



Diminishing returns drive altruists to help extended family

P. Kennedy¹✉, S. Sumner^{1,4}, P. Botha¹, N. J. Welton², A. D. Higginson³ and A. N. Radford¹

Altruism between close relatives can be easily explained. However, paradoxes arise when organisms divert altruism towards more distantly related recipients. In some social insects, workers drift extensively between colonies and help raise less related foreign brood, seemingly reducing inclusive fitness. Since being highlighted by W. D. Hamilton, three hypotheses (bet hedging, indirect reciprocity and diminishing returns to cooperation) have been proposed for this surprising behaviour. Here, using inclusive fitness theory, we show that bet hedging and indirect reciprocity could only drive cooperative drifting under improbable conditions. However, diminishing returns to cooperation create a simple context in which sharing workers is adaptive. Using a longitudinal dataset comprising over a quarter of a million nest cell observations, we quantify cooperative payoffs in the Neotropical wasp *Polistes canadensis*, for which drifting occurs at high levels. As the worker-to-brood ratio rises in a worker's home colony, the predicted marginal benefit of a worker for expected colony productivity diminishes. Helping related colonies can allow effort to be focused on related brood that are more in need of care. Finally, we use simulations to show that cooperative drifting evolves under diminishing returns when dispersal is local, allowing altruists to focus their efforts on related recipients. Our results indicate the power of nonlinear fitness effects to shape social organization, and suggest that models of eusocial evolution should be extended to include neglected social interactions within colony networks.

Altruism—sacrifice of Darwinian fitness to increase that of a recipient—is easily explained when occurring between close relatives¹, through an increase in the altruist's inclusive fitness (success at propagating copies of its genes in the population)^{2–4}. However, paradoxical forms of altruism arise when individuals divert help from the closest relatives and towards more distantly related recipients^{5,6}. This seemingly irrational behaviour appears to reduce inclusive fitness⁷.

Social insect colonies are often impenetrable fortresses⁸. However, in some (especially primitively eusocial) species, between-colony movement (drifting) by both workers^{6,7,9–12} and foundresses¹³ can be extensive. Often, opportunities for reproductive parasitism¹⁴ or nest inheritance¹³ provide clear direct fitness motives. In other cases, drifters lack obvious opportunities for direct fitness^{7,11,12} and help relatives that are more distant kin than recipients in their home colony. In some primitively eusocial *Polistes* paper wasps, the extreme extent of drifting by non-reproductive workers has become clear: in the tropics (where *Polistes* originated), 56% of workers within a population were detected in multiple colonies¹². Drifting workers perform standard cooperative tasks¹² (henceforth, cooperative drifting), creating extended kin groups¹¹ (networks of cooperating colonies). However, despite primitively eusocial insects being long-standing models for understanding trajectories to complex eusociality¹⁵, the driver of cooperative drifting has not been identified.

Three hypotheses have been proposed to explain the evolution of cooperative drifting by non-reproductive primitively eusocial workers. The bet hedging hypothesis¹² suggests that helping multiple related colonies (1, 2 and 3 in Fig. 1a) avoids the risk of investing in a single colony that could succumb to chance failure or predation. By diversifying investments, workers accept reduced expected inclusive fitness for the benefit of reduced variance in inclusive fitness. The indirect reciprocity hypothesis⁵ suggests that helping in

partner colonies (2 in Fig. 1b) leads to other workers (3 in Fig. 1b) —who may be non-relatives—helping the home colony (1 in Fig. 1b). For this to be beneficial, it is assumed that help exchanged between neighbours is worth more than help from natal workers due to social heterosis: a negative relationship between relatedness and ability to improve colony productivity¹⁶. The diminishing returns hypothesis^{7,12} suggests that the marginal benefit provided by a worker diminishes as the number of workers tending the brood increases^{17,18} (1 in Fig. 1c), an effect first highlighted by Michener in 1964¹⁹, and helping in related colonies allows effort to be redirected towards brood that are more in need of care (2 in Fig. 1c). There has been no formal comparison of these hypotheses, so we begin by using inclusive fitness theory to assess the plausibility of each.

Results

Selection for cooperative drifting. First, we show that, under normal conditions, bet hedging¹² cannot select for cooperative drifting. We derive an explicit mean-risk trade-off for bet hedging traits for an organism maximizing its inclusive fitness³ (Methods section 'Bet hedging'):

$$\max_{\mathbf{q}} \left\{ \mathbf{q}^\dagger \boldsymbol{\mu} - \nu \rho \sqrt{\mathbf{q}^\dagger \mathbf{C} \mathbf{q}} \right\} \quad (1)$$

$$\text{subject to } \mathbf{q}^\dagger \mathbf{1} = 1$$

where the organism must choose the optimal weights (\mathbf{q}) to place on different investments, balancing the expectation ($\mathbf{q}^\dagger \boldsymbol{\mu}$, where \dagger denotes transpose) and variance ($\mathbf{q}^\dagger \mathbf{C} \mathbf{q}$) in absolute inclusive fitness returns, for particular values of risk aversion (ν) and the correlation (ρ) between the portfolio's return and whole-population

¹School of Biological Sciences, University of Bristol, Bristol, UK. ²Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

³Centre for Research in Animal Behaviour, University of Exeter, Exeter, UK. ⁴Present address: Centre for Biodiversity and Environment Research (CBER), Division of Biosciences, University College London, London, UK. ✉e-mail: patrick.kennedy@bristol.ac.uk

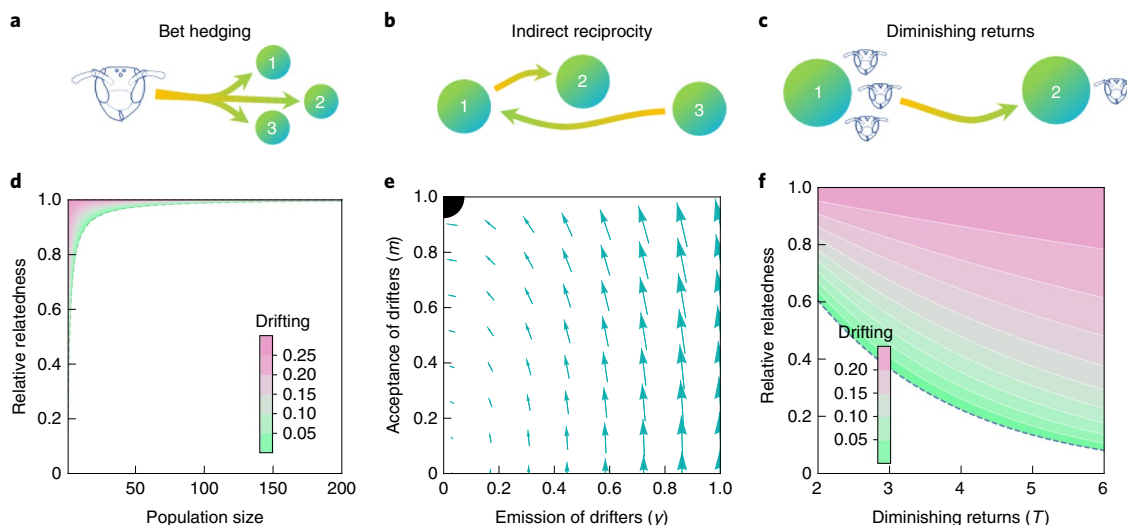


Fig. 1 | Three adaptive hypotheses have been proposed for cooperative drifting. **a–c**, The three adaptive hypotheses for cooperative drifting: bet hedging (**a**), indirect reciprocity (**b**) and diminishing returns (**c**). **d**, Cooperative drifting is only favoured by bet hedging (above the dashed line) at extremely small population sizes and/or when recipients differ only slightly in relatedness to the actor. Purpler colours denote higher levels of drifting at equilibrium, for $n_1=100$ and $n_2=50$ (Methods section ‘Bet hedging’). **e**, Cooperative drifting cannot be explained by the indirect reciprocity hypothesis because free riders invade ($y \rightarrow 0$; $m \rightarrow 1$). The arrows show the direction of selection. The black quarter-circle is the global attractor (see Methods section ‘Indirect reciprocity’). **f**, Diminishing returns can select for cooperative drifting. Stronger diminishing returns and higher relatedness favour more cooperative drifting (see Methods section ‘Diminishing returns’). Relative relatedness is d in equation (3). For illustration, equal reductions in relatedness are plotted for female and male brood ($d_\sigma = d_\varphi = d$). $\psi=1$, $g=0.5$, $x=0.75$ and $z=0.25$.

average reproductive success (\bar{w}). The vector μ contains the offspring production per unit invested (weighted by offspring relatedness), C is the variance-covariance matrix for the investment payoff rates and $\mathbf{1}$ is a vector of ones. A risky portfolio has high variance. In the Methods section ‘Bet hedging’, we show that the value of risk aversion (v) that maximizes inclusive fitness is the coefficient of variation in \bar{w} . However, under demographic stochasticity—which is generated by random colony failures—meaningful fluctuations in \bar{w} will only occur if the population is extremely small (Methods section ‘Bet hedging’), so $v \approx 0$. Moreover, in reasonably sized populations, fluctuations in \bar{w} that happen due to random colony failure occur independent of fluctuations in the portfolio’s returns, so $\rho \approx 0$. Since $v \approx 0$ (and $\rho \approx 0$), the second term of maximand (1) is negligible, so workers should maximize only the expectation of inclusive fitness ($\mathbf{q}^T \mu$). The same logic underpins why bet hedging against randomly occurring clutch failure is an unlikely explanation for birds distributing eggs over multiple nests²⁰ or parasitoids distributing eggs over multiple hosts²¹. In summary, the bet hedging hypothesis for cooperative drifting is mathematically coherent (Fig. 1d) only in an extremely small population and/or in cases where drifting would lead to a very small reduction in expected inclusive fitness $\mathbf{q}^T \mu$, circumstances that only rarely arise in the natural world. One way, for instance, is if there were near-negligible differences in relatedness towards brood in home and neighbouring colonies, but *Polistes* drifters face decreases in relatedness²².

Second, we show that indirect reciprocity⁵ is unlikely to explain cooperative drifting. Using simulations, Nonacs⁵ has argued that drifters will automatically help only at colonies sharing indirect reciprocator alleles because these are the only colonies willing to accept them. However, there is no reason to expect pleiotropy between these two behaviours, so (unlike ref. ⁵) we allow the emission (y) and acceptance (m) of drifters to evolve independently in a kin selection model (Methods section ‘Indirect reciprocity’). As in ref. ⁵, we find that when m is prevented from decreasing while y increases, drifting can evolve (Methods section ‘Indirect reciprocity’).

However, when this unjustified assumption is relaxed, indirect reciprocity collapses due to the invasion of free riding: colonies willingly accept foreign workers¹⁰ ($m=1$) while none of their own workers drift ($y=0$) (Methods section ‘Indirect reciprocity’ and Fig. 1e).

Third, we show that diminishing returns^{7,12} can be a simple driver of cooperative drifting (Methods section ‘Diminishing returns’): as the effect of helping closely related recipients declines, diversion of altruism towards more distantly related recipients can satisfy Hamilton’s rule (the increase in benefit b more than compensates for the decrease in relatedness r). In a kin selection model using the framework of Davies et al.²², we let the probability of a brood surviving to maturity (K) be determined by a simple diminishing returns function: $K = 1 - (1 - h)^T$, where $0 \leq h \leq 1$ is help received from workers. Higher values of T result in stronger diminishing returns. Help (h) received by each brood depends on the worker-to-brood ratio ψ in their colony, which may vary stochastically between and within colonies through time. For simplicity, we consider neighbouring colonies with equal brood numbers lying at different points on a line of possible worker-to-brood ratios. Worker-to-brood ratios greater than $\psi=1$ occur only in colonies in terminal decline, so we focus on $0 < \psi \leq 1$. Drifting reduces a worker’s relatedness to the brood it cares for to the (nonzero) proportions d_φ and d_σ for female and male brood, respectively, relative to raising siblings at home. When a home colony has worker-to-brood ratio ψ , selection favours a small increase in drifting (y) to a colony with a proportion g of the home colony’s workforce when the improvement in indirect fitness from escaping diminishing returns (left-hand side) compensates for reductions in recipient relatedness (right-hand side):

$$\frac{(1-\psi x(1-z)(g+y))^{T-1}}{(1-\psi x(1-z)(1-y))^{T-1}} > \frac{4}{(3d_\varphi+d_\sigma)} \quad (2)$$

where x is the proportion of females that are behaviourally sterile (not reproductive) and z is the proportion of offspring that are male, so $x(1-z)$ is the proportion of offspring that are workers.

To find candidate evolutionarily stable strategies (ESSs) for drifting level y , we set equal the left and right sides of equation (2) and rearrange for y . Letting consanguinity to male and female brood be devalued by the same amount ($d_q = d_g = d$), the ESS drifting level y^* occurs at:

$$y^* = \frac{1}{1 + \sqrt[r]{d}} \left[1 - \sqrt[r]{d}g - \frac{1 - \sqrt[r]{d}}{\psi x(1-z)} \right] \quad (3)$$

Given the possibility to drift to related colonies at sufficiently steeper points on an inclusive fitness returns curve, workers should tolerate a reduction in recipient relatedness (Methods section ‘Diminishing returns’, Fig. 1f and Extended Data Figs. 1 and 2). The ESS drifting level y^* increases whenever a gap in worker payoff increases between two related colonies: y^* increases with higher T (Fig. 1f), higher ψ and lower g , and for scenarios in which there is a larger workforce (lower z and higher x). For example, a worker may be twice as related to a brood on its home colony as to a brood on a partner colony ($d=1/2$). Assume the home colony (with $\psi=1$) has twice as many workers as the partner colony ($g=1/2$). With a diminishing returns factor of $T=4$, three-quarters of females developing as non-reproductive workers ($x=3/4$) and a sex ratio of 3:1 female eggs to male eggs ($z=1/4$), the expected equilibrium level of cooperative drifting from the home colony would be for 13.2% of worker effort to be divested to the less related partner colony ($y^*=0.132$). As workers drift and group sizes consequently change, the marginal payoff of staying in the home colony increases and the marginal payoff of helping in the partner colony decreases until $y=y^*$, where any further drifting would reduce inclusive fitness.

Diminishing payoffs in the wild. To investigate plausible ranges within which diminishing returns to cooperation may exist, we used longitudinal field tracking of brood development and worker numbers in a Neotropical paper wasp. *P. canadensis* is a model species in which cooperative drifting is common and can reduce relatedness to recipient brood considerably (for example, $r=0.56$ at home versus $r=0.19$ on partner colonies)¹². Colonies typically consist of a single queen and behaviourally non-reproductive daughter workers and non-nestmate drifters¹². As workers emerge, colonies can grow in group size from fewer than ten females to over 200 females. In dominance hierarchies below the queen, a small number of high-ranking wasps have an opportunity to inherit the nest on the death of the resident queen²³; most subordinates perform helping behaviours (including foraging, nest hygiene, brood inspection, nest building and nest defence)^{12,24,25}. The lack of a covering nest envelope allows clear observation of individual brood cells (Fig. 2a) and the opportunity to document whole-colony development through time²⁶.

