

8-1-2017

# Evolution of delayed dispersal and subsequent emergence of helping, with implications for cooperative breeding.

Geoff Wild

Judith Korb

Follow this and additional works at: <https://ir.lib.uwo.ca/apmathspub>

 Part of the [Applied Mathematics Commons](#), [Behavior and Ethology Commons](#), and the [Evolution Commons](#)

---

## Citation of this paper:

Wild, Geoff and Korb, Judith, "Evolution of delayed dispersal and subsequent emergence of helping, with implications for cooperative breeding." (2017). *Applied Mathematics Publications*. 6.  
<https://ir.lib.uwo.ca/apmathspub/6>

# Evolution of Delayed Dispersal and Subsequent Emergence of Helping, with Implications for Cooperative Breeding

Geoff Wild

*Department of Applied Mathematics, The University of Western Ontario  
London, Ontario, N6A 5B7, Canada  
gwild@uwo.ca, corresponding author*

Judith Korb

*Institut für Biologie I (Zoologie), Universität Freiburg  
Hauptstrasse 1, D-79104 Freiburg, Germany  
judith.korb@biologie.uni-freiburg.de*

---

## Abstract

Cooperative breeding occurs when individuals help raise the offspring of others. It is widely accepted that help displayed by cooperative breeders emerged only after individuals' tendency to delay dispersal had become established. We use this idea as a basis for two inclusive-fitness models: one for the evolution of delayed dispersal, and a second for the subsequent emergence of helpful behaviour exhibited by non-breeding individuals. We focus on a territorial species in a saturated environment, and allow territories to be inherited by non-breeding individuals who have delayed dispersal. Our first model predicts that increased survivorship and increased fecundity both provide an incentive to non-breeding individuals to delay dispersal, and stay near their natal territory for some period of time. Predictions from the first model can be well understood by ignoring complications arising from competition among relatives. Our second model shows that effects on relatives play a primary role in the advantage of helping. In addition, the second model predicts that increased survivorship and fecundity promote the emergence of help. Together, our models lead us to conclude that the emergence of cooperative-breeding systems is made easier by life-history features associated with high survivorship and fecundity. We discuss the implications of our conclusions for life-history-based hypotheses of cooperative breeding and social evolution.

*Keywords:* alloparental care, helpers-at-the-nest, kin selection, model, social evolution

---

<sup>☆</sup>Final corrected version of this preprint can be found in *Journal of Theoretical Biology* 427:53–64, <https://doi.org/10.1016/j.jtbi.2017.05.038>.

## 1. Introduction

In many cooperatively breeding species, subordinate individuals postpone or even forgo their own reproduction in order to promote the reproductive success of the dominant individuals who occupy breeding territories. Familiar examples include the Florida Scrub Jay (*Aphelocoma coerulescens*), and meerkats (*Suricata suricatta*) (Cockburn, 1998; Clutton-Brock, 2002). Uncovering the adaptive significance of the helpful behaviour displayed by these and other cooperative breeders is key to understanding the evolution of animal societies more broadly.

Attempts to explain the selective advantage of cooperative breeding have emphasized a range of complementary influences (Emlen, 1994). Some have focused on the personal fitness benefits of helping, such as delayed reciprocity (Wiley and Rabenold, 1984), and territory inheritance (Stacey and Ligon, 1991). Others have focused on the indirect fitness benefits of raising related but non-descendant offspring, especially as a response to various environmental constraints (Skutch, 1961; Emlen, 1982a,b).

As plausible as the various explanations for cooperative breeding are, studies in the field have failed to uncover a general effect of species' ecology that is also consistent with the variety of theoretical predictions made (Arnold and Owens, 1998). Consequently, recent research has looked more deeply into the specific life-history features that might promote cooperative breeding (Arnold and Owens, 1998; Beauchamp, 2014). Efforts there have centred on finding evidence, among cooperatively breeding species, for the predominance of similar "K-selected" life-history traits. These traits include high survivorship and low fecundity—traits that are thought to be advantageous in environmentally constrained populations, near carrying capacity. Unfortunately, empirical support for the existing life-history based theories has also been mixed (Hatchwell and Komdeur, 2000).

Despite the lack of empirical support, it seems inappropriate to dismiss life-history-based explanations for cooperative breeding outright, since they are theoretically underdeveloped. That is to say, the explanations are based on generic ideas about life-history evolution, rather than on models tailored for cooperatively breeding species. In this paper, we use simple, yet reasonably comprehensive, models to capture the evolution of cooperative breeding, with the goal of clearly assessing the influence of basic life-history on the origin of these systems.

Biologists have long agreed that the evolution of delayed dispersal of offspring from their

natal site is an important first step in the establishment of cooperative-breeding systems (Brown, 1974; Emlen, 1982a; Koenig et al., 1992; Hatchwell and Komdeur, 2000; Kokko and Lundberg, 2001; Kokko and Ekman, 2002). Evidence also suggests that the evolution of delayed dispersal preceded the emergence of helping in certain eusocial insects (Thorne, 1997). Our approach, therefore, is to separate an individual’s dispersal decision from its decision to help, in contrast to many previous models (Motro, 1993; Pen and Weissing, 2000; Leggett et al., 2012; McLeod and Wild, 2013; Wild and Koykka, 2014). We implement this approach by constructing and analysing a model for the evolution of delayed dispersal (Model I). Results from this model are then used to inform a second model for the evolution of helping following establishment of delayed dispersal (Model II).

Biologists have also long recognized the importance of genetic relatedness among cooperative breeders (Emlen, 1982b; Brown, 1987). Unlike those who incorporate relatedness into their theory as a fixed parameter (Kokko and Johnstone, 1999; Kokko et al., 2001), we model relatedness as a function of species’ population dynamics, which ultimately depends on life-history details. Furthermore, our consideration of relatedness is explicit and allows for inbreeding within groups to build, and so our models differ from recent theoretical work (Koykka and Wild, 2015).

Overall, our models predict that cooperative breeding is promoted as survivorship, or fecundity (or both) is increased. We discuss the implications of our predictions in the final section, paying particular attention to life-history based explanations of cooperative breeding and the evolution of social insects.

## **2. Model I: Delayed-Dispersal**

It is a challenge to build a simple model of the evolution of delayed dispersal, since so many factors (both ecological and social) come into play (Kokko and Ekman, 2002). Here, we concentrate on building a model that is tractable but still reflects key aspects of relevant biological systems. As the reader will see, this has required us to sometimes balance our interest in biological realism against our need for a set of assumptions that help to keep our models mathematically tractable. For the reader’s convenience, we summarize all mathematical notation introduced below in Table 1.

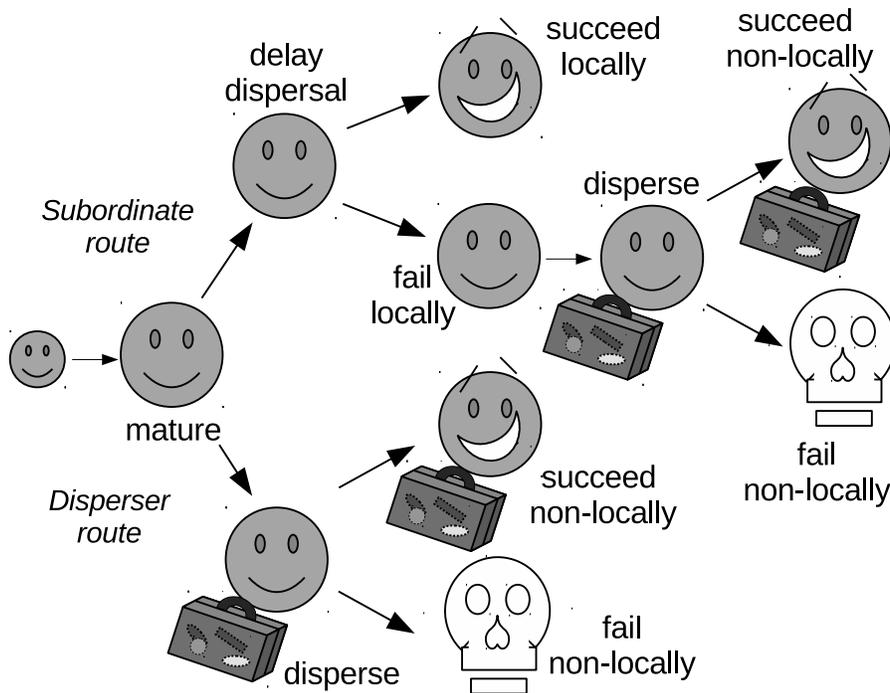
### *2.1. Monomorphic Wild-type Population*

We first consider a genetically monomorphic population of haploid individuals. This “wild-type” population will serve as the backdrop against which we later measure the fitness of mutant individuals. We assume a haploid genetic system because it is simple, and because it mirrors a diploid system with additive interactions between alleles (Johnstone and Cant, 2008).

We track our model wild-type population in discrete time, observing it at the beginning of each time period, e.g. season, year, generation (Figure 1). Each individual in this population begins its life as one of  $b \geq 2$  offspring produced by the two dominant breeders occupying one of a very large number of breeding territories. In keeping with our use of dominant individuals as stand-ins for dominant males and females, respectively, we assume that an offspring inherits all of its genetic material from one or the other dominant on its natal territory, independently, with one-half probability.

As the reader will see, our decision to allow two dominant individuals on a territory means that, in general, genetic lineages experience some degree of inbreeding. In turn, the possibility of inbreeding implies that relatedness among family members is not fixed, but rather can vary in a way that is influenced by the details of the life history. As we have said, this approach to modelling relatedness has not been used by previous authors (e.g. Kokko and Johnstone, 1999; Kokko et al., 2001; Wild and Koykka, 2014), even though it reflects a key means through which life-history features can act to influence the evolution of cooperative breeding.

