch is  $v_B = pv_B + (1 - p)(n_L/n_B)v_L$ . This means that  $v_B = (n_L/n_B)v_L$ . We can normalize the reproductive values such that  $v_B = 1/n_B$  and  $v_L = 1/n_L$ . Note that the probability that the population was in the big state in a previous generation is independent of the present state of the population, such that  $\pi_B = 1 - \pi_L = p$  and  $1 - \pi_B = \pi_L = 1 - p$ . As a consequence, relatedness coefficients obey the following identities:  $r_B = phR_B + (1 - p)hR_L$  and  $r_L = phR_B + (1 - p)hR_L$ . Hence,  $r_B = r_L$  and  $r_{P|B} = r_{P|L}$ .

## Literature Cited Only in the Appendix

- Bulmer, M. G. 1994. Theoretical evolutionary ecology. Sinauer, Sunderland, MA.
- Hamilton, W. D. 1972. Altruism and related phenomena, mainly in social insects. Annual Review of Ecology and Systematics 3:193–232.
- Iizuka, M. 2001. The effective size of fluctuating populations. Theoretical Population Biology 59:281–286.
- Pepper, J. W. 2000. Relatedness in trait group models of social evolution. Journal of Theoretical Biology 206:355–368.
- Wright, S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press, Chicago.