We aimed to assess how a colony’s success at producing new adults is associated with numbers of workers and brood in the colony. Across 56 d, we made over a quarter of a million observations of more than 20,000 individually tracked brood cells on 91 wild post-worker-emergence colonies in Panama. The stage of brood development in every cell in each colony was recorded repeatedly (Methods section ‘*P. canadensis* payoffs’ and Extended Data Fig. 3) and treated as a state in a Markov model (Fig. 2b,c). The number of adults on the nest at night (when all workers were present) was counted at intervals over the observation period. This resulted in a dataset of 123,116 state transitions involving live brood in 85 colonies in 471 colony observations for which the worker number could be predicted by interpolation through the night censuses (Methods section ‘*P. canadensis* payoffs’). To investigate the extent to which there may be variation in payoffs within networks of colonies, we asked how the number of workers correlated with colony success at different points of larval development.

Using between-colony variation in the number of workers and worker-to-brood ratio, while controlling for within-colony

variation and colony state (using the extent of brood cell emptiness as a proxy), a Bayesian hierarchical model predicted that colonies with higher worker-to-brood ratios and worker numbers are associated with a higher brood development pace and fewer brood deaths. The Markov model predicts that the expected time for a single egg-containing brood cell to produce an adult successfully in colonies with different worker and brood numbers (that is, the expected mean first passage time (eMFPT)) decreases as more workers tend the brood (Fig. 3a,b and Extended Data Fig. 4). Multiplying $1/\text{eMFPT}$ by the brood number estimates the expected number of new adults that can be produced per day (whole-colony productivity), which is highest in colonies with many workers and many brood (Fig. 3c,d). The slope of whole-colony productivity with respect to worker number (Fig. 3e,f) then provides a prediction of the marginal increase in the daily number of brood successfully raised associated with each additional worker (that is, plausibility values for the payoff rate). The payoff in Fig. 3e,f represents an empirical estimate of the benefit b in Hamilton’s rule from working on a colony of the given size. An investment in a partner colony 2 is in the inclusive fitness interest of a worker from a home colony 1 if $r_2 b_2 > r_1 b_1$.

The difference in the predicted payoff from the model suggests that, within the main parameter space occupied by colonies, workers are more valuable (lighter colours in Fig. 3e) in colonies with a relatively understaffed workforce faced with large brood-rearing challenges than in colonies with fewer brood to rear. Future models that consider within-colony dynamics may further explain finer-scale variation in brood development rates (see Extended Data Figs. 5–8 and Supplementary Table 3). However, the prediction of variation in plausible payoff rates between colonies with different workforce sizes suggests a context in which workers can increase indirect fitness by helping at less closely related colonies.

Scales of competition. To explore the demographic conditions that can sustain cooperative drifting, we used agent-based haplodiploid simulations (Methods section ‘Individual-based simulation’) of a large population of monogynous colonies distributed over a square lattice (where each colony had eight neighbours in a Moore neighbourhood). To model simple colony growth, we assumed that a unit is a group of workers (for example, ten workers) and let colonies produce a new unit every three time steps (for example, 3 weeks) until a maximum of ten units was reached. We assumed a saturated environment in which each square on the lattice was occupied by one colony. At each time step, a randomly chosen 10% of colonies died and were replaced. The foundress of the replacement colony was drawn either from lottery competition among the local eight colonies in the Moore neighbourhood (when female philopatry was assumed) or globally from the whole population (when female philopatry was relaxed). Males compete globally for mating with females in lottery competition. The number of reproductively destined offspring that colonies produce is a diminishing returns function of the helper effort h on the colony, $1 - (1-h)^T$, as above. Colonies with more helpers therefore produce more reproductively destined females and males, and so have a greater chance in competition for nest sites and mating, respectively. We evolved units’ propensity to drift for workers in colonies with more than five units and showed the results of competition by introducing mutant alleles to resident populations at an initial frequency of 5%. After 1,000 time steps, we recorded the mutant frequency and plotted the average change in frequency over three replicate simulations (Methods section ‘Individual-based simulation’).

Under linear returns, drifting does not invade, regardless of demography (Fig. 4a–c). We then considered moderate diminishing returns ($T=3$) under three conditions: female philopatry and altruism directed at local colonies in the Moore neighbourhood (Fig. 4d); female philopatry and altruism directed at partner colonies whose

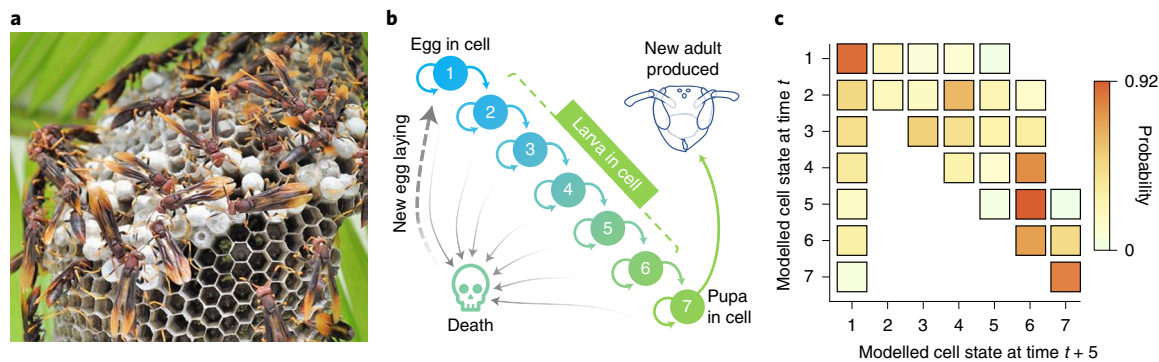


Fig. 2 | Brood development in the Neotropical paper wasp *P. canadensis*. **a**, *P. canadensis* colonies allow easy observation of individual brood cells. **b**, Wasp development involves states from egg (state 1) through larvae (2–6) to pupa (7) and finally a new adult (definitions of states are provided in Extended Data Fig. 3). This can be represented using a Markov model. **c**, Baseline transition probabilities between developmental states using an intercepts-only model (see Methods section ‘*P. canadensis* payoffs’). In using developmental rates to produce a measure of colony productivity as a function of worker behaviour, we set transitions to death as transitions back to the egg (state 1), to exclude confounding effects from between-queen variation in egg-laying rates. Credit: P. Kennedy (a).

queens were parents or daughters of their own queen (Fig. 4e); and female global dispersal with altruism directed at local colonies in the Moore neighbourhood (Fig. 4f). (Female global dispersal with altruism directed at the genealogically close partner colonies was not considered because global female dispersal makes these pairings vanishingly rare.) Cooperative drifting can invade only when we allow female philopatry and kin-directed altruism (Fig. 4e). Under this scenario, neighbouring nests develop genealogical links (spatial kin structure), and when drifters are shared preferentially within these links cooperative drifting occurs at a more local scale than the spatial scale of competition. *Polistes* wasps often form colony clusters within wider aggregations of tens to hundreds of colonies^{11,12,27}, a context in which cooperative drifting can evolve by spatial selection under diminishing returns.

Discussion

Established accounts of the evolution of eusociality assume that actors must choose either to stay as helpers or leave as reproductives^{8,28}. Our results suggest that diminishing returns may drive altruists to diversify their help across recipients: workers in some primitively eusocial societies may increase inclusive fitness by providing altruism to recipients beyond their home colony. Under positive kinship, spatial kin clustering and diminishing returns^{17,29}, worker investments can evolve to become diffusible public goods.

Our model predicts the conditions under which we expect cooperative drifting to have evolved (equation (2) and Fig. 1f). Intuitively, drifting is more likely with stronger diminishing returns (higher T), a larger difference in workforce between nests (smaller g), an increased total workforce (higher ψ) and a greater capacity to target kin (higher d_0 and d_3). For simplicity in equation (2), we assume that all colonies have the same sex ratio, but between-colony sex ratio variation suggests an additional factor: a colony producing mainly brothers has a reduced worker relatedness to the brood, at which point switching colonies may be rational for a worker. In short, drifting offers a simple route to boost inclusive fitness when neighbouring colonies differ in parameters that determine the value of a worker. Differences in worker and brood number arise easily among *P. canadensis* colonies (Fig. 3a), which are subject to several sources of stochasticity. These include fluctuations in worker number due to the high attrition rate of foraging workers¹², frequent loss of brood to parasitoids, presumed loss of brood due to disease (based on workers’ hygienic removal of larvae), episodes of queen replacement, and so on. Fluctuations in brood cohort size translate to fluctuations in workforce size once the brood pupates.

Since Michener¹⁹ highlighted diminishing returns in hymenopteran societies in 1964, a number of studies across ants, bees and wasps have revealed declines in per-capita productivity with rising group size (for example, refs. 18,30–33). This so-called reproductivity effect has not proven universal (for example, refs. 34–36), but its frequent occurrence leads to Michener’s paradox: why do apparently partly redundant helpers exist^{26,30}? Previous tests of the reproductivity effect have used snapshots of per-capita productivity. In contrast, we provide a prediction of plausible ranges for the payoffs of cooperation in a primitively eusocial insect using colony dynamics. Diminishing returns exist, but through cooperative drifting workers can mitigate redundancy arising from stochastic variation in worker-to-brood ratios between colonies.

The extent of drifting across primitively eusocial insects remains to be explored^{5,12}. However, the relatively high levels of drifting observed in Neotropical species such as *P. canadensis* contrast with, for example, the European wasp *P. dominula*, which also forms dense colony aggregations¹³ but shows high aggression towards neighbours. This difference in social organization may be due to differences in the intensity of diminishing returns (for example, due to food availability or parasite density). Higher stochastic predation of workers in some species may undermine workers’ abilities to track need across nests. Alternatively, drifting may be more likely in the tropics: unlike temperate species in which foundresses establish nests synchronously in the spring, tropical species often establish nests throughout the year²⁵, and so nests may be more likely to differ in worker-to-brood ratio. Tropical species may also experience less uncertainty in neighbour relatedness, since nests are more commonly founded by local dispersal from parent nests (simulated in Fig. 4e), although kin spatial structure can be re-established in temperate species by natal philopatry of spring foundresses³⁷. Direct comparisons between species with and without cooperative drifting are needed.

Cooperative drifting has also emerged among complex eusocial species. Ant supercolonies exist when nests with multiple queens (polygyny) exchange workers (polydomy)^{6,38}. Supercoloniality results in remarkably low relatedness cooperation and remains a theoretical challenge. The evolution of supercoloniality is likely to involve informational constraints preventing nepotism⁶, although some positive relatedness may be maintained by cryptic kin structure³⁹. Our results are relevant here: the initial drivers of low relatedness cooperation are unlikely to have been either bet hedging by risk spreading at the expense of the expectation of inclusive fitness (equation (1) and Fig. 1d) or the reciprocity scenario

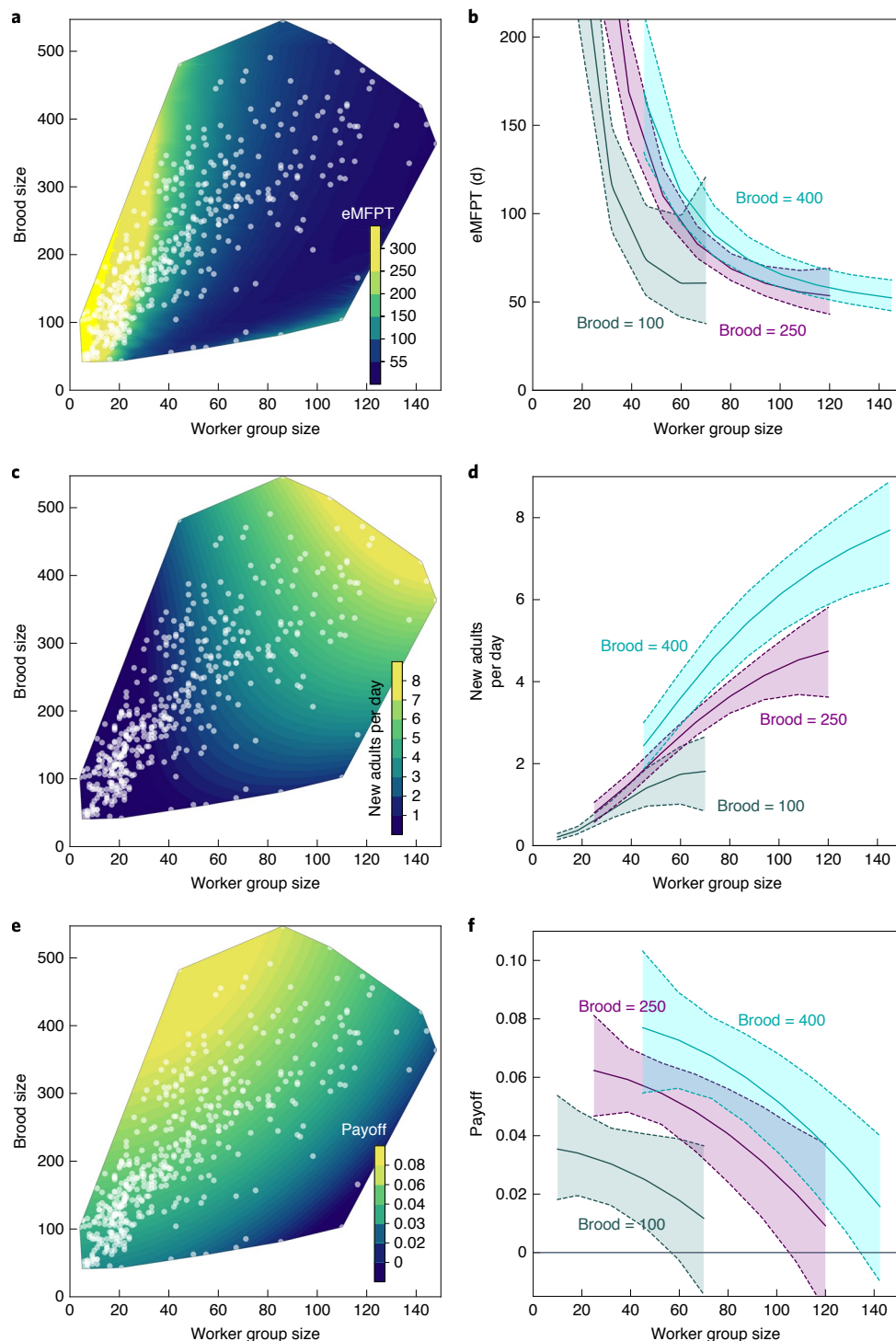


Fig. 3 | Brood-rearing rates in *P. canadensis*. **a**, Expected number of days for an idealized brood cell to successfully produce an adult (eMFPT) predicted by between-colony effects. Brood size denotes the number of prepupal brood (states 1–6). Worker group size denotes the interpolated group size using nighttime censuses. The white dots denote colony observations. To illustrate predictions in the main parameter space, we generated the prediction of eMFPT for 1,000 simulated colonies inside the convex hull set by those colony observations with up to 150 workers and over 40 brood (only ten colony observations (from six colonies) in the dataset of 471 colony observations used to generate the model lay outside of this range). **b**, 95% credible intervals corresponding to **a**. **c**, The predicted whole-colony productivity (daily expected number of new adults, given the eMFPT values) is highest in colonies with many workers and many brood. **d**, 95% credible intervals corresponding to **c**. **e**, Slope of the predicted whole-colony productivity with respect to worker group size, representing the predicted effect of adding a new worker (the payoff provided to the recipients)—an empirical estimate of the potential benefit b in Hamilton's rule. **f**, 95% credible intervals corresponding to **e**. See Methods section '*P. canadensis* payoffs'.