Following birth, each individual offspring matures, and then does one of two things. With probability  $z$  each remains on its natal territory, delaying dispersal for exactly one time period in an attempt to compete locally for a breeding opportunity, should such an opportunity arise. With probability  $1 - z$  each disperses from its natal territory and attempts to compete non-locally for breeding opportunities, wherever these might arise. For convenience we will refer to  $z$  as an individual’s “tendency to delay dispersal.” Those individuals who delay dispersal we call “subordinates” as long as they are on their natal territory, in order to distinguish them from the territory’s two dominant owners. Those who disperse we call “dispersers.” For now, subordinate individuals offer no help to dominants.



**Figure 1:** Cartoon depiction of the possible fate of individuals in our models. Verbal descriptions of events are given in Section 2.1.

*Subordinate route.* An individual who delays its dispersal will compete for local breeding opportunities when these arise. If  $s_\alpha$  is the survival rate of dominants, then, on average, breeding opportunities arise on a given territory at the rate of  $2(1 - s_\alpha)$  per time period. That is to say, breeding opportunities follow the death of a dominant. We assume subordinates, by virtue of the fact that they are native to a territory, always outcompete dispersers for local breeding opportunities. This reflects what is understood to be a key incentive for cooperative breeding (Kokko and Lundberg, 2001; Kokko and Ekman, 2002; Wild and Koykka, 2014).

It should be emphasized that there are downsides to the priority access we afford subordinates. First, subordinates are not guaranteed to survive to compete for the local opportunities that may or may not arise. We use  $s_\beta$  to denote the probability with which a subordinate survives to compete, and so  $1 - s_\beta$  gives the probability with which a subordinate dies before any local competition can occur. Second, subordinates compete against one another on an equal footing (we assume no queuing of subordinates). Since subordinate competitors are all native to a given territory they are genetically related. Consequently, success as a subordinate may come at the expense of a relative, and we use  $c$  to denote the

probability with which this occurs.

Conditioned on its survival, then, a subordinate competes successfully to become a dominant with probability  $m_\alpha$  (or, equivalently, competes for  $m_\alpha$  breeding opportunities). Again conditioned on survival, the same subordinate fails to compete successfully with probability  $m_\beta = 1 - m_\alpha$ . In the latter case, it remains a subordinate until the following time period, when it disperses, and so becomes a “disperser,” with probability 1 (see below). Complete dispersal of subordinates is an assumption that allows us to keep the model simple (it allows us to circumvent the need for classifying territories using a potentially large state space). It is, however, also an assumption that is reflected in the delayed dispersal of the cooperative breeding corvids like the Florida Scrub Jay (Brown, 1974), and the Gray Jay (Waite and Strickland, 1997).

In Appendix A, we develop expressions for  $m_\alpha$ ,  $m_\beta$ , and  $c$  when the wild-type population has reached equilibrium, as it is the equilibrium population in which we later assume mutant behaviour arises (see below, and also Taylor and Frank, 1996). This same assumption can be found in other approaches to modelling behaviour (e.g. Geritz, 2005).

To emphasize the fact that  $m_\alpha$ ,  $m_\beta$ , and  $c$  are calculated at equilibrium, we re-write them as  $\hat{m}_\alpha$ ,  $\hat{m}_\beta$ ,  $\hat{c}$ , respectively. In Appendix A, we also develop an expression for the equilibrium level of genetic relatedness between two offspring born on the same territory during the same time period, denoted  $\hat{r}$ ; i.e. relatedness between potential local competitors.

*Disperser route.* A disperser – either an individual who became a disperser immediately following its birth, or an individual who had been a subordinate on its natal territory, but then failed to secure a breeding opportunity there – competes for breeding opportunities on territories other than the one on which it was born. We assume that each disperser survives dispersal with probability  $s_\omega$ , and, conditioned on its survival, expects to compete successfully for  $n_\alpha$  breeding opportunities. In Appendix A, we provide an expression for  $n_\alpha$  at equilibrium, and we write its equilibrium value as  $\hat{n}_\alpha$ .

We make two technical assumptions in this part of our model. First, we assume that dispersers cannot secure breeding opportunities in time periods other than the one in which they disperse. This can be thought of, effectively, as assuming that a disperser who fails to secure a breeding opportunity dies. Second, we assume that dispersers fill all dominant positions not filled by subordinates. This means that, in principle,  $n_\alpha$  can exceed 1 (our

model for competition uses a typical “sampling with replacement” scheme; see Appendix A). More importantly, it also means that breeding habitat is saturated. In this way the model incorporates ideas from early theory (Brown, 1974; Emlen, 1982a,b), as well as evidence from cooperatively breeding birds in the field (Pruett-Jones and Lewis, 1990).

## 2.2. The Inclusive-Fitness Effect of Delayed Dispersal

To understand how selection acts to change the tendency to delay dispersal, we use a version of Hamilton’s rule (Hamilton, 1964) that reflects the assumptions laid out above. Specifically, we consider a mutant allele that increases the tendency of its immature carrier to delay dispersal, and we develop an expression for the effect the allele has on its carrier’s inclusive fitness. We stress that this inclusive-fitness effect is calculated under the assumption that the population is at equilibrium, and that, relative to the wild-type, the mutant allele changes the tendency to delay dispersal by a small amount only.

We use  $\Delta w(z)$  to denote the inclusive-fitness effect of the mutant allele, where  $z$  now reflects the population-wide average tendency to delay dispersal (Taylor and Frank, 1996). When  $\Delta w(z)$  is positive (resp. negative) a mutant individual that increases (resp. decreases) its tendency to delay dispersal puts its genetic lineage at an advantage. It follows that when  $\Delta w(z)$  is positive (resp. negative) selection favours an increase (resp. decrease) in the tendency to delay, and so the sign of  $\Delta w(z)$  reflects the direction of the selection gradient acting on  $z$ . In Appendix B, we use the approach described in Taylor and Frank (1996) to show that

$$\Delta w(z) = \underbrace{s_\beta \hat{m}_\alpha v_\alpha + s_\beta \hat{m}_\beta v_\beta - s_\omega \hat{n}_\alpha v_\alpha}_{\text{direct effects}} - \underbrace{s_\beta \hat{m}_\alpha \hat{c}(v_\alpha - v_\beta)}_{\text{indirect effects}} \hat{r} \quad (1)$$

where  $v_\alpha = 1$  is the reproductive value of a dominant individual, and  $v_\beta = s_\omega \hat{n}_\alpha v_\alpha$  is the reproductive value of a subordinate individual. Reproductive value measures the extent of an individual’s genetic contribution to generations in the long-run. Here, reproductive value serves both as an “exchange rate” between the evolutionary significance of dominants and subordinates, respectively, and as the “gold standard” measure of evolutionary success. It is worth noting at this point that terms in (1) depend on life-history parameters  $b$ ,  $s_\alpha$ , and  $s_\beta$ , but  $s_\omega$  cancels out of equation (1).

Equation (1) has a straightforward biological interpretation. To see what this interpretation entails, consider an individual who (for lack of a better term) intended to disperse

immediately after its birth, but remained on its natal site as a subordinate for one time period instead. Now ask, how has this decision altered the evolutionary success of this individual's genetic lineage?

First, by staying on its natal site, the individual in question either survives to become a dominant individual (probability  $s_\beta \hat{m}_\alpha$ ), survives and remains a subordinate ( $s_\beta \hat{m}_\beta$ ), or fails to survive (probability  $1 - s_\beta$ ). The value of a dominant individual to its genetic lineage is  $v_\alpha$ , the value of a subordinate to its genetic lineage is  $v_\beta$ , and the value of a dead individual to its genetic lineage is, of course, zero. Combining this information shows us that, by staying, the individual in question increases the success of its lineage by  $s_\beta \hat{m}_\alpha v_\alpha + s_\beta \hat{m}_\beta v_\beta$ , which is represented by the first two terms of equation (1).

Second, by staying on its natal site rather than leaving, the individual in question has forfeited fitness gains it might have otherwise made elsewhere in the population. Those gains would have required the individual to survive, and would have included gains made by winning a dominant position. The amount of evolutionary success forfeited, in this case, is therefore equal to  $s_\omega \hat{n}_\alpha v_\alpha$ , which is the third term in equation (1).

Third, the decision to stay has implications for other individuals belonging to the same lineage as the individual in question. By staying for one additional time period, the individual in question has increased the amount of competition among kin for local dominant vacancies. Specifically, any success this individual has in securing a dominant position (we can quantify this success as  $\hat{m}_\alpha$ ) comes at the expense of one other subordinate with probability  $\hat{c}$ . The displaced subordinate would have been valued by its lineage in the amount  $v_\alpha$ , but is now valued at  $v_\beta$ . Moreover, the lineage to which the displaced subordinate belongs is the same as that to which our focal individual belongs with probability  $\hat{r}$ . It follows that, by staying, the individual in question further reduces the success of its own lineage by  $\hat{m}_\alpha \hat{c} (v_\alpha - v_\beta) \hat{r}$ , which is the fourth term in (1).

Overall, we can group terms in (1) in the usual way (Brown, 1987); that is, according to whether they represent gains/losses made through personal survival and reproduction (“direct effects” on fitness), or whether they represent gains/losses made through the production of non-descendant kin (“indirect effects” on fitness).

### 3. Method of Analysis and Results I: Delayed-Dispersal

#### 3.1. Simulation algorithm

We analysed our delayed-dispersal model using a numerical simulation of the evolution of  $z$  (Figure 2). The simulation relied on the fact that the sign of  $\Delta w(z)$  indicates the direction in which selection acts to change  $z$  (Rousset, 2004). As previously mentioned, if  $\Delta w(z) > 0$  then  $z$  is increasing over time, and if  $\Delta w(z) < 0$  then  $z$  is decreasing over time. Of course, if  $\Delta w(z) = 0$  then  $z$  is at an evolutionary steady state.