proposed by ref. ⁵ (Fig. 1e). In principle, diminishing returns may initially have favoured partial diversion of altruism to more distantly related colonies. However, supercolony and primitively eusocial

cooperative drifting are not completely analogous. Supercolony may have been a relatively small step for ants that had already evolved high within-colony polygyny—and consequently reduced

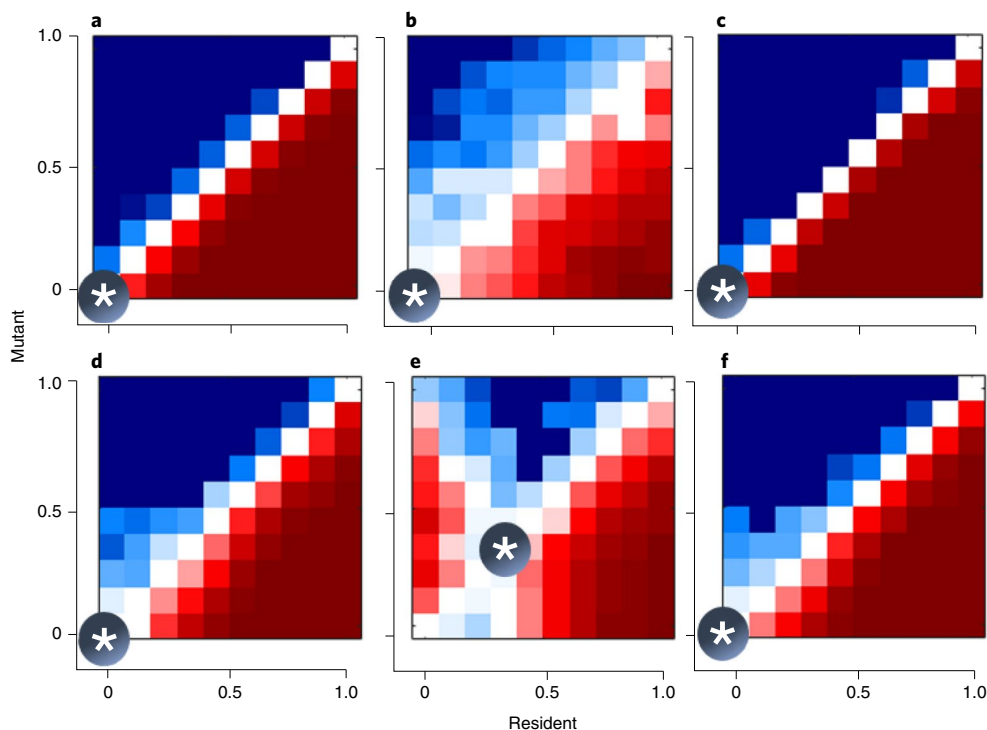


Fig. 4 | Evolution of cooperative drifting in a spatially explicit social haplodiploid simulation. a–f, Results of competition from agent-based simulations of the invasion of drifting. Red denotes mutants invading from a starting frequency (p_{initial}) of 5% of the population, whereas blue denotes mutants declining. White denotes a negligible change in frequency. Convergence-stable strategies are marked by asterisks. The conditions were no diminishing returns ($T=1$) (**a–c**) and moderate diminishing returns ($T=3$) (**d–f**), with female philopatry and altruism directed at local colonies in the Moore neighbourhood (**a** and **d**), female philopatry and altruism directed at partner colonies with queens who were parents or daughters of their own queen (**b** and **e**) or global female dispersal with altruism directed at local colonies in the Moore neighbourhood (**c** and **f**). Nonlinear returns drive inter-colony cooperation only under both local female dispersal and kin-directed altruism (**e**).

relatedness⁴⁰—for other reasons. Unlike primitively eusocial wasps, the first step to explaining cooperative drifting in ants is explaining polygyny⁴¹.

Manipulating colony networks by adjusting the worker-to-brood ratio (ψ) may offer tests of whether wasps make strategic adjustments to investments (y). Empirical studies are needed to identify whether host workers discern cooperative drifters and adjust acceptance thresholds (m) adaptively^{42,43} according to need. Future theoretical work could assess the tension between selfish and cooperative drifting in determining the acceptance of foreign workers. Wasps with high resource-holding potential may exploit the relaxation of nest boundaries to drift for direct fitness (for example, joining dominance hierarchies on multiple nests to maximize the chance of nest inheritance). Models of the mechanisms individual workers might use to distribute their effort would be useful, potentially inspired by resource-use models in foraging theory⁴⁴.

Nonlinear payoffs exert strong effects on social evolution: diminishing returns can limit the tragedy of the commons⁴⁵, promote polymorphic equilibria⁴⁶ and increase sharing in reproductive skew games⁴⁷. However, the extent to which diminishing returns shape investment patterns remains little quantified, despite clear theoretical predictions. A tempting explanation for divestment across recipients is that actors help different recipients in proportion to relatedness (an idea known as the proportional altruism model⁴⁸). This argument commits the ‘gamblers’ fallacy⁴⁹: instead, it is always best to invest in the single recipient who carries the highest inclusive fitness returns at any one time⁵⁰. To explain altruism towards more distant relatives, in the 1980s Altmann⁴⁹, Weigel¹⁷ and Schulman and Rubenstein²⁹ highlighted diminishing returns to investment by a single individual. Here, we have considered diminishing returns

to investment by multiple individuals. In both cases, diminishing returns provide a simple explanation for helping multiple recipients, which awaits empirical study in many species. Our results indicate the power of nonlinear fitness effects to shape social organization, and suggest that models of eusocial evolution should be extended to include neglected social interactions within colony networks.

Methods

Tables of notation and details relating to the models are available in the Supplementary Information, and additional figures are provided as Extended Data Figs. 1–9.

Bet hedging. We consider when an inclusive fitness maximizing actor should divide its investments between separate recipients to minimize the risk that its investments will be lost². Notation is summarized in Supplementary Table 1. Following Grafen³, we start with the Price equation under uncertainty and treat the target of selection as an individual maximand. The expected change in allele frequency due to selection $\Delta\bar{G}$ —where average reproductive success for the population is \bar{w} , and I_i is the absolute inclusive fitness—is equal to the covariance over individuals i between the expectation of relative inclusive fitness $\frac{I_i}{\bar{w}}$ and the individual’s genotype G_i :

$$\mathbb{E}_\omega[\Delta\bar{G}] = \text{Cov}_i(G_i, \mathbb{E}_\omega[\frac{I_i}{\bar{w}}]) \tag{4}$$

where expectations \mathbb{E} are taken across possible states of the population $\omega \in \Omega$ that may occur. The absolute inclusive fitness is the sum of all effects b_{jt} of actor i on the absolute fitness of recipient j (including the actor itself) across each role t (that is, $I_i = \sum_j^N r_{ij} b_{jt}$). For derivation of equation (4), see the first term on the right-hand side of equation (6) in Grafen³. Under fitness additivity and frequency independence, the quantity $\mathbb{E}_\omega[\frac{I_i}{\bar{w}}]$ is considered to represent a maximand of organismal behaviour³: it is a target of selection (as it covaries with genotype) that is under the control of the actor since I_i is composed of the effects of the actor’s behaviours.

The expectation of a ratio of random variables can be approximated by the Taylor series expansion. The Taylor series is an acceptable approximation when

\bar{w} does not rise to greater than $2\mathbb{E}_\omega[\bar{w}]$ (ref. 51). We expand the maximand to the second order as follows:

$$\mathbb{E}_\omega\left[\frac{I_i}{\bar{w}}\right] \approx \frac{1}{\mathbb{E}_\omega[\bar{w}]} \left(\mathbb{E}_\omega[I_i] - \frac{\sigma_\omega[\bar{w}]}{\mathbb{E}_\omega[\bar{w}]} \rho \sigma_\omega[I_i] \right) \quad (5)$$

where $\sigma_\omega[\cdot]$ denotes standard deviation over the states $\omega \in \Omega$, and ρ denotes the correlation between I_i and \bar{w} . The division by $\mathbb{E}_\omega[\bar{w}]$ does not affect the optimal decision, so we focus on the terms in the brackets.

To make more explicit the quantity under a focal organism's control, we can describe as \mathbf{q} the vector of investment weights (the proportions of its total resource) that the individual chooses to place on different investments that affect the reproductive success of itself and other individuals. The expectation of the focal individual's absolute inclusive fitness ($\mathbb{E}_\omega[I_i]$) is given by $\mathbf{q}^\dagger \boldsymbol{\mu}$, where $\boldsymbol{\mu}$ is the vector of expected inclusive fitness payoffs from the different potential options (the expectation over the different states $\omega \in \Omega$). By convention, \dagger denotes transpose. Likewise, the standard deviation of a focal individual's absolute inclusive fitness appearing in equation (5) ($\sigma_\omega[I_i]$) is given by $\sqrt{\mathbf{q}^\dagger \mathbf{C} \mathbf{q}}$, where \mathbf{C} is the variance-covariance matrix of the payoffs over the states $\omega \in \Omega$, and $\mathbf{q}^\dagger \mathbf{C} \mathbf{q}$ gives the variance (over the states $\omega \in \Omega$) of the individual's absolute inclusive fitness.

The aim for the organism can then be expressed using the following maximand (echoing financial portfolios⁵²):

$$\max_{\mathbf{q}} \left\{ \mathbf{q}^\dagger \boldsymbol{\mu} - \nu \rho \sqrt{\mathbf{q}^\dagger \mathbf{C} \mathbf{q}} \right\} \quad (6)$$

where the sum of \mathbf{q} is 1. The coefficient of variation in population average reproductive success (ν) is not affected by the organism's choice of \mathbf{q} . Whereas in economics, individuals have subjective risk aversions, risk aversion in biology is imposed by the environment: a higher ν makes individuals more averse to having an inclusive fitness investment portfolio that has correlated fluctuations with population average reproductive success. An environment can have high ν or low ν , and this is imposed on the organism. The bet hedging hypothesis implies that an optimization trade-off exists within \mathbf{q} , balancing the expectation and the variance in inclusive fitness profit (measured in the absolute number of offspring equivalents produced).

Equation (7) highlights that variance in inclusive fitness ($\mathbf{q}^\dagger \mathbf{C} \mathbf{q}$) only matters when the level of stochasticity ν (that is, the coefficient of variation in population average reproductive success \bar{w}) and the correlation ρ are non-negligible. This is generally only true when there is environmental stochasticity that makes a genotype's total reproductive success fluctuate in a correlated fashion between environmental states. For instance, there may be some distinct environmental states when one type of colony does worse, such as big colonies failing when there is a drought. However, this would require a genotype starting from a position of overinvesting workers in one type of colony, such that it may then be able to reduce its variance by diverting some effort towards a different type of colony that does not fail in that environmental state.

The bet hedging hypothesis proposed by Sumner et al.¹² is based on hedging against individual risks to colonies: demographic stochasticity. In general, demographic stochasticity can only generate meaningful fluctuations in \bar{w} when the population (or deme) size is very small^{53,54}. Next, we illustrate the effect of population size, which is plotted in Fig. 1d of the main text.

To illustrate the bet hedging hypothesis in a specific example, we switch to a neighbour-modulated perspective⁵⁵. Notation is summarized in Supplementary Table 2. We focus on the effects experienced by recipients due to the actions of related actors. Thus, instead of focusing on the expected relative inclusive fitness $\mathbb{E}_\omega\left[\frac{I_i}{\bar{w}}\right]$ of an actor i , we focus on the expected relative fitness $\mathbb{E}_\omega\left[\frac{w_j}{\bar{w}}\right]$ of a recipient j . The Taylor approximation allows us to write the condition for selection of a trait of interest in a population undergoing pure demographic stochasticity as:

$$\begin{aligned} \text{Cov}_j \left(G_j, \frac{1}{\mathbb{E}_\omega[\bar{w}]} \left(\mathbb{E}_\omega[w_j] - \frac{\text{Cov}_\omega[w_j, \bar{w}]}{\mathbb{E}_\omega[\bar{w}]} \right) \right) &> 0 \\ \therefore \text{Cov}_j \left(G_j, \mathbb{E}_\omega[w_j] - \frac{\text{Var}_\omega[w_j]}{N \mathbb{E}_\omega[\bar{w}]} \right) &> 0 \end{aligned} \quad (7)$$

since $\text{Cov}_\omega[w_j, \bar{w}] = \rho_{w_j, \bar{w}} \text{Var}_\omega[w_j]$ and, under pure demographic stochasticity⁵⁶, $\rho_{w_j, \bar{w}} = \frac{1}{N}$.

We consider the following scenario. Workers invest in colony reproductive success (where w_j is the reproductive success of the colony's queen) with linear returns $w_j = A n_j$, where n_j is the number of workers helping at queen j 's colony and A is a constant. After this investment period, queens are exposed to random catastrophe (such as a predation of the nest) with independent probability θ , which reduces their reproductive success to a proportion k of its value. Workers must decide during the investment period whether to invest solely in their home colony, where the queen is the closest related queen, or to divest some of their investment to neighbour colonies.

We assume that there are two sizes of colony with equal numbers of brood: colony type 1 and colony type 2. Type 1 and type 2 colonies have a high (n_1) and low (n_2 , where $0 < n_2 < n_1$) number of workers, respectively. For each genotype, we let half of the colonies be type 1 and half be type 2, then pair each type 1 colony

with a type 2 colony. For simplicity, we assume a haploid asexual population (that is, workers share their queen's allele for the trait of interest). We ask whether workers on a type 1 colony should divest part of their investments to a type 2 colony in order to hedge against the risk of their investments on the home nest being lost to random colony failure. If workers divest from a type 1 colony, they are paired with a foreign queen on a type 2 colony identical at the focal locus with probability α or a random queen (who may or may not be identical at the focal locus) on a type 2 colony with probability $1 - \alpha$. A mutant worker from a type 1 colony with divestment level y will expend a proportion y of its effort on the neighbouring type 2 colony and a proportion $1 - y$ on its own colony. We assume a population monomorphic for a resident strategy \bar{y} .

The absolute fitness w_2 of a queen on an n_2 colony carrying a mutant allele y if no catastrophe occurs is:

$$w_2 = A \{ n_2 + n_1 (\alpha y + (1 - \alpha) [P y + (1 - P) \bar{y}]) \} \quad (8)$$

where P is the frequency of the mutant allele in the population. The absolute fitness w_1 of a queen on an n_1 colony carrying the mutant allele is as follows if no catastrophe occurs:

$$w_1 = A n_1 (1 - y) \quad (9)$$

There are then four outcomes for any given queen: (1) starts with few workers (n_2) and no catastrophe occurs: w_2 ; (2) starts with few workers (n_2) and catastrophe occurs: $k w_2$; (3) starts with many workers (n_1) and no catastrophe occurs: w_1 ; and (4) starts with many workers (n_1) and catastrophe occurs: $k w_1$.

The probability of each outcome is: (1) starts with few workers (n_2) and no catastrophe occurs: $\frac{1-\theta}{2}$; (2) starts with few workers (n_2) and catastrophe occurs: $\frac{\theta}{2}$; (3) starts with many workers (n_1) and no catastrophe occurs: $\frac{1-\theta}{2}$; and (4) starts with many workers (n_1) and catastrophe occurs: $\frac{\theta}{2}$.

Accordingly, the variance of w over population states ω is:

$$\begin{aligned} \text{Var}_\omega[w] &= \frac{1}{2} \\ & \left((1 - \theta) ((w_2 - \mathbb{E}_\omega[w])^2 + (w_1 - \mathbb{E}_\omega[w])^2) + \theta ((k w_2 - \mathbb{E}_\omega[w])^2 + (k w_1 - \mathbb{E}_\omega[w])^2) \right) \end{aligned} \quad (10)$$

where:

$$\mathbb{E}_\omega[w] = \frac{1-\theta}{2} (w_1 + w_2) + \frac{\theta}{2} k (w_1 + w_2) \quad (11)$$

To ask whether a small increase in the level of altruism divested to relatives ($y - \bar{y}$) can invade, we evaluate the derivative of the approximation of expected relative fitness (which covaries with genetic value in equation (8)). We set this to equal 0 to find the candidate equilibria y^* :

$$\left. \frac{\partial \left(\mathbb{E}_\omega \left[\frac{w_j}{\bar{w}} \right] - \frac{\text{Var}_\omega \left[\frac{w_j}{\bar{w}} \right]}{N \mathbb{E}_\omega[\bar{w}]} \right)}{\partial y} \right|_{y=\bar{y}} = 0 \quad (12)$$

As y^* is too complex to give intuition, we plot y^* for a range of population sizes (N) and differences in relatedness between the home and partner colony in Fig. 1d (that is, when relative relatedness is 1, there is no difference between a worker's relatedness to the home colony and the partner colony). The region in which divestment to multiple relatives is favoured is very narrow.

Indirect reciprocity. In this section and the 'Diminishing returns' section below, we apply the Taylor-Frank method⁵⁸ to a haplodiploid population, using the framework of Davies et al.²², where x is the level of altruistic sterility, z is the sex ratio (proportion male), y is the probability of drifting and m is the probability of accepting incoming drifters. An overview with further details is given in Supplementary Section 2, with notation in Supplementary Table 1 and the consanguinity values for haplodiploids in Supplementary Table 2.

We model the indirect reciprocity hypothesis for drifting between unrelated colonies proposed by ref. 5 in a haplodiploid population. Reference⁵ argues that a colony's willingness to accept drifters (m) is an honest signal of its cooperative willingness to emit drifters, allowing the evolution of drifting in the absence of genealogical kinship. We follow Nonacs⁵ in letting m be the tag used passively to determine whether to drift to a target colony: if the colony will let the worker in, it enters. Accordingly, we look for ESS, y^* and m^* .

We randomly match each colony with a local unrelated recipient colony. Capital letters denote the mean value of the trait in a social group. We assume that the probability with which females develop as sterile helpers is the same in all colonies (x). The following events occur:

1. Of helper females on a focal (home) colony, a proportion Y_f will be emitted to a recipient colony as drifters (that is, Y_f is the mean value of the drifting trait in the home colony). Of these, a proportion M_r will on average be accepted (since M_r is the average level of willingness to accept drifters among members of the recipient colony that receives drifters from the focal colony).

Any workers that attempt to drift but are rejected by the neighbour colony return to their home colony and work there.

- Of helper females on a donor colony to the home colony, a proportion Y_d will be emitted to the home colony as drifters. Of these, a proportion M_f will be accepted.

To include social heterosis, we let help have the effect u on an unrelated partner colony and the effect a on the home colony. When $u > a$, workers can make a bigger contribution on an unrelated partner colony (having unrelated genotypes in the workforce amplifies productivity). When $u < a$, workers can make a bigger contribution on their own home colony (having unrelated genotypes in the workforce impairs productivity).

Brood developing on the home colony receive h units of help:

$$h = [a(1 - M_f Y_f) + u M_f Y_d]x \tag{13}$$

In this section, we assume linear returns to cooperation, such that $K = h$, where K is a brood member's probability of successfully developing and h is the help received during its development (see Supplementary Information for a table of notations).

In case A below, we show that the result of ref. ⁵ can be recovered in the improbable scenario where workers that accept incoming drifters are unable to stop themselves from drifting in turn. In case B, we show that cooperative drifting does not invade when this constraint is removed.

Case A: constrained pleiotropy between emission and acceptance of drifters. Here, we force m and y into pleiotropy, so that an increase in the trait value of one is accompanied by an increase in the trait value of the other. We assume $M_f = \beta Y_f$, $M_r = \beta Y_r$, and $\bar{m} = \beta \bar{y}$, where β is a constant ($\beta > 0$). Substituted into equation (13):

$$h = [a(1 - \beta Y_f Y_f) + u \beta Y_f Y_d]x \tag{14}$$

In the Supplementary Information, we provide the background to the relative fitness functions. The relative fitness of both male and female brood is:

$$W_\sigma = W_\varphi = \frac{x}{K} (\beta u Y_f Y_d + a(1 - \beta Y_f Y_f)) \tag{15}$$

In the absence of kinship between colonies (as assumed by ref. ⁵), selection favours a small increase in drifting y when it increases the fitness of the home brood (sisters and brothers on the home colony), which occurs when:

$$c_\varphi \frac{\partial W_\varphi}{\partial Y_f} p_{\text{sis}} + c_\sigma \frac{\partial W_\sigma}{\partial Y_f} p_{\text{bro}} > 0 \tag{16}$$

where c_φ is the class reproductive value for females, which for haplodiploids is $\frac{2}{3}$, and c_σ is the class reproductive value for males, which for haplodiploids is $\frac{1}{3}$. The terms p_{sis} and p_{bro} are haplodiploid consanguinities to sisters and brothers, respectively, which are $\frac{3}{8}$ and $\frac{1}{4}$. Evaluating with both Y_f and Y_r at the population average drifting level \bar{y} , the effect on the fitness of female and male brood in the home nest due to a small increase in drifting by workers from the home nest is:

$$\left. \frac{\partial W_\varphi}{\partial Y_f} \right|_{Y_f=\bar{y}} = \left. \frac{\partial W_\sigma}{\partial Y_f} \right|_{Y_f=\bar{y}} = (u - a) \frac{x\beta\bar{y}}{K} \tag{17}$$

Substituting the relevant consanguinities (p_{sis} and p_{bro}) and equation (17) into the condition for selection on drifting (equation (16)) gives:

$$(u - a) \frac{x\beta\bar{y}}{3K} > 0 \tag{18}$$

Since $\frac{x\beta\bar{y}}{3K}$ has a positive real value, drifting is favoured in this scenario if:

$$u > a \tag{19}$$

Accordingly, cooperative drifting can invade under the hypothesis proposed by ref. ⁵ when increasing the emission of drifters ($y > 0$) to non-relatives is the unavoidable price of increasing the acceptance of drifters ($m > 0$) from non-relatives. Due to a constraint forcing pleiotropy between the traits (β), it is worth paying the price of losing home workers when the incoming non-relatives increase the colony's productivity more than home workers (through social heterosis, $u > a$).

Case B: absence of pleiotropy between emission and acceptance of drifters. We now remove the assumption of pleiotropy, so that m and y are treated independently. Again, assuming that drifters would be unrelated to brood they care for on partner nests, as in ref. ⁵, the condition for selection to favour a small increase in drifting is the same as equation (16). The relative fitness of male and female brood is now:

$$W_\sigma = W_\varphi = \frac{x}{K} (M_f u Y_d + a(1 - M_f Y_f)) \tag{20}$$

Evaluating again with both Y_f and Y_r at the population average drifting level \bar{y} , and with M_f and M_r at the population average acceptance level \bar{m} , the effect on the fitness of female and male brood in the home nest due to a small increase in drifting by workers on the home nest is:

$$\left. \frac{\partial W_\varphi}{\partial Y_f} \right|_{Y_f=\bar{y}} = \left. \frac{\partial W_\sigma}{\partial Y_f} \right|_{Y_f=\bar{y}} = -a \frac{m\bar{x}}{K} \tag{21}$$

As long as a recipient colony does not reject incoming drifters ($\bar{m} > 0$), a small increase in drifting by workers leads to a loss in relative fitness for their sibling brood. Again substituting the relevant consanguinities (p_{sis} and p_{bro}) and equation (21) into the condition for selection on drifting (equation (16)) gives:

$$a \frac{m\bar{x}}{3K} < 0 \tag{22}$$

Since $\frac{m\bar{x}}{3K}$ has a positive real value, drifting is favoured if:

$$a < 0 \tag{23}$$

(that is, if home workers are actively deleterious to the home colony's productivity). Thus, if workers provide any benefit to their home colony ($a > 0$), they should not be emitted as drifters to unrelated colonies.

Selection favours an increase in acceptance (m) of any incoming drifters from other colonies when:

$$c_\varphi \frac{\partial W_\varphi}{\partial M_f} p_{\text{sis}} + c_\sigma \frac{\partial W_\sigma}{\partial M_f} p_{\text{bro}} > 0 \tag{24}$$

A small increase by workers in acceptance to the home nest of incoming drifters increases the relative fitness of their brother and sister brood on the home nest:

$$\left. \frac{\partial W_\varphi}{\partial M_f} \right|_{M_f=\bar{m}} = \left. \frac{\partial W_\sigma}{\partial M_f} \right|_{M_f=\bar{m}} = u \frac{y\bar{y}}{K} \tag{25}$$

Substituting equation (25) into equation (24) gives the condition for an increase in acceptance (m):

$$\bar{y} \frac{u\bar{x}}{3K} > 0 \tag{26}$$

Assuming incoming drifters would contribute positively to productivity ($u > 0$), $\frac{u\bar{x}}{3K}$ has a positive real value, and acceptance is favoured if:

$$\bar{y} > 0 \tag{27}$$

Selection favours acceptance ($m > 0$) whenever drifters are being emitted by other colonies ($\bar{y} > 0$). While selection favours the minimization of drifting to unrelated colonies in all circumstances in which workers are helpful (equation (27)), it favours maximum acceptance m of foreign cooperative drifters (that is, free riding) (equation (23)). Accordingly, willingness to accept drifters (higher m) is not an honest signal of willingness to emit drifters (higher y). In the absence of complex and implausible social insect colony versions of the stabilizing mechanisms known to sustain indirect reciprocity (including between-colony monitoring of whole-colony reputation, uncheatable physical greenbeards or punishment of free-riding colonies by cooperative colonies), the proposed effect of indirect reciprocity⁵ cannot drive drifting. We plot the direction of selection in Fig. 1e.

Finally, we note that, in principle, social heterosis ($u > a$) between related nests could drive drifting by kin selection, rather than by indirect reciprocity: a large benefit that could be provided to distant kin would compensate for their more distant relatedness, and so satisfy Hamilton's rule. However, this would require acutely strong social heterosis at levels unknown in any social insect: high synergies between genotypes would need to ensure that halving relatedness to recipients would more than double the benefit a worker can provide.

Diminishing returns. We consider a situation in which colonies differ in their worker-to-brood ratio ψ , in the range $0 < \psi \leq 1$. For simplicity, we consider neighbour colonies with equal absolute brood numbers. Variation in the worker-to-brood ratio can arise for any reason. For instance, *P. canadensis* colonies show considerable variation in worker number for a given brood size (as shown in the horizontal range of the scatter points of Fig. 3), often through chance worker mortality (estimated at 7% per day¹²). We ask when workers on a home colony with ψ should invest in a partner colony that has a proportion $g < 1$ of the worker-to-brood ratio of the home colony. The diminishing returns hypothesis predicts that workers should not invest in other colonies, despite ψ variation, when the returns from cooperation are linear: the increment in the payoffs of altruism is the same regardless of ψ . However, when there are diminishing returns to cooperation, a worker on a home colony with high ψ may experience weaker indirect fitness returns on its home colony than on a neighbour colony.

We assume that ψ variation between colonies occurs unpredictably from the perspective of an individual. Accordingly, wasps are blind to their colony type when deciding whether to pursue altruistic sterility (worker phenotype) or reproduction. Let the sex ratio (z) and proportion of females that are sterile altruists (x) be common to all nests. We assume that the partner colony will emit no drifters and ask what the equilibrium level of drifting will be from the home colony to the partner colony. The average level of drifting y from the home colony to the partner colony is Y . Both y and Y are evaluated at population average value \bar{y} .

We consider the extent of drifting that should evolve between two types of nest (1 and 2), which differ in their worker-to-brood ratio (ψ). Using the framework of Davies et al.²² (described in the Supplementary Information), we let the absolute fitnesses of the focal female brood in the home type 1 colony (subscript 1) and the partner type 2 colony (subscript 2), respectively, be:

$$w_{\varnothing,1} = (1-x)K_1 \quad (28)$$

$$w_{\varnothing,2} = (1-x)K_2 \quad (29)$$

The rate of producing reproductives on a home colony of type s is $K_s = 1 - (1-h_s)^T$ (see Supplementary Information for details). The total help received by a brood on each colony type is:

$$h_1 = x(1-z)\psi(1-Y) \quad (30)$$

$$h_2 = x(1-z)\psi(g+Y) \quad (31)$$

The population average levels of help on colonies of each type in the population are:

$$\bar{h}_1 = x(1-z)\psi(1-\bar{y}) \quad (32)$$

$$\bar{h}_2 = x(1-z)\psi(g+\bar{y}) \quad (33)$$

The population average relative fitnesses on each nest type, for both male and female brood, assuming that colony types 1 and 2 are equally frequent in the population, are (Supplementary Information):

$$W_{\sigma,1} = W_{\varnothing,1} = \frac{2K_1}{K_1+K_2} \quad (34)$$

$$W_{\sigma,2} = W_{\varnothing,2} = \frac{2K_2}{K_1+K_2} \quad (35)$$

where \bar{K}_s is the population average value of K_s for colonies of type s .

As above (equation (16)), let c_{\varnothing} be the class reproductive value for females ($\frac{2}{3}$). Let c_{σ} be the class reproductive value for males ($\frac{1}{3}$). Selection favours an increase in cooperative drifting from nest type 1 to nest type 2 (y) when the net effect on all potentially affected parties (sisters, brothers, partner-colony female brood and partner-colony male brood) leads to an expected increase in the success of a mutant allele for drifting (we assume no effect on self-fitness, since drifters are already committed to being behaviourally sterile helpers):

$$c_{\varnothing} \left(\frac{\partial W_{\varnothing,1}}{\partial Y} p_{\text{sis}} + \frac{\partial W_{\varnothing,2}}{\partial Y} d_{\varnothing} p_{\text{sis}} \right) + c_{\sigma} \left(\frac{\partial W_{\sigma,1}}{\partial Y} p_{\text{bro}} + \frac{\partial W_{\sigma,2}}{\partial Y} d_{\sigma} p_{\text{bro}} \right) > 0 \quad (36)$$

with traits evaluated at their population average values ($y = Y = \bar{y}$), and where d_{\varnothing} is the devaluation in consanguinity to female brood on the partner colony (relative to sisters on the home colony) and d_{σ} is the devaluation in consanguinity to male brood on the partner colony (relative to brothers on the home colony). Although we are focusing on scenarios in which a worker's consanguinity to a brood is lower on the partner colony than on the home colony, and hence 'devalued' ($d_{\varnothing}, d_{\sigma} < 1$), there can also be scenarios in which a worker is more consanguineous with brood on the partner colony, which may only apply to one sex. For instance, nephews are more consanguineous to a female than brothers are ($p_{\text{nephew}} = \frac{3}{8}$ but $p_{\text{bro}} = \frac{1}{4}$). If the partner-colony male brood are nephews, d_{σ} would be greater than 1. If the partner-colony female brood are nieces, d_{\varnothing} remains below 1 (since $p_{\text{niece}} = \frac{2}{16}$ but $p_{\text{sis}} = \frac{3}{8}$).