The simulation, itself, consisted of two replicate populations. One population initially had a  $z$  near zero, while the other initially had a  $z$  near one. All other life-history parameters in the two replicates were set to identical values. Values of  $z$  in each replicate population were updated, independently, by adding a multiple of  $\Delta w(z)$  to the most recent rate. Updating continued until  $z$  values in the replicate populations were sufficiently close to one another. By using two replicate populations, then, we were able to use a clear stopping criterion for our simulation. Overall, our simulation produced a prediction for the stable tendency to delay dispersal (the stable  $z$ ) under given set of life-history conditions.

#### 3.2. Simulation-based predictions

Intermediate phenotypes ( $0 < z < 1$ ) are predicted to be stable in only a narrow range of parameter space. As shown in Figure 3a-c, outside of this narrow range selection results in either (1) all immature individuals delay dispersal, and forgo the chance to mature and reproduce independently for one time period ( $z = 1$ ), or (2) all immature dispersing in the same year in which they were born, mature and attempt to reproduce as dominant individuals ( $z = 0$ ).

In general, delaying dispersal (resp. dispersing immediately) during the first year is favoured more readily when fecundity and/or survivorship is high (resp. low) (Figure 3a-c). This result is due to the assumption that individuals born on a territory have an advantage over non-native dispersers when competing for dominant vacancies, and it can be understood sufficiently well in terms of direct-fitness effects only (Figure 3d-f). As fecundity is increased, the competitive pressure from subordinates for dominant vacancies intensifies, and the chance of a disperser securing a dominant territory is reduced disproportionately. As dominant survivorship increases, the competitive pressure from subordinates also increases—not because

**Input:** Birth rate,  $b$   
Survival rate of dominant,  $s_\alpha$   
Survival rate of subordinate,  $s_\beta$   
Small step size,  $STEP > 0$   
Small tolerance,  $TOL > 0$

**Output:** Numerical prediction for stable tendency to delay dispersal

**Algorithm**

**Step 0:** Initialize two variables (i.e. replicates),  $Z1 = 0$  and  $Z2 = 1$

**Step 1:** Use  $z = Z1$ , and inputs  $b, s_\alpha, s_\beta$  to compute  $GRAD = \Delta w(z)$  using equation 1

**Step 2:** Update  $Z1$  by setting its new value equal to  $Z1 + GRAD * STEP$

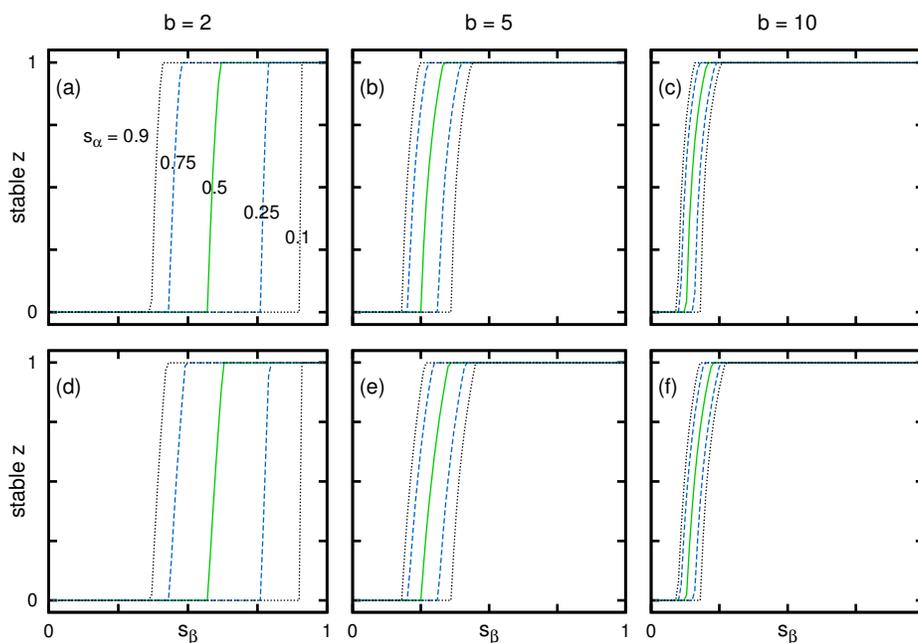
**Step 3:** If  $Z1 < 0$  set it equal to 0, if  $Z1 > 1$  set it equal to 1

**Step 4:** Repeat Steps 1 to 3 using  $Z2$  in place of  $Z1$

**Step 5:** If  $|Z1 - Z2| > TOL$  go back to Step 1, else return  $(Z1 + Z2)/2$

**Figure 2:** Description of the numerical simulation procedure used to determine the stable delay-dispersal phenotype,  $z$ , meaning that value of  $z$  favoured by selection in the long-run.

there are more natives, but because there are fewer vacancies over which individuals compete. Therefore, by delaying dispersal, an individual can profit from a valuable competitive advantage it would otherwise lose. Should that same individual not compete successfully, it may still survive to compete as a disperser in the next time period. In this way, an individual who delays dispersal is no worse off than one who did not (the absence of opportunity cost is common in models of cooperative breeding in stable habitats (McLeod and Wild, 2013)). In fact, delaying dispersal naturally affords individuals the opportunity to compete for vacancies twice: on its natal territory the first time, and elsewhere the second time. Increased survivorship as a subordinate, then, incentivizes delayed dispersal by improving one's chances of competing a second time.



**Figure 3:** Results for Model I: delayed dispersal. Long-term, stable tendency to delay dispersal (stable  $z$ ) as a function of subordinate survival rate ( $s_\beta$ ) for varying dominant survival rate  $s_\alpha$ , and varying birth rate  $b$ . Panels (a)-(c) show predictions from the complete inclusive-fitness model, while panels (d)-(f) uses direct-fitness effects only. Predictions do not depend on  $s_\omega$ .

The fact that direct-fitness considerations, alone, provide us with good understanding of the long-term evolution of  $z$  is also consistent with the quasi-threshold effect evident in Figure 3. When indirect-fitness considerations are small, compared to the direct-fitness ones,

there will be little to oppose the pursuit of selfish interests over the course of evolution. The answer to the question, “Is it advantageous to delay dispersal?”, then, will almost always be a clear “yes” or “no” because only one individual’s perspective need be adopted. This kind of clarity ultimately means selection will drive  $z$  to take a value of 0 (when the clear answer is “no”) or 1 (when the clear answer is “yes”) over time.

#### 4. Model II: Subordinate Helping

From the previous section, we predict that delayed dispersal evolves readily when fecundity and survivorship are sufficiently high. Once delayed dispersal has emerged, non-breeding relatives from different generations have the opportunity to interact, and so the subordinate helping rate, denoted  $h$ , may be expected to evolve to non-zero levels. In this section we extend our first model to determine conditions under which subordinate helping,  $h > 0$ , is advantageous.

##### 4.1. Monomorphic Wild-type Population

We again consider the wild-type population described above. We assume that life-history parameters are such that the action of selection has resulted in all individuals delaying dispersal ( $z = 1$ ), and we assume that the wild-type subordinate does not help ( $h = 0$ ).

##### 4.2. The Inclusive-Fitness Effect of Subordinate Help

Consider a subordinate on a territory at the very beginning of a time period. Suppose that this mature subordinate individual carries a mutant allele that causes it to interact with one of the  $b$  immature individuals produced by the dominants on its territory just prior to its own dispersal (see Figure 1).

We assume that the interaction in question is costly for the subordinate who is about to disperse. Specifically, the interaction diminishes the physical condition of this subordinate so that the likelihood with which it survives as a disperser is changed by  $\Delta s_\omega < 0$ . We also assume that the interaction in question is beneficial for the immature individual. In this case, the interaction improves the condition of the immature individual so that its survival, following maturation, as a subordinate is changed by  $\Delta s_\beta > 0$  (recall that we have assumed  $z = 1$ ). The assumptions we make in this model reflect typical interactions among

cooperative breeders, with non-breeding individuals paying a cost to ultimately promote the production of non-descendant kin (Heinsohn and Legge, 1999).

In Appendix C, we show that the helpful act has changed the inclusive fitness of the subordinate actor by

$$\Delta w(h) = \underbrace{\Delta s_\beta(\hat{m}_\alpha v_\alpha + \hat{m}_\beta v_\beta - \hat{m}_\alpha \hat{c}(v_\alpha - v_\beta))}_{\text{indirect effect}} \hat{r} + \underbrace{\Delta s_\omega \hat{m}_\alpha v_\alpha}_{\text{direct effect}}. \quad (2)$$

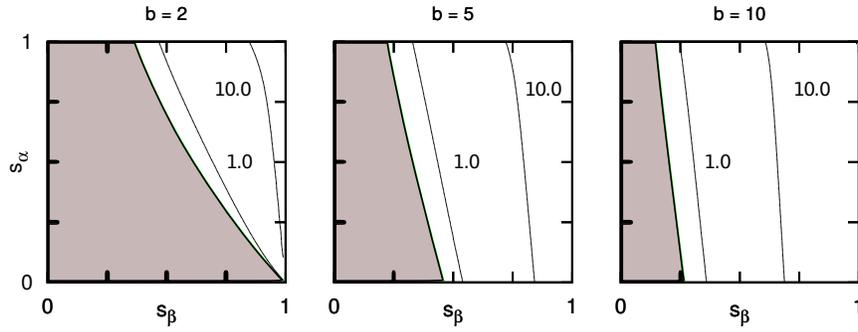
We also argue, in Appendix C, that  $\hat{r}$ , the relatedness between two different individuals born on the same territory during the same time period, is the relevant measure of relatedness for studying the advantage of costly subordinate help.

Like equation (1), equation (2) has a straightforward interpretation, but we will dispense with the details of the interpretation, here. It is important, though, to note that when  $\Delta w(h)$  is positive (resp. negative) a helpful act confers an advantage (resp. disadvantage) to a subordinate actor's lineage when the act, itself, is directed toward a recipient born on the same territory as the actor.