A small increase in drifting (Y) by workers from the home type 1 nest leads to a loss of relative fitness for the workers' sibling brood on their home nest:

$$\frac{\partial W_{\varnothing,1}}{\partial Y} \Big|_{Y=\bar{y}} = \frac{\partial W_{\sigma,1}}{\partial Y} \Big|_{Y=\bar{y}} = -\frac{2\psi T x(1-z)(1-\psi x(1-z)(1-\bar{y}))^{T-1}}{K_1+K_2} \quad (37)$$

A small increase in drifting (Y) by workers from the home type 1 nest leads to an increase in relative fitness for the related brood on the partner type 2 nest to which they drift:

$$\frac{\partial W_{\varnothing,2}}{\partial Y} \Big|_{Y=\bar{y}} = \frac{\partial W_{\sigma,2}}{\partial Y} \Big|_{Y=\bar{y}} = \frac{2\psi T x(1-z)(1-\psi x(1-z)(g+\bar{y}))^{T-1}}{K_1+K_2} \quad (38)$$

Substituting equations (37) and (38) into equation (36) and dividing both sides by $\frac{\psi T x(1-z)}{\delta(K_1+K_2)}$ gives the condition for selection to favour a small increase in drifting:

$$\frac{(1-\psi x(1-z)(g+\bar{y}))^{T-1}}{(1-\psi x(1-z)(1-\bar{y}))^{T-1}} > \frac{4}{(3d_{\varnothing}+d_{\sigma})} \quad (39)$$

which is equation (2).

To find candidate ESSs for drifting (y^*), we set the left and right sides of equation (39) equal to one another and rearrange for y . The ESS level of drifting is given by:

$$y^* = \frac{1}{1 + T^{-1} \sqrt{\frac{3}{4}d_{\varnothing} + \frac{1}{4}d_{\sigma}}} \left[1 - T^{-1} \sqrt{\frac{3}{4}d_{\varnothing} + \frac{1}{4}d_{\sigma}} \right] g - \frac{1 - T^{-1} \sqrt{\frac{3}{4}d_{\varnothing} + \frac{1}{4}d_{\sigma}}}{\psi x(1-z)} \quad (40)$$

which is equation (3) (where, for simplicity, we assume that consanguinity to male and female brood is devalued by the same proportion; $d_{\varnothing} = d_{\sigma} = d$). Higher levels of relatedness (higher d_{\varnothing} and d_{σ}) and stronger diminishing returns (higher T) select for higher levels of drifting y^* at equilibrium (Extended Data Fig. 1). We plot the equilibria at illustrative values in Fig. 2f. In Extended Data Fig. 2, we plot the candidate equilibria at different values of sex ratio z and female helping x .

***P. canadensis* payoffs.** To obtain empirical measures of productivity in *P. canadensis*, we tracked a cohort of developing brood on 91 free-living post-emergence colonies over a 56-d period (from 14 June to 8 August 2016). Colonies were clustered in six aggregations on the north coast of Panama (15.2 ± 7.7 colonies per aggregation (mean \pm s.d.)). Five aggregations were in clearings between lowland tropical forest and the Panama Canal (former US Army Base Fort Sherman, San Lorenzo National Park, Colón Province) and one aggregation was in a mangrove swamp (Galeta Point, Colón Province). We used an observational approach to quantify productivity. Key parameters of brood rearing can be quantified effectively in unmanipulated colonies, including natural rates of stochastic failure, predation, parasitism, queen turnover, workforce fluctuations and male production. Associations between brood transition rates and workforce size are correlational, so we view our results as plausible ranges within which causal effects can exist.

To measure the impact of workers at different developmental states, we split brood into stereotypical categories in a sequence. Each category, and its notation, is listed in Extended Data Fig. 3. We examined each brood cell at 5-d intervals, using a ladder to access colonies and a flashlight to illuminate each cell. A small number of observations were made the following morning due to issues with field site access. The current classification of the brood cell was dictated to a second observer, who recorded it on a hexagonal grid of the nest. Accordingly, brood classification was done by a single observer blind to the previous state of the cell.

We censused adult group size by recording the total adult numbers at night (20:00–23:00; six or seven censuses per colony across the monitoring period). We used a red light (which wasps are unable to see) to avoid disturbance. Nests that were difficult to observe were counted multiple times on a given night and averages were taken. A small number of males emerged and stayed on the nests. Adult males observed in daylight surveys during the brood counts on 5-d intervals were used to interpolate male numbers through time; otherwise, males were assumed to be absent. Female numbers (henceforth, group sizes) during each 5-d interval were defined as the interpolated total number of adults minus the interpolated male number. We assumed that any changes in group size between night censuses occurred gradually without sudden jumps, to avoid imposing artificial step changes in the model. We therefore estimated group size during each interval as the mean (across 5 d) of the fitted group sizes generated by a cubic spline interpolation through the night censuses. For 5-d brood observation intervals in which the first night census occurred within 1 d of the beginning of the interval, we allowed a limited extrapolation of 1 d in order to approximate the mean group size over the 5-d interval. All intervals that would require any other extrapolation of group size were excluded from the analysis below in which group size was used as a predictor. To obtain estimates of the payoff rates, we fit a Markov model (shown in Fig. 3) to the brood development data, asking how colony productivity changed in association with different worker and brood numbers.

Statistical methods. The (relatedness-weighted) marginal effect of a worker on the development time of a larva is an incomplete measure of indirect fitness payoff. This is because *Polistes* nests experience considerable individual brood death¹², which returns the cell to the start of the developmental process (once the queen has redeposited an egg). A worker's major contribution might be to prevent inefficiency by minimizing the rate of brood death. To accommodate both effects, we therefore analysed brood development as a Markov model, where the target parameter for inference is the expected time for a brood cell to produce a new adult (time to absorption of the Markov model³⁷). The cell may cycle through repeat deaths before finally producing an adult. We treated the Markov transition matrix as a function of the predictor variables worker number, worker-to-brood ratio, brood cell emptiness, and interactions. To estimate the transition matrix, we used Bayesian mixed models with colony as a random effect. Using the resulting estimate of the time to absorption for *P. canadensis* colonies of different worker and brood sizes, we then obtained the predicted colony productivity rate, defined as the expected number of new adults produced per unit time. Because we were interested in the marginal effect of each additional worker on brood production, we took as an estimate of Hamiltonian benefit B the partial derivative of this rate with respect to the worker number, evaluated at the point in demographic parameter space characterizing any focal colony.

In principle, workers absconding due to a colony entering the declining phase and productivity falling might generate the appearance of a decline in worker numbers causing a decline in productivity, while in reality the decline in productivity caused the decline in worker number. To mitigate this as far as is

possible, we included brood cell emptiness as a measure of colony state and allowed for its interactions.

Because brood states are categorical and longitudinal data were collected in discrete time steps, we give a multinomial likelihood for the number of transitions observed from each developmental state i into each state j , which provides information on the probability π_{ij} of a brood transition from developmental state i to state j . We model the number of transitions $\phi_{i \rightarrow j}$ (defined over a 5-d interval) where the self-transition $i \rightarrow i$ provides the reference category ($\phi_{i \rightarrow i} = 1$) and all other transitions are described by log-linear functions of the predictors x_p in the vector \mathbf{x} (that is, $\ln[\phi_{i \rightarrow (j \neq i)}] = \alpha_{ij} + \beta_{ij,p} \cdot x_p$):

$$\pi_{ij} = \frac{\phi_{i \rightarrow j}}{1 + \sum_{j \neq i} \phi_{i \rightarrow j}} = \frac{e^{(\alpha_{ij} + \beta_{ij,p} \cdot x_p)}}{1 + \sum_{j \neq i} e^{(\alpha_{ij} + \beta_{ij,p} \cdot x_p)}} \quad (41)$$

The predictors include random effects for colony ID (defined below in equation (43)). The $\phi_{i \rightarrow j}$ transitions are the entries in the following matrix:

$$\Phi = \begin{bmatrix} 1 & \phi_{1 \rightarrow 2} & \phi_{1 \rightarrow 3} & \phi_{1 \rightarrow 4} & \phi_{1 \rightarrow 5} & - & - & - & \phi_{1 \rightarrow 9} \\ - & 1 & \phi_{2 \rightarrow 3} & \phi_{2 \rightarrow 4} & \phi_{2 \rightarrow 5} & \phi_{2 \rightarrow 6} & - & - & \phi_{2 \rightarrow 9} \\ - & - & 1 & \phi_{3 \rightarrow 4} & \phi_{3 \rightarrow 5} & \phi_{3 \rightarrow 6} & - & - & \phi_{3 \rightarrow 9} \\ - & - & - & 1 & \phi_{4 \rightarrow 5} & \phi_{4 \rightarrow 6} & - & - & \phi_{4 \rightarrow 9} \\ - & - & - & - & 1 & \phi_{5 \rightarrow 6} & \phi_{5 \rightarrow 7} & - & \phi_{5 \rightarrow 9} \\ - & - & - & - & - & 1 & \phi_{6 \rightarrow 7} & - & \phi_{6 \rightarrow 9} \\ - & - & - & - & - & - & 1 & \phi_{7 \rightarrow 8} & \phi_{7 \rightarrow 9} \\ - & - & - & - & - & - & - & 1 & - \\ - & - & - & - & - & - & - & - & 1 \end{bmatrix} \quad (42)$$

State 1 is the egg, states 2–6 are larval states, state 7 is the pupa, state 8 is the adult (successful pupation) and state 9 is death during development. The final two rows (adulthood (8) and death (9)) are absorbing.

The model is solved in discrete time because brood were observed at intervals. Accordingly, brood are free to transition from one state to a state further downstream than the next step in the sequence; they have passed through the transitional states during the 5-day window. Some transitions are not biologically possible during a 5-day window (such as 1 to 6 or 3 to 7), so are not permitted in the Markov transition matrix (represented by a dash). We present two models with increasing complexity. Model 1 focuses on the baseline transition rates (that is, intercepts and random effects only) for the complete dataset of observed transitions between live-brood-containing cells, which allows us to estimate the baseline productivity rate of *P. canadensis* colonies (Fig. 2c).

Next, we present an indicator-variable-selection model (model 2) to identify the marginal change in productivity associated with each additional worker at different points in the parameter space typifying *P. canadensis* colonies. We exclude all colony observations lying outside of the night census range and a small number of colony observations with fewer than ten (state 2–6) larvae. These steps focus the payoff model on a core dataset of 471 colony observations in 85 colonies, comprising 123,116 observed brood transitions (from the total dataset of 168,811 observed brood transitions on 1,027 colony observations in 91 colonies). The model parameters were estimated using a Bayesian approach computed using Markov chain Monte Carlo (MCMC) simulation in JAGS (using rJAGS⁵⁸ in R). We confirmed MCMC convergence using potential scale reduction factors (\hat{R}) for five chains ($\hat{R} < 1.1$ for the large majority of parameters). After an adaptation phase of 5,000 iterations and a burn-in of 1,000 iterations, we sampled parameter slope values for 15,000 iterations with a thinning of 4. For the posterior predictive model (described below), to generate the inference about payoff rates using simulated input values for colony worker and brood numbers, we continued running MCMC sampling for a further 10,000 iterations with a thinning of 10.

Within and between effects. We used a within-between formulation⁵⁹, which follows the de-meaning procedure suggested by Mundlak⁶⁰. We split each time-varying predictor (that is, predictors whose values could differ between different colony observations within the same colony) into a between-colony component and a within-colony component. The between-colony component was the mean value \bar{x}_C of the relevant predictor for the focal colony C (across its colony observations). The within-colony component was the deviation from this mean in any one colony observation.

We denote the slope dealing with a between-colony component with the subscript b (for example, $\beta_{ij,workers|b}$) and the slope dealing with a within-colony component with the subscript w (for example, $\beta_{ij,workers|w}$). We also permit $x_{workers}$ to interact with the two other predictors: $\beta_{ij,empty,workers}$ denotes an interaction between x_{empty} and $x_{workers}$ and $\beta_{ij,ratio,workers}$ denotes an interaction between x_{ratio} and $x_{workers}$. We allow these interactions at both the between-colony and within-colony levels. The random intercept for the transition $i \rightarrow j$ for colony C is $\beta_{ij,C}$. Thus, the fully saturated model for the transition from state i to state j (subject to pruning of variables during the indicator-variable-selection process detailed in the next section) is:

$$\begin{aligned} \ln[\phi_{i \rightarrow (j \neq i)}] = & \alpha_{ij} + \beta_{ij,C} + \beta_{ij,empty|w}(x_{empty} - \bar{x}_{empty,C}) \\ & + \beta_{ij,empty|b}\bar{x}_{empty,C} + \beta_{ij,workers|w}(x_{workers} - \bar{x}_{workers,C}) \\ & + \beta_{ij,workers|b}\bar{x}_{workers,C} + \beta_{ij,ratio|w}(x_{ratio} - \bar{x}_{ratio,C}) \\ & + \beta_{ij,ratio|b}\bar{x}_{ratio,C} + \beta_{ij,empty,workers|w}((x_{empty} - \bar{x}_{empty,C})(x_{workers} - \bar{x}_{workers,C})) \\ & + \beta_{ij,empty,workers|b}(\bar{x}_{empty,C}\bar{x}_{workers,C}) \\ & + \beta_{ij,ratio,workers|w}((x_{ratio} - \bar{x}_{ratio,C})(x_{workers} - \bar{x}_{workers,C})) \\ & + \beta_{ij,ratio,workers|b}(\bar{x}_{ratio,C}\bar{x}_{workers,C}) \end{aligned} \quad (43)$$

The random effect of colony ID is:

$$\beta_{ij,C} \sim N\left(0, \frac{1}{\tau_{ij}}\right) \quad (44)$$

where the precision τ is drawn from a gamma distribution:

$$\tau_{ij} \sim \text{Gamma}(0.001, 0.001) \quad (45)$$

Model averaging. In an indicator-variable-selection procedure, we construct the saturated model (equation (43)) and introduce a series of binary indicator variables that switch on or switch off each predictor throughout MCMC⁶¹. The duration of MCMC time with each coefficient switched on is proportional to the coefficient's marginal likelihood of contribution. For each predictor p 's regression slope $\beta_{ij,p}$ (within the vector $\beta_{ij,p}$ in equation (41)), we can therefore annex a binary coefficient $g_{ij,p}$ that switches between 0 and 1, and then track the mean of the posterior distribution for $g_{ij,p}$. The closer this mean is to 1, the more frequently the corresponding regression slope $\beta_{ij,p}$ is retained in the model.