## 5. Method of Analysis and Results II: Subordinate Helping

By simple algebraic rearrangement of the condition  $\Delta w(h) > 0$ , we can arrive at a critical cost-benefit ratio,  $-(\Delta s_\omega/s_\omega)/(\Delta s_\beta/s_\beta)$ , above which subordinate helping is disadvantageous. By expressing the change in survival ( $\Delta s_\omega$  or  $\Delta s_\beta$ , respectively) relative to survival rates themselves ( $s_\omega$  and  $s_\beta$ , respectively), we obtain a critical ratio that does not depend on  $s_\omega$ .

Study of the critical cost-benefit ratio shows that increases in both survivorship and fecundity promotes the emergence of subordinate help. Evidence of this is presented in Figure 4. In that figure, we see larger values of  $s_\alpha$ ,  $s_\beta$ , and  $b$ , respectively, increase the range of conditions under which the advantage of help is maintained under large relative costs. As an example of what we mean, consider what happens as the fecundity parameter  $b$  is increased in Figure 4 (from left to right): the region of  $s_\alpha, s_\beta$ -space over which the advantage of help is maintained increases in area, when the cost-benefit ratio exceeds 10.0. Carrying on with this example, if we assume  $s_\omega = s_\beta$ , then our critical cost-benefit ratio takes a more usual form,  $-\Delta s_\omega/\Delta s_\beta$ . Certainly, the condition  $-\Delta s_\omega/\Delta s_\beta < 10.0$  is not overly restrictive and suggests the indirect benefits in equation (2) can translate to appreciable inclusive-fitness gains.



**Figure 4:** Results for Model II: subordinate helping. Contours for critical cost-benefit ratio,  $-(\Delta s_\omega/s_\omega)/(\Delta s_\beta/s_\beta)$ , when all individuals become subordinates ( $z = 1$ ). Darker black contours are shown in the space of dominant and subordinate survival rates ( $s_\alpha$  and  $s_\beta$ , respectively), and calculated for various fecundity parameters ( $b$ ). Contour labels 1.0, or 10.0 give the value of the cost-benefit ratio which, if not exceeded, ensures subordinate helping a selective advantage. Shaded region shows where complete delayed dispersal,  $z = 1$ , is not stable. To ensure  $z = 1$  is stable, life-history parameters must lie in the unshaded region of  $s_\beta, s_\alpha$ -space.

## 6. Discussion and Conclusions

### 6.1. Comparison with Previous Models

Models of the evolution of cooperative breeding often assume an individual's tendency to delay dispersal is linked to its tendency to help (Motro, 1993; Pen and Weissing, 2000; Leggett et al., 2012; McLeod and Wild, 2013; Wild and Koykka, 2014). Those that have sought to separate the link between staying and helping have either treated relatedness as a fixed parameter (Kokko and Johnstone, 1999; Kokko et al., 2001), have neglected the possibility that relatedness builds within social groups while placing severe limits on subordinate numbers (Koykka and Wild, 2015), or have focused on modelling delayed dispersal only (Kokko and Lundberg, 2001; Kokko and Ekman, 2002).

Here, we provide an explanation for the evolution of cooperative breeding that separates delaying dispersal from helping, while providing equal coverage to both behaviours. Our explanation also makes clear assumptions that show how relatedness within social groups, among other things, changes with changing life-history parameters, and ultimately how those changes impact the origin of cooperative breeding. This brings theory for cooperative breeding in line with that of many other social behaviours, including more generic forms of

altruism (Taylor, 1992a,b), natal dispersal (Taylor, 1989), reproductive effort (Pen, 2000), sex allocation (Wild and Taylor, 2004, 2005), and sexual conflict (Wild et al., 2011). We have also allowed variable numbers of subordinates to inhabit breeding territories (cf. Koykka and Wild, 2015). In all of these ways, then, our work represents a step forward in our understanding of cooperative breeding.

## 6.2. *Delayed Dispersal*

Our model for the evolution of delayed dispersal makes two main predictions. First, it predicts that indirect effects of delayed dispersal have minimal impact on the evolution of this trait. We found that predictions of a modified version of equation (1), one that ignored effects on relatives, made predictions that were very close to those made by the full inclusive-fitness model. Previous models for the evolution of delayed dispersal have purposefully neglected the indirect effects of delayed dispersal (Kokko and Lundberg, 2001; Kokko and Ekman, 2002). The importance of competition among siblings has been identified by theoretical models of dispersal in species without dominance hierarchies (Hamilton and May, 1977; Taylor, 1988), and is considered to be a prominent force shaping sociality in cooperative breeders (Gaston, 1978; Griffin and West, 2002). We see here, however, that ignoring those effects in cooperative breeders is unlikely to lead one too far astray provided the assumptions of the model are met. Chief among these assumptions is the degree of ‘habitat elasticity’ (Taylor, 1992b) experienced by subordinates. Subordinate individuals in our model may compete with relatives for breeding opportunities, but they do not compete with relatives in order to simply remain on their natal territory. Instead, a territory effectively expands in an ‘elastic’ manner to accommodate variable numbers of subordinates, thus reducing inclusive-fitness costs of delaying dispersal.

The kind of habitat elasticity experienced by subordinates in our model is not at all far-fetched. Though cooperative-breeding vertebrates often occupy saturated habitats, it is suitable *breeding habitat* that is considered to be saturated. Non-breeding individuals are often able to remain near their natal sites, “floating” in available marginal habitat and reducing competition with kin (Stacey and Ligon, 1991). This also applies to social insects, whose territories expand as their colonies grow. Expansion, here, is either due to (1) the fact that social insects are central site foragers, where increasing forager numbers increases the foraging area that can be explored and exploited, or (2) the fact that territories occupied

have vast food supplies that, in turn, allow for the coexistence of vast numbers of insects with minimal local competition (Korb and Heinze, 2008).

The second, perhaps more important, prediction made by our delayed-dispersal model concerns the effect of life-history parameters. We found that increased longevity and increased fecundity both promote delayed dispersal. In keeping with our previous comments, these effects are well understood in terms of direct consequences for personal fitness. These points are discussed further below, but it is worth stressing that our results confirm previous conjectures about life-history effects on delayed dispersal, particularly fecundity (Kokko and Ekman, 2002).

### *6.3. Subordinate Helping, Life-History Features, and Cooperative Breeding*

Our second model predicts that increased survivorship and fecundity, respectively, increase the maximum cost that can support an advantage of helping. This is similar to the predictions made by our model for delayed dispersal. Since cooperative breeding is thought to have arisen as a result of the emergence of delayed dispersal followed by helping (Brown, 1974; Emlen, 1982a), our models predict that life-histories that consist of low mortality and high fecundity promote the evolution of cooperative breeding, overall.

Life-history-based hypotheses for the advantage of cooperative breeding have focused on the role played by mortality, especially the mortality of dominant breeders. It is often argued that low mortality results in overcrowded populations with limited availability of breeding habitat due to low rates of turnover (Arnold and Owens, 1998; Beauchamp, 2014). Our results agree with this predicted association between the occurrence of cooperative breeding and mortality/survival.

Life-history hypotheses have also tended to downplay the effects fecundity might exert in the emergence of cooperative breeding systems. Authors predict that cooperative breeding will be associated with smaller clutch sizes, owing to a trade-off between survival and fecundity (Arnold and Owens, 1998; Beauchamp, 2014). Essentially this perspective characterizes fecundity as a life-history response to cooperative breeding, rather than a life-history feature that might directly promote help.

In contrast to the typical life-history-based views, our results show that, all else being equal, increases in fecundity can promote cooperative breeding via increasing incentives to both delay dispersal and to help. Incentives, here, stem from the fact that increased fecundity

can also result in low territory turnover, and a build-up of within-territory relatedness—or, more simply, high rates of territory retention by a given genetic lineage.

Certainly, our modelling shows that fecundity and survival both have a primary role to play in promoting and shaping cooperative breeding, but what practical implications might this have? One answer has to do with the mixed support life-history hypotheses have received (Arnold and Owens, 1998; Beauchamp, 2014). In the presence of a trade-off between survivorship and fecundity, one might not expect any clear life-history pattern to be associated with cooperative breeding. Macroscopic features of the population, like low rates of territory turn-over, could be associated equally well with high fecundity and low mortality, or low fecundity and high mortality, or even intermediate levels of both. Granted, comparative analyses have found little differences in birth rates of cooperatively breeding and non-cooperatively breeding species (Arnold and Owens, 1998; Beauchamp, 2014). Variation in birth rates, however, was still reported by these analyses and could definitely confound efforts to extract a clear “life-history signal” from the data collected. In addition, comparative analyses cannot explain why longevity alone does not lead to cooperative breeding more frequently in certain taxa (Arnold and Owens, 1998). Perhaps such puzzles presented by the absence of cooperative breeding could be solved by giving greater consideration to the role played by fecundity in directly promoting cooperative breeding, and may require deeper investigation into the details of relevant survival-fecundity trade-offs. Alternatively, it may be necessary to develop measures of territory turnover rates, rather than indices of habitat saturation, that combine fecundity and survival field data in meaningful ways to improve comparative analyses. To this end, terms like  $\hat{m}_\alpha$  or  $\hat{c}$ , that appear in equations (1) and (2) might serve as useful guides for field biologists.

We may extend the above considerations to social insects. Social hymenoptera (ants and some bees and wasps) as well as termites, which evolved eusociality independently, are characterized by an apparent reversal (or absence) of the fecundity/longevity trade-off. The most fecund individuals (queens, and in termites also kings) outlive the non-reproducing workers and soldiers by one to two orders of magnitude (Keller and Genoud, 1997; Keller, 1998; Monroy Kuhn and Korb, 2016). An extreme example comes from fungus-growing *Macrotermes* termites in which the lifespan of queens is more than 15-20 years while the workers only live 2-3 months (Traniello and Leuthold, 2000). At the same time, these queens are the most fertile individuals amongst all animals as they lay up to 20 000 eggs

per day. The evolution of this reshaping of the fecundity/longevity trade-off is still poorly understood but one may speculate that, after the evolution of helping, feedback between sociality, increased survival and fecundity may explain it.