For the random effects, we used independent Bernoulli priors for $g_{ij,p}$ to allow the MCMC sampler to turn the specific random effect on or off directly. For other predictors, we used a product of the independent priors for $g_{ij,p}$ and $\beta_{ij,p}$ ⁶²:

$$f(\beta_{ij,p}|g_{ij,p}) = g_{ij,p}f(\beta_{ij,p}|g_{ij,p} = 1) + (1 - g_{ij,p})f(\beta_{ij,p}|g_{ij,p} = 0) \quad (46)$$

$\beta_{ij,p}$ is indistinguishable from zero when $g_{ij,p}$ switches the predictor off. We model the indicators as Bernoulli random variables with a 50:50 prior ($g_{ij,p} \sim \text{Bernoulli}(0.5)$), representing our starting point of indifference between either including or dropping the slope $\beta_{ij,p}$. For interactions, we set a prior of Bernoulli(0.2). To accommodate the within-between formulation, we apply the indicator variables to the whole fixed effect predictor: the predictor cannot be turned on for one component (for example, the between-colony component) and switched off for the other component (for example, the within-colony component).

The effects on the transition probabilities whose 95% credible intervals do not overlap zero are shown in Extended Data Fig. 3.

Model checking: residual deviance. The residual deviances D_i for each multinomial model corresponding to the seven initial states i are given by summing over the residual deviance contributions for each colony observation k , where E_{kij} is the observed number of transitions from state i to state j for colony observation k and \hat{E}_{kij} is the corresponding expected number of transitions from the fitted model⁶³:

$$D_i = \sum_k 2 \left(\sum_j E_{kij} \log \left(\frac{E_{kij}}{\hat{E}_{kij}} \right) \right) \quad (47)$$

The deviance explained is presented in Extended Data Fig. 5.

In Extended Data Figs. 6–8, we plot residual deviance contributions for each colony observation against predictors. There is no evidence of structure in the plots: most colony observations fit relatively well, but some colony observations exhibit very high deviance. In Supplementary Table 3, we provide the mean residual deviances for each of the seven starting state models (residual deviances divided by 471 n_i , where n_i is the number of states into which a cell can transition from starting state i , and 471 is the number of colony observations). The models for starting states 1 and 6 contain observations with high residual deviances. Inspection of potential outlier colony observations (colony observations with unusually high residual deviance contributions) suggests that these high residual deviances may be partly reflective of episodes of unusually high mortality on colonies, with substantial death of large larvae (starting state 6) and substantial cannibalism of eggs (starting state 1). Future models focusing on predictors of severe mortality episodes and other on-colony dynamics may be required to reduce residual deviance in these scenarios.

Expected time to adulthood. To calculate the worker effect on the transition matrix, controlling for oviposition rate, we submit all estimated transition-to-death probabilities ($\phi_{i \rightarrow j} \sqrt{i}$) to the prediction matrices as transitions to new eggs. This isolates the potential effect of workers (as opposed to the egg layer) to obtain per-cell efficiency without the confounding effect of variation between queens in the rate at which replacement eggs are laid following the death of larvae.

The expected time to absorption (eMFPT in Fig. 3a), in which a brood cell transitions from an egg to adulthood (\hat{m}_{1-8}) via intermediate states, can be obtained using the linear algebra for a discrete Markov process via the fundamental matrix method of Kemeny and Snell⁵⁷. Following Grinstead and Snell⁶⁴, we obtain the fundamental matrix \mathbf{N} by inverting the matrix $\mathbf{I}_t - \mathbf{Q}$, where \mathbf{I}_t is the identity matrix for the transient states and \mathbf{Q} is a square matrix of transition probabilities between each transient brood state with length equal to the number of transient states (that is, all states apart from adulthood and death). Accordingly:

$$\mathbf{N} = (\mathbf{I}_t - \mathbf{Q})^{-1} \quad (48)$$

We solve for \mathbf{N} for each of the recorded iterations in MCMC separately, in order to sample the posterior predictive distribution using simulated colonies across the parameter space of worker and brood number, with the number of empty cells set to the population average value. The element n_{ij} in \mathbf{N} is the frequency with which the brood cell is expected to visit state j given a current state i . The vector \mathbf{t} of times to absorption (using the values at any one recorded iteration of MCMC) is then:

$$\mathbf{t} = \mathbf{N}\mathbf{1} \quad (49)$$

where $\mathbf{1}$ is a column vector of 1s. The i th element of \mathbf{t} is the duration (in step numbers) from state i to successful production of a new adult.

To obtain the $\mathbf{I}_t - \mathbf{Q}$ matrix within MCMC, we employ the between-colony effects, which provide a measure of the quantities of interest (workforce size and worker-to-brood ratio) that abstracts away extraneous between-colony variation (in the random intercepts terms $\beta_{j,c}$) and within-colony variation that is likely to be confounded by colony ageing (in the within-colony effects). We then derive the predicted whole-colony rate of brood production by converting the per-cell time to absorption to a productivity rate per unit time and scaling this rate by the brood cohort size on each colony (using as a measure of cohort size the number of prepupal brood, that is states 1–6).

In the main text, we plot the eMFPT in Fig. 3a and the estimated whole-colony productivity in Fig. 3c.

Indirect fitness payoff calculation. Extracting predictions at the population-mean level of cell emptiness (Extended Data Fig. 9), we plot the posterior predictive distribution for whole-colony productivity using 1,000 simulated points (in Mathematica). That is, to extract the shape of the posterior predictive distribution, we fit the smooth interpolation to 1,000 closely packed samples monitored in MCMC.

The diminishing returns hypothesis predicts that a worker can maximize inclusive fitness by shifting from a home colony 1 with worker-to-brood ratio ψ_1 and brood number t_1 (where she is related to the brood by r_1) to a partner colony 2 with worker-to-brood ratio ψ_2 and brood number t_2 (where she is related to the brood by r_2):

$$r_2 \left. \frac{\partial w(\psi, t)}{\partial \psi} \right|_{\psi = \psi_2, t = t_2} > r_1 \left. \frac{\partial w(\psi, t)}{\partial \psi} \right|_{\psi = \psi_1, t = t_1} \quad (50)$$

where she has a larger effect on the production rate w of offspring equivalents by changing ψ_2 on 2 than she would have by changing ψ_1 on 1. The end result of the Markov model is a model of these partial derivatives (estimates of the payoffs driving the benefit term in Hamilton's rule) for colonies of different worker numbers and brood numbers (Fig. 3e).

Individual-based simulation. To simulate the spatial invasion of cooperative drifting under nonlinear returns to cooperation, we consider the evolution of a decision rule in a haplodiploid population on a square lattice in a spatially explicit individual-based simulation. Nodes represent colony sites. Each colony has a single monogamous reproductive pair of a diploid queen and haploid drone. We simulate a death–birth updating process: with each time step, 10% of nodes are selected, and their resident colonies die. A new colony at each updating node is then founded by a new queen drawn from within the dispersal range for females, with probabilities proportional to the payoffs of each of the surrounding eight colonies. We consider two discrete conditions for the female dispersal range: (1) female philopatry (local dispersal within the Moore neighbourhood (the surrounding eight nodes)); and (2) global female dispersal (dispersal from any distance on the lattice with equal probability).

Drones are selected from anywhere in the lattice³⁷, with probabilities proportional to the queen's fitness at each colony on the lattice. Colonies produce female workers as they age: every three time steps, they gain a unit of workers, until reaching a maximum number of units (set to ten in the main text results; Fig. 4). Queen fitness w (production of reproductively destined brood) is determined by:

$$w = 1 - (1 - h)^T \quad (51)$$

where h is the total worker effort received by the colony rescaled between 0 (no workers) and 1 (maximum number of workers possible). We consider two discrete conditions for diminishing returns: no diminishing returns ($T=1$) and moderate diminishing returns ($T=3$).

Each worker has a finite amount of help to distribute. We consider selection on a single locus y , which is the probability of drifting by workers on colonies above a threshold worker number (five 'units' in Fig. 4). Drifter help is distributed at random within each worker's local helping window. We consider two discrete conditions for the size of the helping window: (1) workers help indiscriminately within the local Moore neighbourhood; and (2) workers restrict the helping window only to colonies with immediate genealogical ties to the home colony (defined as the parent colony from which the mother queen came or a daughter colony on which a sister is queen).

Figure 4 summarizes the role of local cooperation.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The transitions data for *P. canadensis* are available in the Supplementary Information.

Code availability

The statistical code and individual-based simulation code are available in the Supplementary Information.

Received: 19 June 2020; Accepted: 10 December 2020;

Published online: 15 February 2021

References

- Gardner, A., West, S. A. & Wild, G. The genetical theory of kin selection. *J. Evol. Biol.* **24**, 1020–1043 (2011).
- Gardner, A. The purpose of adaptation. *Interface Focus* **7**, 20170005 (2017).
- Grafen, A. Optimization of inclusive fitness. *J. Theor. Biol.* **238**, 541–563 (2006).
- Bourke, A. F. G. The validity and value of inclusive fitness theory. *Proc. R. Soc. B Biol. Sci.* **278**, 3313–3320 (2011).
- Nonacs, P. Go high or go low? Adaptive evolution of high and low relatedness societies in social Hymenoptera. *Front. Ecol. Evol.* **5**, 87 (2017).
- Helanterä, H., Strassmann, J. E., Carrillo, J. & Queller, D. C. Uniclonal ants: where do they come from, what are they and where are they going? *Trends Ecol. Evol.* **24**, 341–349 (2009).
- Hamilton, W. D. The genetical evolution of social behaviour. II. *J. Theor. Biol.* **7**, 17–52 (1964).
- Queller, D. C. & Strassmann, J. E. The many selves of social insects. *Science* **296**, 311–313 (2002).
- Tsuchida, K. & Itô, Y. Internidal drifting and dominance behaviour in *Polistes jadvigae* Dalla Torre workers (Hymenoptera: Vespidae). *J. Ethol.* **5**, 83–85 (1987).
- Page, R., Post, D. & Metcalf, R. Satellite nests, early males, and plasticity of reproductive behavior in a paper wasp. *Am. Nat.* **134**, 731–748 (1989).
- Pickering, J. *Sex Ratio, Social Behaviour, and Ecology in Polistes (Hymenoptera, Vespidae), Pachysomoides (Hymenoptera, Ichneumonidae) and Plasmidium (Protozoa, Haemosporida)*. PhD Thesis, Harvard Univ. (1980).
- Sumner, S., Lucas, E., Barker, J. & Isaac, N. Radio-tagging technology reveals extreme nest-drifting behavior in a eusocial insect. *Curr. Biol.* **17**, 140–145 (2007).
- Grinstead, L. & Field, J. Market forces influence helping behaviour in cooperatively breeding paper wasps. *Nat. Commun.* **8**, 13750 (2017).
- Beekman, M. & Oldroyd, B. P. When workers disunite: intraspecific parasitism by eusocial bees. *Annu. Rev. Entomol.* **53**, 19–37 (2008).
- Gadagkar, R. Origin and evolution of eusociality: a perspective from studying primitively eusocial wasps. *J. Genet.* **69**, 113–125 (1990).
- Nonacs, P. & Kapheim, K. M. Social heterosis and the maintenance of genetic diversity. *J. Evol. Biol.* **20**, 2253–2265 (2007).
- Weigel, R. M. The distribution of altruism among kin: a mathematical model. *Am. Nat.* **118**, 191–201 (1981).
- Grinstead, L. & Field, J. Predictors of nest growth: diminishing returns for subordinates in the paper wasp *Polistes dominula*. *Behav. Ecol. Sociobiol.* **72**, 88 (2018).
- Michener, C. Reproductive efficiency in relation to colony size in hymenopterous societies. *Insectes Soc.* **11**, 317–341 (1964).
- Bulmer, M. G. Risk avoidance and nesting strategies. *J. Theor. Biol.* **106**, 529–535 (1984).
- Hopper, K., Rosenheim, J., Prout, T. & Oppenheim, S. Within-generation bet hedging: a seductive explanation? *Oikos* **101**, 219–222 (2003).
- Davies, N. G., Ross, L. & Gardner, A. The ecology of sex explains patterns of helping in arthropod societies. *Ecol. Lett.* **19**, 862–872 (2016).
- Sumner, S., Kelstrup, H. & Fanelli, D. Reproductive constraints, direct fitness and indirect fitness benefits explain helping behaviour in the primitively eusocial wasp, *Polistes canadensis*. *Proc. R. Soc. B Biol. Sci.* **277**, 1721–1728 (2010).