Insects have higher maximum fecundities than vertebrates and it has been speculated that this may explain why eusociality rarely evolved in the latter (Alexander et al., 1991; Korb, 2009). Strikingly, the single eusocial vertebrates are rodents, those mammals with the highest fecundity. Hence, high fecundity, which allows for large families with many offspring to care for, may facilitate the evolution of altruism and set the stage for the co-evolution of sociality, and increased survival and fecundity.

#### *6.4. Limitations and Future Work*

Models are always limited by their assumptions, and several key assumptions limit those presented here. One of the most important assumptions we made, was the assumption of weak selection: an individual's tendency to delay dispersal was slight, as were the effects of helping on the survival of actors and recipients alike. In mathematical terms, weak selection means that we cannot use our model to definitively assess the evolutionary stability of strategies, as these have been traditionally defined (Rousset, 2004). In biological terms, this means our model cannot be used to study non-additive effects, like any possible synergy that could occur among helpers as a helpful gene invades (Grafen, 1985). Extensions of this work that consider higher-order effects of behaviours (e.g. using the analytic tools of adaptive dynamics (Dercole and Rinaldi, 2008)) could, therefore, be fruitful.

A second key feature of our work was the assumption that delayed dispersal had become established prior to the emergence of helping. Though this idea has some long-standing support in the biological literature (Brown, 1974; Emlen, 1982a; Koenig et al., 1992; Hatchwell and Komdeur, 2000; Kokko and Lundberg, 2001; Kokko and Ekman, 2002), an alternative scenario – where helping arises first, then delay dispersal emerges to enhance helping – has not been explored theoretically. That said, there is little, if any, evidence that such an alternative scenario has actually occurred, as it would likely be recognized as within-generation altruism rather than cooperative breeding.

Our focus on the origins of helping is yet another limitation of our work. Indeed, few studies have tracked the coevolution of delayed dispersal and helping beyond their origins as independent traits, and those that have tracked said coevolution have presented relatively

complicated results (Koykka and Wild, 2015). Clearly, origins of helping along with the subsequent maintenance of cooperative breeding via coevolution deserve further attention.

Finally, we chose to model the tendency to delay dispersal as a condition-independent strategy. In natural populations, though, delays in dispersal can be informed by an individual’s physiological condition, as well as local population densities, and the availability of suitable alternative habitats (Pruett-Jones and Lewis, 1990). Helping may also be expressed in a similarly facultative manner (Holman, 2014). While our decision to ignore condition dependence likely did not influence our basic conclusions (e.g., population elasticity would have reduced kin competition even with condition-dependence), accounting for condition-dependent strategies might be important when investigating the levels to which helping and dispersal evolve once helping becomes established. Thus, ignoring condition dependence as we have done could limit our ability to extrapolate our conclusions beyond simply the emergence of cooperative breeding to include maintenance and longer-term evolution of these systems. The role of condition dependence in the origins of cooperative breeding is one of several interesting lines of further research.

## **Appendix A. Monomorphic Wild-type Population Without Subordinate Help**

### *Appendix A.1. Equilibrium State*

In this section, we develop the mathematical expressions that characterize the equilibrium state of the wild-type population, in the absence of subordinate help, as a distribution of the state of the territories in the population. We define the “state” of a territory as the number of subordinates found there at the beginning of a given time period. Since these probabilities are identical for each territory, the population converges to its equilibrium state after only a single time period.

Let  $\hat{p}_j$  denote the frequency of territories supporting  $j = 0, 1, \dots, b$  subordinates at equilibrium. We express equilibrium frequencies in terms of the number of dominant vacancies on a given territory,  $k = 0, 1, 2$ , and the number of surviving subordinates who have also delayed dispersal,  $\ell = 0, 1, \dots, b$ . For example, to determine an expression for  $\hat{p}_0$  we note that there are exactly  $k$  vacancies on a given territory in any given time step with probability

$$\binom{2}{k} s_\alpha^{2-k} (1 - s_\alpha)^k. \tag{A.1}$$

Conditioned on the event that  $k$  dominant vacancies arise, the territory in question ends up with  $j = 0$  subordinates when no more than  $k$  of the  $b$  offspring produced have survived and delayed dispersal (if fewer than  $k$  survive and delay dispersal, then vacancies are filled by dispersers, as described in the main text). The probability with which this conditional event occurs is given by

$$\sum_{\ell=0}^k \binom{b}{\ell} (zs_{\beta})^{\ell} (1 - zs_{\beta})^{k-\ell}. \quad (\text{A.2})$$

Combining the previous observations using the Law of Total Probability, we get

$$\hat{p}_0 = \sum_{k=0}^2 \binom{2}{k} s_{\alpha}^{2-k} (1 - s_{\alpha})^k \sum_{\ell=0}^k \binom{b}{\ell} (zs_{\beta})^{\ell} (1 - zs_{\beta})^{k-\ell}. \quad (\text{A.3})$$

Similar conditioning arguments results in the following overall description for  $\hat{p}_j$ :

$$\hat{p}_j = \begin{cases} \sum_{k=0}^2 \binom{2}{k} s_{\alpha}^{2-k} (1 - s_{\alpha})^k \sum_{\ell=0}^k \binom{b}{\ell} (zs_{\beta})^{\ell} (1 - zs_{\beta})^{b-\ell} & j = 0, \\ \sum_{k=0}^2 \binom{2}{k} s_{\alpha}^{2-k} (1 - s_{\alpha})^k \binom{b}{k+j} (zs_{\beta})^{k+j} (1 - zs_{\beta})^{b-(k+j)} & 1 \leq j \leq b-2, \\ \sum_{k=0}^1 \binom{2}{k} s_{\alpha}^{2-k} (1 - s_{\alpha})^k \binom{b}{k+b-1} (zs_{\beta})^{k+(b-1)} (1 - zs_{\beta})^{b-(k+b-1)} & j = b-1 \\ s_{\alpha}^2 (zs_{\beta})^b & j = b. \end{cases} \quad (\text{A.4})$$

The equilibrium state of the population, then, is given by  $\hat{p}_0, \dots, \hat{p}_b$ . Note that  $\hat{p}_0$  is the probability of the event that all surviving subordinates fill local vacancies left by dominants, and  $\hat{p}_{j \geq 1}$  is the probability of the event that all but  $j$  surviving subordinates fill local vacancies left by dominants. As a check, note that  $\sum_j \hat{p}_j = 1$ , as expected.

### Appendix A.2. Success of Dispersers

We continue to consider a monomorphic wild-type population without subordinate help, but now we determine the number of dominant positions each surviving disperser expects to secure at equilibrium. This number, denoted  $\hat{n}_{\alpha}$ , is simply the expected number of vacancies available to dispersers,

$$1 \times \underbrace{2s_{\alpha}(1 - s_{\alpha})(1 - zs_{\beta})^b + (1 - s_{\alpha})^2 b z s_{\beta} (1 - zs_{\beta})^{b-1}}_{\text{prob one vacancy for dispersers}} + 2 \times \underbrace{(1 - s_{\alpha})^2 (1 - zs_{\beta})^b}_{\text{prob two vacancies for dispersers}} \quad (\text{A.5})$$

divided by the number of surviving dispersers,  $s_\omega[(1-z)b + \sum_j j\hat{p}_j]$  (dispersing individuals born in the current time period plus former subordinates). Using the expressions in (A.4) it is possible to show that  $\sum_j j\hat{p}_j$ , is equal to the number of surviving subordinates from the previous time period less the number of vacancies filled by this group. Mathematically, the number of surviving dispersers,  $s_\omega[(1-z)b + \sum_j j\hat{p}_j]$ , is given by

$$s_\omega \left[ \begin{array}{c} (1-z)b + \underbrace{bz s_\beta}_{\# \text{ surviving subordinates}} \\ \underbrace{-(2(1-s_\alpha) - 2s_\alpha(1-s_\alpha)(1-zs_\beta)^b - (1-s_\alpha)^2(bzs_\beta(1-zs_\beta)^{b-1} + 2(1-zs_\beta)^b))}_{\# \text{ vacancies filled by non-dispersers}} \end{array} \right] \quad (\text{A.6})$$

and so the expected number of territories won by a surviving disperser is,

$$\hat{n}_\alpha = \frac{2s_\alpha(1-s_\alpha)(1-zs_\beta)^b + (1-s_\alpha)^2bz s_\beta(1-zs_\beta)^{b-1} + 2(1-s_\alpha)^2(1-zs_\beta)^b}{s_\omega[(1-z)b + \sum_j j\hat{p}_j]}. \quad (\text{A.7})$$

This expression for  $\hat{n}_\alpha$  will be used in the development of fitness expressions in a subsequent section.

Careful consideration of equation (A.7) reveals that one disperser can win multiple dominant positions; that is,  $\hat{n}_\alpha$  might be greater than one. This situation is typical of social evolutionary models of behaviour—for example in models of sex-ratio evolution where males secure multiple female mates (Wild and Taylor, 2004). This model feature arguably implies that the model itself over-estimates the success of dispersers. However, the model feature could be understood as a single individual being part of several different mated pairs, as in any one of a number of cases of extra-pair parentage found in the literature (Petrie and Kempnaers, 1998). Regardless of interpretation, the model feature is an important technical part of our model, because it ensures that vacancies which cannot be filled locally never go unfilled, and so the population does not tend to extinction over time.

### *Appendix A.3. Success of Subordinates*

Now we turn our attention to the expected number of dominant and subordinate positions won by a surviving non-disperser in a wild-type population at equilibrium, denoted  $\hat{m}_\alpha$  and  $\hat{m}_\beta$ , respectively.

We calculate  $\hat{m}_\alpha$  by (a) fixing attention on an immature individual that stays on its natal territory and survives to become a subordinate (possibly only temporarily), then (b) conditioning on the dispersal and survival of various territory inhabitants to get

$$\begin{aligned}
\hat{m}_\alpha &= (1 - s_\alpha)^2 \times \\
&\left[ \sum_{\ell=0}^1 \binom{b-1}{\ell} (zs_\beta)^\ell (1 - zs_\beta)^{b-1-\ell} + \sum_{\ell=2}^{b-1} \binom{b-1}{\ell} (zs_\beta)^\ell (1 - zs_\beta)^{b-1-\ell} \frac{2}{\ell+1} \right] \\
&\quad + 2s_\alpha(1 - s_\alpha) \sum_{\ell=0}^{b-1} \binom{b-1}{\ell} (zs_\beta)^\ell (1 - zs_\beta)^{b-1-\ell} \frac{1}{\ell+1} \\
&= \frac{2(1 - s_\alpha)(1 - (1 - zs_\beta)^b) - (1 - s_\alpha)^2 b z s_\beta (1 - zs_\beta)^{b-1}}{b z s_\beta}. \quad (\text{A.8})
\end{aligned}$$

Equation (A.8) can be understood as follows. In the case that two local vacancies arise (probability  $(1 - s_\alpha)^2$ ), the focal individual secures a local dominant position with probability 1 provided no more than one other subordinate survived (summarized by the first term in square brackets in equation A.8). If two or more additional subordinates are able to compete for the vacancies, then the focal individual secures a local dominant position with probability less than 1 (summarized by the second term in square brackets in equation A.8). If only one vacancy arises (probability  $2s_\alpha(1 - s_\alpha)$ ), the focal individual secures the dominant position with probability 1 only when no other subordinate is present to contest the vacancy (summarized by the final summation in equation A.8). As a check, notice that when we add the numerator of  $\hat{m}_\alpha$  in (A.8) to the numerator of  $\hat{n}_\alpha$  in equation (A.7), the sum is  $2(1 - s_\alpha)$  which is the total dominant vacancy rate.

By conditioning on the dispersal and survival of all but a focal subordinate, we find that the probability with which that subordinate does not have to compete for a local dominant position against a relative (i.e. the probability that said position is *uncontested* by a relative of the subordinate) is

$$\hat{u} = \frac{2s_\alpha(1 - s_\alpha)(1 - zs_\beta)^{b-1} + (1 - s_\alpha)^2 \left( (1 - zs_\beta)^{b-1} + (b-1)(1 - zs_\beta)^{b-2} \right)}{\hat{m}_\alpha}. \quad (\text{A.9})$$

It follows that the fraction of sites that are contested by relatives at equilibrium is,  $\hat{c} = 1 - \hat{u}$ , or

$$\hat{c} = 1 - \frac{2s_\alpha(1 - s_\alpha)(1 - zs_\beta)^{b-1} + (1 - s_\alpha)^2 \left( (1 - zs_\beta)^{b-1} + (b-1)(1 - zs_\beta)^{b-2} \right)}{\hat{m}_\alpha}. \quad (\text{A.10})$$

A similiar line of reasoning gives us

$$\begin{aligned} \hat{m}_\beta &= (1 - s_\alpha)^2 \sum_{\ell=2}^{b-1} \binom{b-1}{\ell} (zs_\beta)^\ell (1 - zs_\beta)^{b-1-\ell} \left(1 - \frac{2}{\ell+1}\right) \\ &\quad + 2s_\alpha(1 - s_\alpha) \sum_{\ell=1}^{b-1} \binom{b-1}{\ell} (zs_\beta)^\ell (1 - zs_\beta)^{b-1-\ell} \left(1 - \frac{1}{\ell+1}\right) + s_\alpha^2 = 1 - \hat{m}_\alpha, \end{aligned} \quad (\text{A.11})$$

which makes intuitive sense, because a subordinate who survives either becomes a dominant or remains a subordinate. In fact, this calculation serves as yet another check on the correctness of (A.8).

#### *Appendix A.4. Calculation of Relatedness*

Here, we provide an explicit expression for  $\hat{r}$ , the coefficient of relatedness between two individuals born on the same territory during the same time period. Again, to make our calculations we assume that the population is at equilibrium, genetically monomorphic, and that subordinates do not help.

We begin by considering the relatedness between the two dominant breeders that occupy the same territory in a given time step, denoted  $f$ . Technically,  $f$  gives the probability that one breeder carries an allele at some focal locus that is identical by descent (IBD) to the allele carried by the other breeder at the same locus (Michod and Hamilton, 1980). We make the usual assumption that two individuals born on different territories have a coefficient of relatedness equal to zero (e.g. Taylor, 1992a). After any given time period, then, the relatedness between breeders can either increase to one, decrease to zero, or remain unchanged. It follows that the relatedness between dominant breeders in the next time step, denoted  $f'$ , can be expressed recursively as,

$$f' = f \times (1 - P_1 - P_0) + 1 \times P_1 + 0 \times P_0 \quad (\text{A.12})$$

where  $P_1$  is the probability with which the relatedness increases to 1, and  $P_0$  is the probability with which it decreases to zero. We find

$$P_1 = \frac{1}{2} 2s_\alpha(1 - s_\alpha)[1 - (1 - zs_\beta)^b] + \frac{1}{2}(1 - s_\alpha)^2[1 - (1 - zs_\beta)^b - bzs_\beta(1 - zs_\beta)^{b-1}]. \quad (\text{A.13})$$

The first term in equation (A.13) gives the probability that exactly one dominant dies, at least one subordinate is available to succeed the deceased dominant, and the surviving dominant made the genetic contribution to the successor. The second term in equation (A.13) gives

the probability that both dominants die, at least two subordinates are available to succeed the deceased dominants, and the two successors both received genetic contributions from the same breeder. A similar argument allows one to arrive at the remaining terms in (A.13), and also leads one to conclude that

$$P_0 = 2s_\alpha(1 - s_\alpha)(1 - zs_\beta)^b + (1 - s_\alpha)^2[(1 - zs_\beta)^b + bzs_\beta(1 - zs_\beta)^{b-1}]. \quad (\text{A.14})$$

At equilibrium,  $f = f' = \hat{f}$  where

$$\hat{f} = \frac{P_1}{P_1 + P_0}. \quad (\text{A.15})$$

We now use  $\hat{f}$  to express the relatedness between two individuals born on the same territory during the same time period, denoted  $\hat{r}$ . Conditioning on the identity of the dominant breeder(s) who made the genetic contribution to the individuals being compared, we find

$$\hat{r} = \frac{1 + \hat{f}}{2}. \quad (\text{A.16})$$

This is the measure of relatedness featured in the main text.

## Appendix B. The Selective Advantage of Delayed Dispersal

### *Appendix B.1. Key Assumptions*

Suppose that the tendency to delay dispersal is controlled at a single locus, and suppose further that there are two alleles at this locus, one wild-type allele and one mutant allele (no longer is the model population genetically monomorphic). A mutant remains on its natal site at a rate that differs only slightly from the wild-type rate,  $z$ ; in other words, selection is assumed to be weak.

### *Appendix B.2. Components of Fitness*

In order to determine whether the mutant is at an advantage/disadvantage, and ultimately the direction in which selection moves  $z$ , we use the “direct-fitness approach” (Taylor and Frank, 1996).

The direct-fitness approach requires expressions for fitness. For this model, we develop fitness expressions for an immature individual at the very beginning of a given time period (the “focal individual”). Each focal individual has two distinct components of fitness. The first component represents the likelihood with which the focal individual belongs to the

dominant class in the next time period, and is denoted  $w_\alpha$ . The second component represents the likelihood with which it belongs to the subordinate class in the next time period, and is denoted  $w_\beta$ . We deal with each component in turn, below.

Consider  $w_\alpha$ , the likelihood that a focal individual belongs to the dominant class in the next time period, or more accurately, the expected number of dominant positions won by the focal individual in the next time period. We will break the calculation of  $w_\alpha$  into two parts: the number of dominant positions won on its natal territory in the next time period (i.e. locally), and the number of dominant positions won on territories other than its natal territory in the next time period (i.e. non-locally).

Let  $z_0$  denote the focal individual's tendency to delay dispersal, and let  $z_1$  denote the tendency of another individual, born on the same territory and at the same time as the focal individual, to delay its own dispersal. Conditioning on the event that the focal individual stays on its natal patch and survives (probability  $z_0 s_\beta$ ), then applying the same argument that led to equation (A.8), we find that the number of local dominant positions next time is

$$z_0 s_\beta \frac{2(1 - s_\alpha)(1 - (1 - z_1 s_\beta)^b) - (1 - s_\alpha)^2 b z_1 s_\beta (1 - z_1 s_\beta)^{b-1}}{b z_1 s_\beta}. \quad (\text{B.1})$$

In order to determine the number of non-local successes, we condition on the event that the focal individual disperses and survives to compete for a vacancy (probability  $(1 - z_0) s_\omega$ ). Because selection is weak, we can approximate the number of vacancies available using the expression for  $\hat{n}_\alpha$  already derived for a wild-type population at equilibrium (see Appendix A). Consequently, the number of non-local dominant position next time is

$$(1 - z_0) s_\omega \hat{n}_\alpha. \quad (\text{B.2})$$

We arrive at an expression for  $w_\alpha$  by simply summing equations (B.1) and (B.2):

$$w_\alpha = z_0 s_\beta \frac{2(1 - s_\alpha)(1 - (1 - z_1 s_\beta)^b) - (1 - s_\alpha)^2 b z_1 s_\beta (1 - z_1 s_\beta)^{b-1}}{b z_1 s_\beta} + (1 - z_0) s_\omega \hat{n}_\alpha. \quad (\text{B.3})$$

Now, consider the likelihood with which the focal individual belongs to the subordinate class at the beginning of the next time period,  $w_\beta$ . Given the assumptions of the model, an individual can only become a subordinate if it remains on its natal territory, and so we condition on the event that the focal individual has delayed dispersal and has survived (probability  $z_0 s_\beta$ ). Using the approach that led to equation (A.11), we get

$$w_\beta = z_0 s_\beta \left( 1 - \frac{2(1 - s_\alpha)(1 - (1 - z_1 s_\beta)^b) - (1 - s_\alpha)^2 b z_1 s_\beta (1 - z_1 s_\beta)^{b-1}}{b z_1 s_\beta} \right). \quad (\text{B.4})$$

### *Appendix B.3. Combining Components of Fitness*

Expressions (B.3) and (B.4) can be combined into a single measure of fitness using coefficients of reproductive value. “Reproductive value” is defined as the genetic contribution made by an individual to the population in the very distant future (Taylor, 1996). If we use  $v_\alpha$  to denote the reproductive value of a dominant individual, then reproductive value of a subordinate becomes  $v_\beta = s_\omega \hat{n}_\alpha v_\alpha$ . The expression for  $v_\beta$  can be understood as the product of three terms: the probability with which a subordinate later survives as a disperser to compete on a new territory, the expected number of dominant positions it wins, and its value as a dominant. This measure of reproductive value is approximated to zeroth order in the strength of selection.