24. Ferreira, P. G. et al. Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* **14**, R20 (2013).
25. Southon, R. J. et al. High indirect fitness benefits for helpers across the nesting cycle in the tropical paper wasp *Polistes canadensis*. *Mol. Ecol.* **28**, 3271–3284 (2019).
26. Wenzel, J. W. & Pickering, J. Cooperative foraging, productivity, and the central limit theorem. *Proc. Natl Acad. Sci. USA* **88**, 36–38 (1991).
27. Lengronne, T. et al. Little effect of seasonal constraints on population genetic structure in eusocial paper wasps. *Ecol. Evol.* **2**, 2610–2619 (2012).
28. Davies, N. G. & Gardner, A. Monogamy promotes altruistic sterility in insect societies. *R. Soc. Open Sci.* **5**, 172190 (2018).
29. Schulman, S. R. & Rubenstein, D. I. Kinship, need, and the distribution of altruism. *Am. Nat.* **121**, 776–788 (1983).
30. Kramer, B. H., Scharf, I. & Foitzik, S. The role of per-capita productivity in the evolution of small colony sizes in ants. *Behav. Ecol. Sociobiol.* **68**, 41–53 (2014).
31. Tindo, M., Kenne, M. & Dejean, A. Advantages of multiple foundress colonies in *Belonogaster juncea juncea* L.: greater survival and increased productivity. *Ecol. Entomol.* **33**, 293–297 (2008).
32. Naug, D. & Wenzel, J. Constraints on foraging success due to resource ecology limit colony productivity in social insects. *Behav. Ecol. Sociobiol.* **60**, 62–68 (2006).
33. Clouse, R. Some effects of group size on the output of beginning nests of *Mischocyttarus mexicanus* (Hymenoptera: Vespidae). *Florida Entomol.* **84**, 418–425 (2001).
34. Smith, A. R., Wcislo, W. T. & O'Donnel, S. Survival and productivity benefits to social nesting in the sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Behav. Ecol. Sociobiol.* **61**, 1111–1120 (2007).
35. Bouwma, A. M., Nordheim, E. V. & Jeanne, R. L. Per-capita productivity in a social wasp: no evidence for a negative effect of colony size. *Insectes Soc.* **53**, 412–419 (2006).
36. Shreeves, G. & Field, J. Group size and direct fitness in social queues. *Am. Nat.* **159**, 81–95 (2002).
37. Parsons, P. J., Grinsted, L. & Field, J. Partner choice correlates with fine scale kin structuring in the paper wasp *Polistes dominula*. *PLoS ONE* **14**, e0221701 (2019).
38. Helanterä, H. Supercolonies of ants. In *Encyclopedia of Social Insects* (ed. Starr, C. K.) (Springer, 2020); https://doi.org/10.1007/978-3-319-90306-4_121-1
39. Holzer, B., Keller, L. & Chapuisat, M. Genetic clusters and sex-biased gene flow in a unicolonial *Formica* ant. *BMC Evol. Biol.* **9**, 69 (2009).
40. Boomsma, J. J., Huszár, D. B. & Pedersen, J. S. The evolution of multiqueen breeding in eusocial lineages with permanent physically differentiated castes. *Anim. Behav.* **92**, 241–252 (2014).
41. Keller, L. Social life: the paradox of multiple-queen colonies. *Trends Ecol. Evol.* **10**, 355–360 (1995).
42. Mora-Kepfer, F. Context-dependent acceptance of non-nestmates in a primitively eusocial insect. *Behav. Ecol. Sociobiol.* **68**, 363–371 (2013).
43. Couvillon, M. J. et al. En garde: rapid shifts in honeybee, *Apis mellifera*, guarding behaviour are triggered by onslaught of conspecific intruders. *Anim. Behav.* **76**, 1653–1658 (2008).
44. McNamara, J. Optimal patch use in a stochastic environment. *Theor. Popul. Biol.* **21**, 269–288 (1982).
45. Foster, K. R. Diminishing returns in social evolution: the not-so-tragic commons. *J. Evol. Biol.* **17**, 1058–1072 (2004).
46. Sibly, R. M. & Curnow, R. N. Selfishness and altruism can coexist when help is subject to diminishing returns. *Heredity* **107**, 167–173 (2011).
47. Cant, M. A. A tale of two theories: parent–offspring conflict and reproductive skew. *Anim. Behav.* **71**, 255–263 (2006).
48. Dugatkin, L. A. *The Altruism Equation: Seven Scientists Search for the Origins of Goodness* (Princeton Univ. Press, 2006).
49. Altmann, S. Altruistic behaviour: the fallacy of kin deployment. *Anim. Behav.* **27**, 958–962 (1979).
50. Dawkins, R. Twelve misunderstandings of kin selection. *Z. Tierpsychol.* **51**, 184–200 (1979).
51. Rice, S. H. *The Expected Value of the Ratio of Correlated Random Variables* (Texas Tech Univ., 2015); <https://go.nature.com/361lo0B>
52. Zivot, E. *Introduction to Computational Finance and Financial Econometrics with R* (2016); <https://go.nature.com/2Kz6IhL>
53. Shpak, M. Evolution of variance in offspring number: the effects of population size and migration. *Theory Biosci.* **124**, 65–85 (2005).
54. Starrfelt, J. & Kokko, H. Bet-hedging—a triple trade-off between means, variances and correlations. *Biol. Rev.* **87**, 742–755 (2012).
55. Taylor, P. & Frank, S. How to make a kin selection model. *J. Theor. Biol.* **180**, 27–37 (1996).
56. Rice, S. H. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. *BMC Evol. Biol.* **8**, 262 (2008).
57. Kemeny, J. G. & Snell, J. *Finite Markov Chains* (Springer, 1983).
58. Plummer, M. rjags: Bayesian graphical models using MCMC. Version 4-9 (2019); <http://mcmc-jags.sourceforge.net>
59. Bell, A., Fairbrother, M. & Jones, K. Fixed and random effects models: making an informed choice. *Qual. Quant.* **53**, 1051–1074 (2019).
60. Mundlak, Y. On the pooling of time series and cross section data. *Econometrica* **46**, 69–85 (1978).
61. Hooten, M. B. & Hobbs, N. T. A guide to Bayesian model selection for ecologists. *Ecol. Monogr.* **85**, 3–28 (2015).
62. Ntzoufras, I. Gibbs variable selection using BUGS. *J. Stat. Softw.* **7**, 1–19 (2002).
63. Dias, S., Sutton, A. J., Ades, A. E. & Welton, N. J. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med. Decis. Mak.* **33**, 607–617 (2013).
64. Grinstead, C. M. & Snell, J. L. *Introduction to Probability* 2nd edn (American Mathematical Society, 2000).

Acknowledgements

P.K. was funded by a Smithsonian Tropical Research Institute (STRI) Short-Term Fellowship (hosted by W. Wcislo), the National Geographic Society (GEF-NE 145-15) and a European Research Council Grant to A.N.R. (award number 682253). S.S. was funded by the Natural Environment Research Council (NE/M012913/2). A.D.H. was funded by the Natural Environment Research Council (NE/L011921/1). A.N.R. was funded by a European Research Council Consolidator Grant (award number 682253). Simulations were run on the University of Bristol's high-performance computing facility, BlueCrystal. B. Wharam and A. C. Chadwick assisted with data entry. We are very grateful for the support of the Panamanian authorities for this project; data were collected in accordance with Panamanian law under a Ministerio del Ambiente research permit (SE/A-46-16) and Ministerio de Economía y Finanzas authorization for field site access (024-2016). We thank the generous support of STRI at Galeta Field Station.

Author contributions

P.K., S.S. and A.N.R. planned the field data collection. P.K. and P.B. collected the field data. P.K. and A.D.H. conducted the modelling. P.K., N.J.W. and A.N.R. conducted the statistical analysis. P.K., S.S. and A.N.R. interpreted the results. P.K. drafted the manuscript and all authors contributed to its development.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-020-01382-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41559-020-01382-z>.

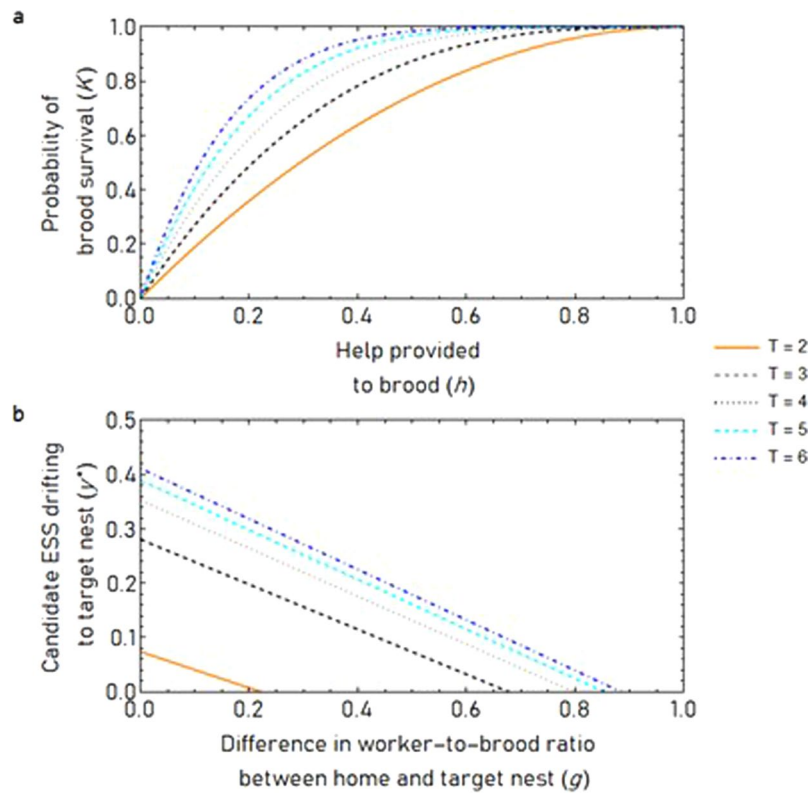
Correspondence and requests for materials should be addressed to P.K.

Peer review information *Nature Ecology & Evolution* thanks Michael Cant, Petri Rautiala and Christina Riehl for their contribution to the peer review of this work. Peer reviewer reports are available.

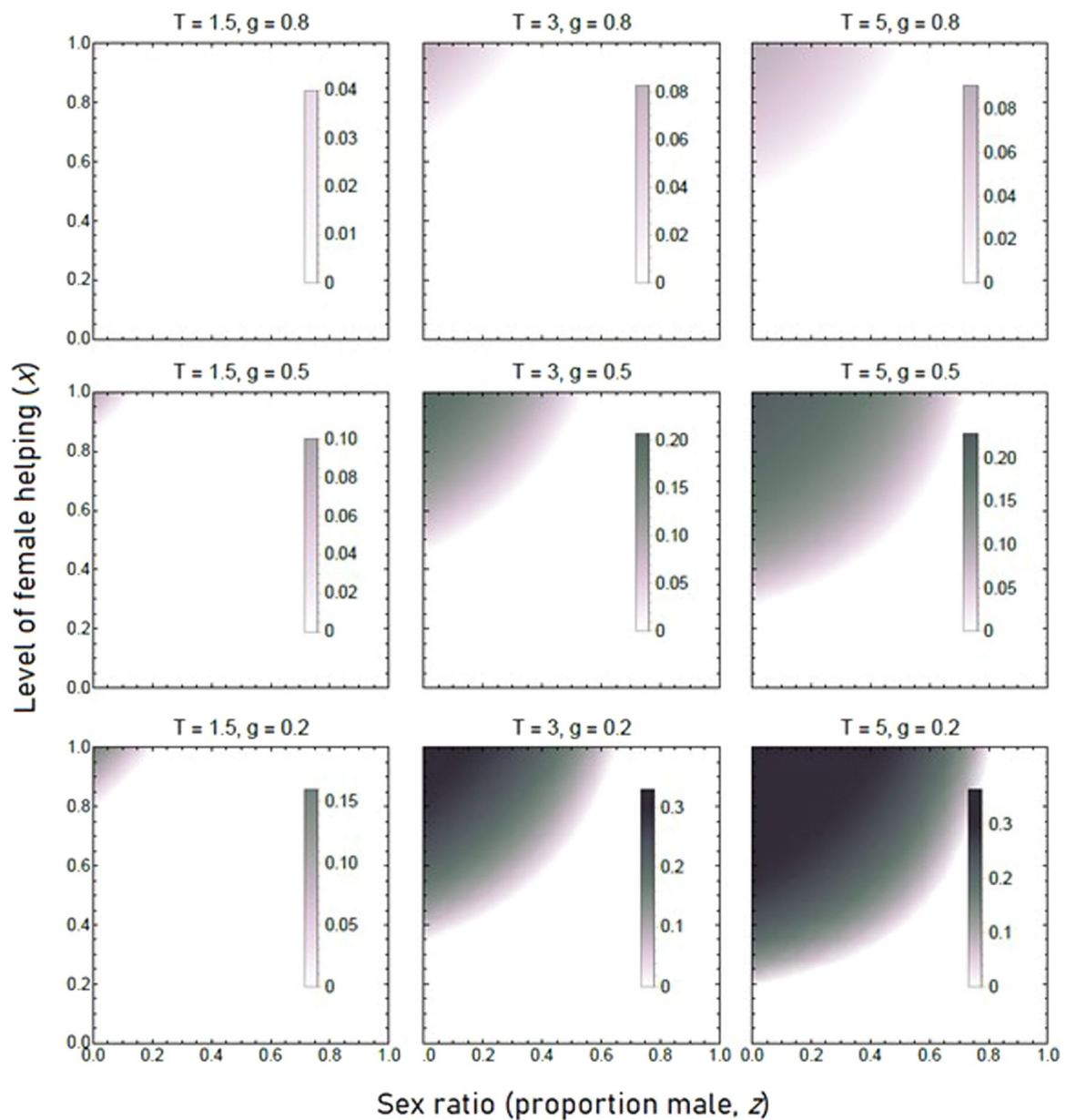
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

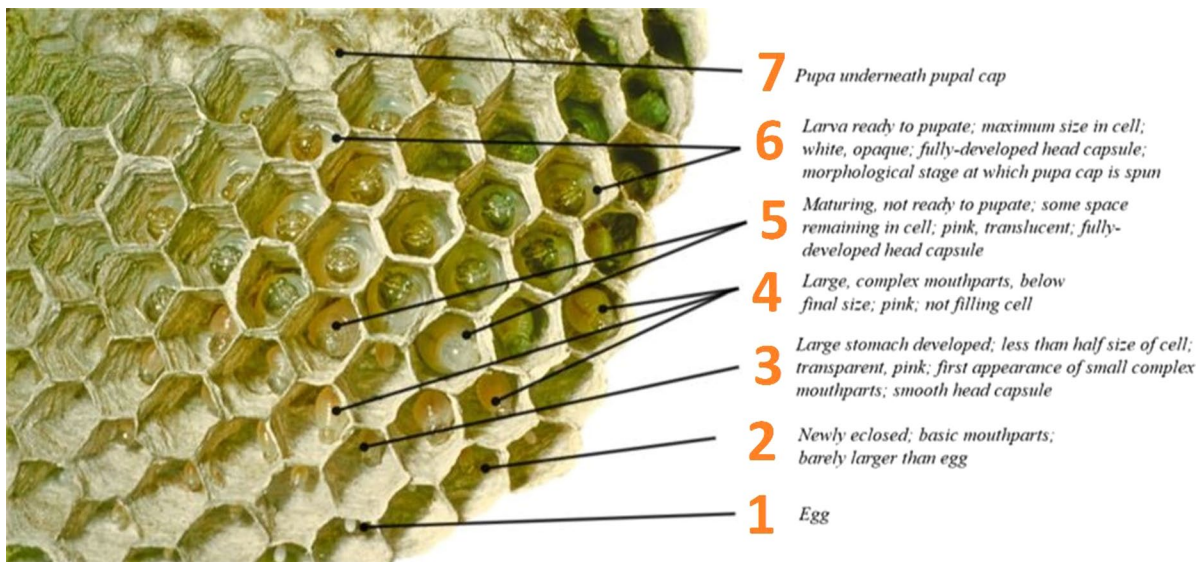
© The Author(s), under exclusive licence to Springer Nature Limited 2021



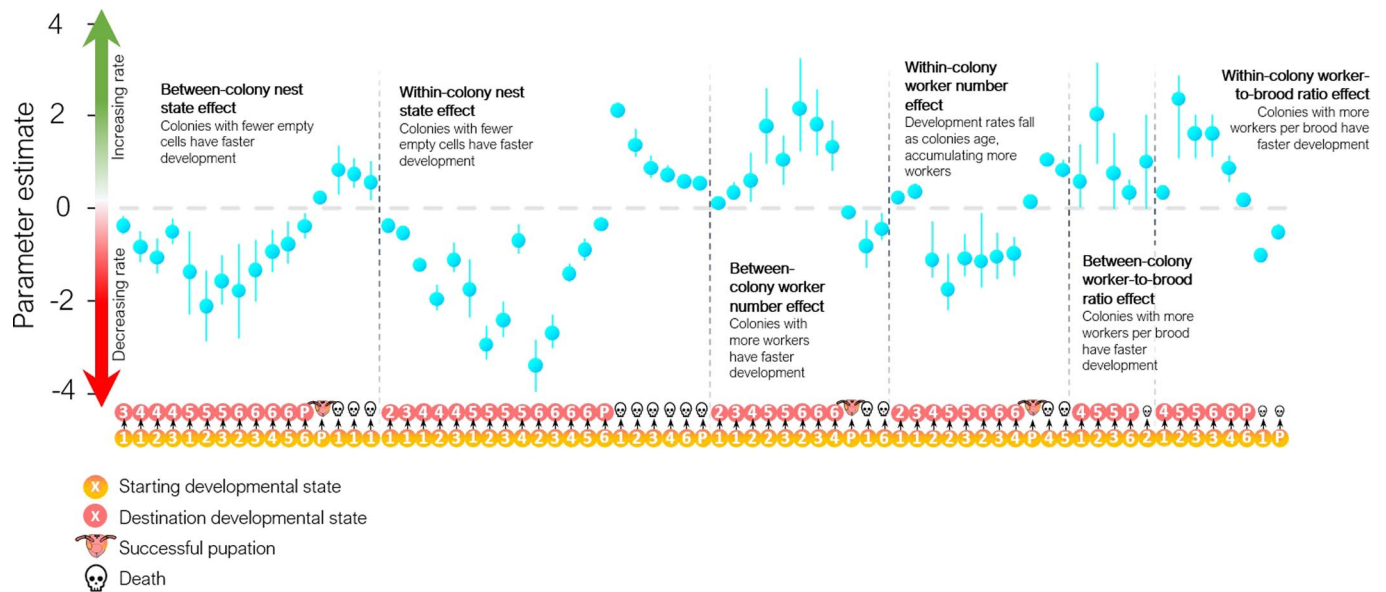
Extended Data Fig. 1 | Diminishing returns can allow the invasion of cooperative drifting. **a**, The parameter T captures the level of diminishing returns in the model. **b**, Candidate equilibrium levels of drifting predicted by Eq. 3 of the main text, for illustrative values: $d=0.5$, $x=0.75$, $\phi=1$, $z=0.25$.