With reproductive value defined, we get

$$w = v_\alpha w_\alpha + v_\beta w_\beta \tag{B.5}$$

as our measure of the overall fitness of a focal individual.

### *Appendix B.4. The Inclusive-Fitness Effect of Philopatry*

Using Taylor and Frank (1996) as a guide, we calculate  $\Delta w(z)$  by (a) treating  $z_1$  as a function of  $z_0$ , (b) differentiating  $w$  with respect to  $z_0$ , paying careful attention to the Chain Rule from freshman calculus, (c) substituting the relatedness between the focal individual and its newborn neighbour, denoted  $\hat{r}$  (see Appendix A), in place of the ordinary derivative  $dz_1/dz_0$ , and (d) evaluating the entire expression at  $z_0 = z_1 = z$  and. These steps lead to equation (1) in the main text.

## **Appendix C. Model Extension: the Advantage of Subordinate Help**

### *Appendix C.1. Preliminary Comments*

In this section, we consider a scenario in which selection has led all individuals to delay dispersal for one time period (i.e.,  $z = 1$ ). We develop an expression for the inclusive-fitness advantage/disadvantage created by a helpful subordinate individual on a territory at the beginning of a time step, in this case. We dispense with the direct-fitness methodology used in a previous section, and instead use a more traditional inclusive-fitness approach (Taylor, 1990). However, we do assume the reader is familiar with the contents of previous appendices.

*Appendix C.2. Marginal Changes in Fitness*

Suppose the helpful individual in question reduces its own condition to improve the condition of an immature individual on its territory. Let  $s'_\omega$  denote the eventual survival rate of the helper that results from its reduced condition, and let  $\Delta s_\omega = s'_\omega - s_\omega < 0$  denote the marginal effect helping has on helper's eventual survival as a disperser. Similarly, let  $s'_\beta$  denote the eventual survival rate of the immature individual whose condition is improved by the helper, and let  $\Delta s_\beta = s'_\beta - s_\beta > 0$  denote the marginal improvement in survival rate that occurs. In keeping with our assumption of weak selection, both  $\Delta s_\omega$  and  $\Delta s_\beta$  are assumed to be small.

Following equation (A.8), we see that the likelihood with which the recipient of the help becomes a dominant individual in the next time period is  $s'_\beta \hat{m}_\alpha$ , whereas without help the likelihood would have been  $s_\beta \hat{m}_\alpha$ . It follows that

$$\Delta s_\beta \hat{m}_\alpha \tag{C.1}$$

is one of the immediate benefits of help. It should be understood that  $\hat{m}_\alpha$  in (C.1) is calculated with  $z = 1$ .

A second immediate benefit of helping is realized in the case where the recipient belongs to the subordinate class in the next time period. The second benefit is calculated in a manner analogous to the one in (C.1), and is expressed as

$$\Delta s_\beta \hat{m}_\beta, \tag{C.2}$$

where it is understood that  $\hat{m}_\beta$  is calculated with  $z = 1$ .

Because it improves survival, helping increases the extent of competition among kin for dominant positions. This is a deferred cost of helping, and it is incurred when the recipient of the help displaces a relative. We express the cost as

$$\Delta s_\beta \hat{m}_\alpha \hat{c} \tag{C.3}$$

where it is, once again, understood that  $z = 1$ .

Finally, the helper also changes its own fitness. Specifically, it reduces the likelihood with which it fills a dominant vacancy, following its own inevitable dispersal, by an amount equal to

$$s'_\omega \hat{n}_\alpha - s_\omega \hat{n}_\alpha = \Delta s_\omega \hat{n}_\alpha. \tag{C.4}$$

The fitness changes described above must be weighted by measures of reproductive value so that they be combined into a single, overall expression for inclusive-fitness change. The appropriate measures are again  $v_\alpha$  and  $v_\beta$  presented in Appendix B. To be clear, changes in lines (C.1) and (C.4) must be weighted by  $v_\alpha$ , the change in line (C.2) must be weighted by  $v_\beta$ , and the change in line (C.3) must be weighted by  $v_\alpha - v_\beta$ . Lines (C.1) to (C.3) must also be weighted by relatedness coefficients. These coefficients are calculated below.

### *Appendix C.3. Relatedness*

We will use the calculations presented in Appendix A in order to determine the correct coefficient of relatedness to use in our second model. It is important to remember that the calculations carried out below assume that  $z = 1$  and  $h = 0$ .

Fix attention on a subordinate at the beginning of time period  $t$ . During time period  $t - 1$  we know that, on the territory with which the subordinate is associated, there were more than enough other subordinates to fill any dominant vacancies that might have arisen (otherwise, we would not have found the focal subordinate there, by assumption). Let  $q_i$  denote the probability that  $i = 0, 1, 2$  dominant positions were filled at the end of time  $t - 1$  on the focal subordinate's natal territory. We will not concern ourselves with developing explicit expressions for  $q_i$ s since these ultimately drop out of our calculation. We will, however, use  $q_i$ s to guide our calculation of relatedness between the focal subordinate and an immature individual on the same territory.

Consider a uniform-random immature individual on the same territory as the focal subordinate, at the beginning of time period  $t$ . With probability  $q_0$  the dominant individuals that occupy the focal territory during time period  $t$  also occupied the territory during time period  $t - 1$ . In this case, the focal subordinate and the immature may have the same parent (probability  $1/2$ ) so that they are related by a factor of 1. Alternatively, the focal subordinate and the immature individual may have different parents (probability  $1/2$ ) so that they are related by a factor of  $\hat{f}$ . With probability  $q_0$ , then, the relatedness between the focal subordinate and a given immature individual on the same territory is  $(1 + \hat{f})/2 = \hat{r}$ .

With probability  $q_1$  one of the dominant individuals at the beginning of time period  $t$  also occupied the territory during time  $t - 1$ , while the other dominant individual is breeding for the first time during time period  $t$ . In this case, the parent of the immature individual under consideration may be the first-time dominant (probability  $1/2$ ) so that the subordinate and

the immature individual are related by a factor  $\hat{r}$ . Alternatively, the parent of the immature may be the more experienced dominant (probability  $1/2$ ). The more experienced dominant may also be the parent of the focal subordinate (probability  $1/2$ ), so that the relatedness between subordinate and the immature is 1. If the more experienced dominant is not the parent of the focal subordinate (probability  $1/2$ ), then the relevant relatedness is  $\hat{f}$  since the subordinate's parent would have been found alongside the more experienced dominant during time  $t - 1$ . With probability  $q_1$ , then, the relatedness between the focal subordinate and a given immature individual is  $\hat{r}/2 + (1/2)(1 + \hat{f})/2 = \hat{r}$ .

With probability  $q_2$ , both dominant individuals are first-time dominants during period  $t$ . By assumption, both would have been born during the same time period as the focal subordinate (time period  $t - 1$ ). It follows that the relatedness between the focal subordinate and any immature produced during time period  $t$  is  $\hat{r}$ .

Putting the pieces of our calculations together, we find that the relatedness between a focal subordinate and the immature individual that receives its help is  $q_0\hat{r} + q_1\hat{r} + q_2\hat{r} = \hat{r}$ .

#### *Appendix C.4. The Inclusive-Fitness Effect of Helping*

Weighting marginal fitness changes by the appropriate measures of reproductive value, and the appropriate relatedness coefficients we find that the inclusive-fitness effect of subordinate help is given by equation (2) in the main text.

### **Acknowledgements**

Comments made by J. Kasperski, and an anonymous reviewer improved various drafts of this work. GW is supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. 549848). JK is supported by the German Science Foundation (KO1895/19-1).

### **References**

- Alexander, R.D., Noonan, K.M., Crespi, B.J., 1991. The evolution of eusociality, in: Sherman, P.W., Jarvis, J.U., Alexander, R.D. (Eds.), *The Biology of the Naked Mole-rat*. Princeton University Press, Princeton, NJ, pp. 3–44.
- Arnold, K.E., Owens, I.P.F., 1998. Cooperative breeding in birds: a comparative test of the life history hypothesis. *Proc. R. Soc. B* 265, 739–745.

- Beauchamp, G., 2014. Do avian cooperative breeders live longer? *Proc. R. Soc. B* 281, 20140844.
- Brown, J.L., 1974. Alternative routes to sociality in jays—with a theory for the evolution of altruism and communal breeding. *Am. Zool.* 14, 63–80.
- Brown, J.L., 1987. *Helping and Communal Breeding in Birds*. Monographs in Behavior and Ecology, Princeton University Press, Princeton, NJ.
- Clutton-Brock, T., 2002. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* 296, 69–72.
- Cockburn, A., 1998. Evolution of helping behavior in cooperatively breeding birds. *Annu. Rev. Ecol. Syst.* 29, 141–177.
- Dercole, F., Rinaldi, S., 2008. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press, Princeton, NJ.
- Emlen, S.T., 1982a. The evolution of helping. I. an ecological constraints model. *Am. Nat.* 119, 29–39.
- Emlen, S.T., 1982b. The evolution of helping. II. the role of behavioral conflict. *Am. Nat.* 119, 40–53.
- Emlen, S.T., 1994. Benefits, constraints and the evolution of the family. *Trends Ecol. Evol.* 9, 282–285.
- Gaston, A.J., 1978. The evolution of group territorial behavior and cooperative breeding. *Am. Nat.* 112, 1091–1110.
- Geritz, S.A.H., 2005. Resident-invader dynamics and the coexistence of similar strategies. *J. Math. Biol.* 50, 67–82.
- Grafen, A., 1985. Hamilton's rule OK. *Nature* 318, 310–311.
- Griffin, A.S., West, S.A., 2002. Kin selection: fact and fiction. *Trends Ecol. Evol.* 17, 15–21.
- Hamilton, W.D., 1964. The genetical evolution of social behavior. I and II. *J. Theor. Biol.* 7, 1–52.

- Hamilton, W.D., May, R.M., 1977. Dispersal in stable habitats. *Nature* 269, 578–581.
- Hatchwell, B.J., Komdeur, J., 2000. Ecological constraints, life-history traits and the evolution of cooperative breeding. *Anim. Behav.* 59, 1079–1086.
- Heinsohn, R., Legge, S., 1999. The cost of helping. *Trends Ecol. Evol.* 14, 53–57.
- Holman, L., 2014. Conditional helping and evolutionary transitions to eusociality and cooperative breeding. *Behav. Ecol.* 25, 1173–1182.
- Johnstone, R.A., Cant, M.A., 2008. Sex differences in dispersal and the evolution of helping and harming. *Am. Nat.* 172, 318–330.
- Keller, L., 1998. Queen lifespan and colony characteristics in ants and termites. *Insectes Sociaux* 45, 235–246.
- Keller, L., Genoud, M., 1997. Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature* 389, 958–960.
- Koenig, W.D., Pitelka, F.A., Carmen, W.J., Mumme, R.L., 1992. Delayed dispersal in cooperative breeders. *Q. Rev. Biol.* 67, 111–150.
- Kokko, H., Ekman, J., 2002. Delayed dispersal as a route to breeding: territorial inheritance, safe havens, and ecological constraints. *Am. Nat.* 160, 468–484.
- Kokko, H., Johnstone, R.A., 1999. Social queuing in animal societies: a dynamic model of reproductive skew. *Proc. R. Soc. Lond. B* 266, 571–578.
- Kokko, H., Johnstone, R.A., Clutton-Brock, T.H., 2001. The evolution of cooperative breeding through group augmentation. *Proc. R. Soc. B* 268, 187–196.
- Kokko, H., Lundberg, P., 2001. Dispersal, migration, and offspring retention in saturated habitats. *Am. Nat.* 157, 188–202.
- Korb, J., 2009. Termites: An alternative road to eusociality and the importance of group benefits in social insects, in: Gadau, J., Fewell, J.H. (Eds.), *Organization of Insect Societies. From Genome to Sociocomplexity*. Harvard University Press, Cambridge, MA, pp. 128–147.

- Korb, J., Heinze, J., 2008. The ecology of social life: a synthesis, in: Korb, J., Heinze, J. (Eds.), *Ecology of Social Evolution*. Springer, Berlin/Heidelberg, pp. 245–260.
- Koykka, C., Wild, G., 2015. The association between the emergence of cooperative breeding and clutch size. *J. Evol. Biol.* in press.
- Leggett, H.C., El Mouden, C., Wild, G., West, S.A., 2012. Promiscuity and the evolution of cooperative breeding. *Proc. R. Soc. B* 279, 1405–1411.
- McLeod, D.V., Wild, G., 2013. Ecological constraints influence the emergence of cooperative breeding when population dynamics determine the fitness of helpers. *Evolution* 67, 3221–3232.
- Michod, R.E., Hamilton, W.D., 1980. Coefficients of relatedness in sociobiology. *Nature* 288, 694–697.
- Monroy Kuhn, J.M., Korb, J., 2016. Social insects: aging and the re-shaping of the fecundity/longevity trade-off. *Curr. Opin. Insect Sci.* 16, vii–x.
- Motro, U., 1993. Helpers at parents' nest: a game theoretic approach. *J. Theor. Biol.* 163, 127–134.
- Pen, I., 2000. Reproductive effort in a viscous population. *Evolution* 54, 293–297.
- Pen, I., Weissing, F.J., 2000. Towards a unified theory of cooperative breeding: the role of ecology and life history re-examined. *Proc. R. Soc. B* 267, 2411–2418.
- Petrie, M., Kempenaers, B., 1998. Extra-pair paternity in birds: explaining variation between species and populations. *Trends Ecol. Evol.* 13, 52–58.
- Pruett-Jones, S.G., Lewis, M.J., 1990. Sex ratio and habitat limitation promote delayed dispersal in superb fairy-wrens. *Nature* 348, 541–542.
- Rousset, F., 2004. *Genetic Structure and Selection in Subdivided Populations*. Monographs in Population Biology, Princeton University Press, Princeton, NJ.
- Skutch, A.F., 1961. Helpers among birds. *Condor* 63, 198–226.

- Stacey, P.B., Ligon, J.D., 1991. The benefits-of-philopatry hypothesis for the evolution of cooperative breeding: variation in territory quality and group size effects. *Am. Nat.* 1137, 831–846.
- Taylor, P.D., 1988. An inclusive fitness model for dispersal of offspring. *J. Theor. Biol.* 130, 363–378.
- Taylor, P.D., 1989. Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* 36, 125–143.
- Taylor, P.D., 1990. Allele-frequency change in a class-structured population. *Am. Nat.* 135, 95–106.
- Taylor, P.D., 1992a. Altruism in viscous populations—an inclusive fitness model. *Evol. Ecol.* 6, 352–356.
- Taylor, P.D., 1992b. Inclusive fitness in a homogeneous environment. *Proc. R. Soc. B* 249, 299–302.
- Taylor, P.D., 1996. Inclusive fitness and genetic models of behaviour. *J. Math. Biol.* 34, 654–674.
- Taylor, P.D., Frank, S.A., 1996. How to make a kin selection model. *J. Theor. Biol.* 180, 27–37.
- Thorne, B.L., 1997. Evolution of eusociality in termites. *Annu. Rev. Ecol. Syst.* 28, 27–54.
- Traniello, J.F., Leuthold, R.H., 2000. Behavior and ecology of foraging in termites, in: Abe, T., Bignell, D.E., Higashi, M. (Eds.), *Termites: Evolution, Sociality, Symbiosis and Ecology*. Kluwer Academic Publishers, Netherlands, pp. 141–168.
- Waite, T.A., Strickland, D., 1997. Cooperative breeding in Gray Jays: philopatric offspring provision juvenile siblings? *Condor* 99, 523–525.
- Wild, G., Koykka, C., 2014. Inclusive-fitness logic of cooperative breeding with benefits of natal philopatry. *Phil. Trans. R. Soc. B* 369, 20130361.
- Wild, G., Pizzari, T., West, S.A., 2011. Sexual conflict in viscous populations: the effect of the timing of dispersal. *Theor. Popul. Biol.* 80, 298–316.

- Wild, G., Taylor, P.D., 2004. Kin selection models for the co-evolution of the sex ratio and sex-specific dispersal. *Evol. Ecol. Res.* 6, 481–502.
- Wild, G., Taylor, P.D., 2005. A kin-selection approach to the resolution of a sex-ratio conflict between mates. *J. Theor. Biol.* 236, 126–136.
- Wiley, R.H., Rabenold, K.N., 1984. The evolution of cooperative breeding by delayed reciprocity and queuing for favorable social positions. *Evolution* 38, 609–621.

**Table 1:** Summary of notation used in the main text.

Symbol(s)	Explanation
$b$	Number of immature offspring produced on a territory during a given time period, and interpreted as fecundity or birth rate.
$c, \hat{c}$	Fraction of dominant vacancies per time period per territory that are contested by related individuals. A hat denotes the equilibrium value of the parameter.
$h$	Subordinate helping rate.
$m_\alpha, \hat{m}_\alpha$	Number of local dominant vacancies secured by a subordinate. This parameter can also be interpreted as a probability. A hat denotes the equilibrium value of the parameter.
$m_\beta, \hat{m}_\beta$	Likelihood that a subordinate does not compete successfully for a breeding opportunity as a dominant. Of course, $m_\beta = 1 - m_\alpha$ . A hat denotes the equilibrium value of the parameter.
$n_\alpha, \hat{n}_\alpha$	Number of dominant vacancies secured by a disperser. A hat denotes the equilibrium value of the parameter.
$\hat{r}$	Genetic relatedness between two different individuals born on the same territory during the same time period, at equilibrium.
$s_\alpha$	Survival rate of a dominant individual.
$s_\beta, \Delta s_\beta$	Survival rate of a subordinate, and the positive change in this rate following receipt of help from a subordinate born in the previous time period.
$s_\omega, \Delta s_\omega$	Survival rate of a disperser, and the negative change in this rate following donation of help to an individual born in the current time period.
$v_\alpha$	Reproductive value of a dominant individual.
$v_\beta$	Reproductive value of a subordinate individual, $v_\beta = s_\omega \hat{n}_\alpha v_\alpha$ .
$z$	The probability that an individual delays its dispersal from its natal territory for one time period for the chance to compete to inherit a dominant position.