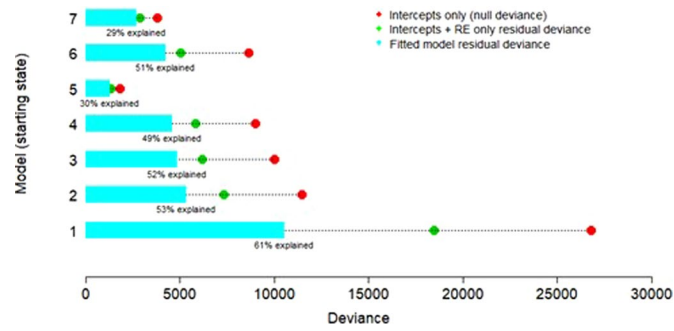
Extended Data Fig. 2 | Candidate equilibrium drifting levels from a colony with $\phi_1=1$. We assume a monomorphic sex ratio \bar{z} and level of female helping \bar{x} common to all colonies. Diminishing returns become more likely drivers of cooperative drifting as the level of helping rises (top left corner). $\bar{m} = 1$, $\phi_1=1$, $d_0=d_3=0.5$.



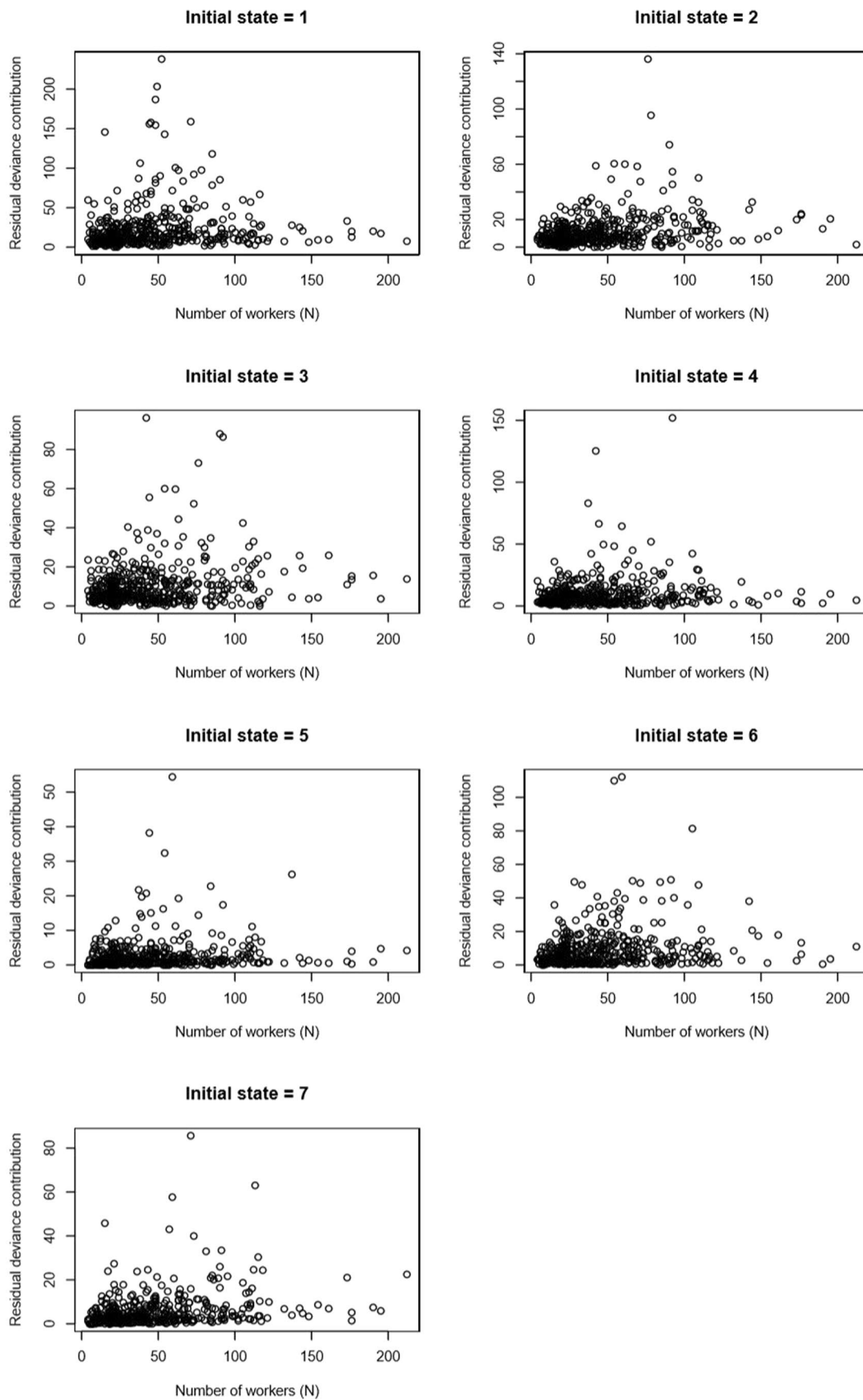
Extended Data Fig. 3 | Repeat observations ($n = 257,867$) were made of all brood cells on 91 post-emergence colonies over a 56-day period, recording eight developmental states, empty cells, and death due to parasitism (sarcophagid flies and ichneumonid wasps). Brood pass through a clear developmental sequence of changing size, colour, and mouthpart complexity, which allows categorisation of larvae into stereotyped morphological categories on their transition from egg to pupa. Image adapted with permission from J. Pickering (<http://go.nature.com/2Lq1QLY>).



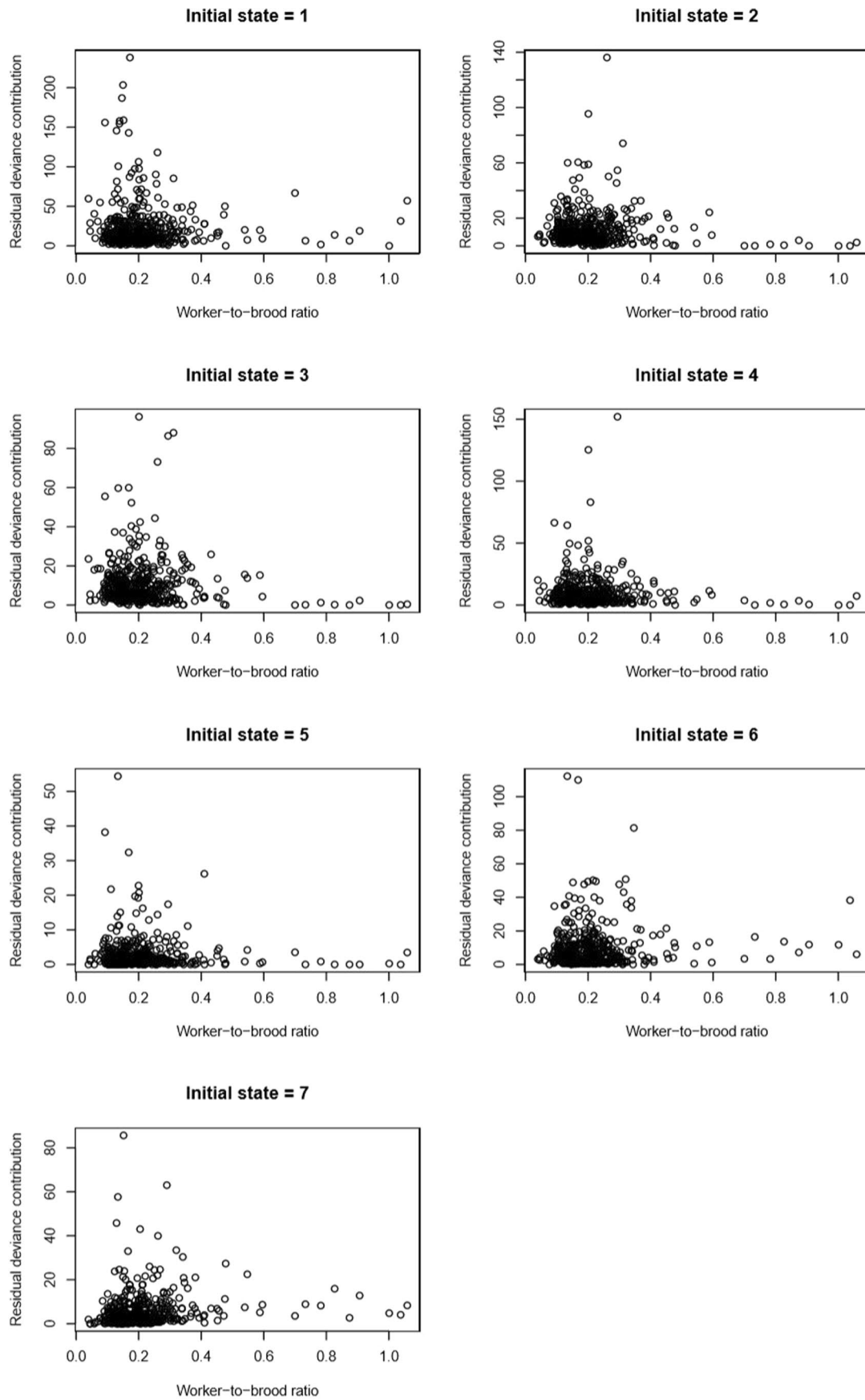
Extended Data Fig. 4 | Main effects on the probability of transitions in the Markov model of wasp development. Transitions corresponding to steps through the Markov model states (represented in Fig. 2b of the main text) are shown along the horizontal axis. For instance, the first transition on the left is 1→3, and the effect is negative (below the dashed line): accordingly, the model fits a negative association (at a between-colony level) between nest emptiness ('nest state effect') and the probability that an individual egg (state 1) transitions to larval state 3 in a 5-day interval. Only main effects whose 95% credible intervals do not overlap zero are shown. In the main text, we focus on whether increased worker number and/or worker-to-brood ratio is associated with higher productivity. When considering between-colony variation ('between-colony worker number effect' in the figure), worker number increases brood development pace and reduces brood death. Higher worker-to-brood ratios ('between-colony worker-to-brood ratio effect' in the figure) increase brood developmental pace.



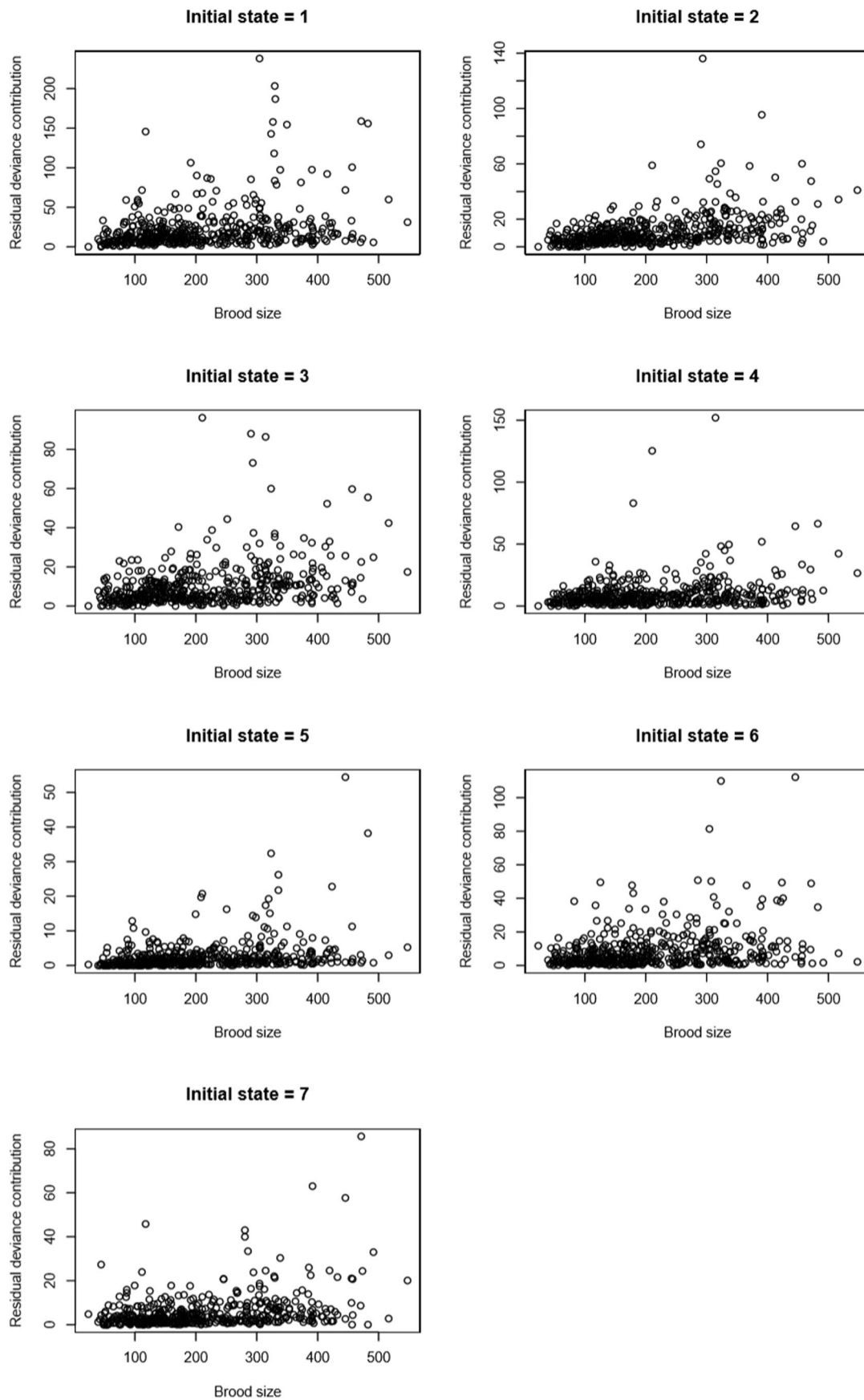
Extended Data Fig. 5 | Posterior mean of the total residual deviance. Deviance explained is the posterior mean of the total residual deviance of the fitted model as a percentage of the null posterior mean residual deviance (for an intercepts-only model). There are seven models, corresponding to the seven starting states.



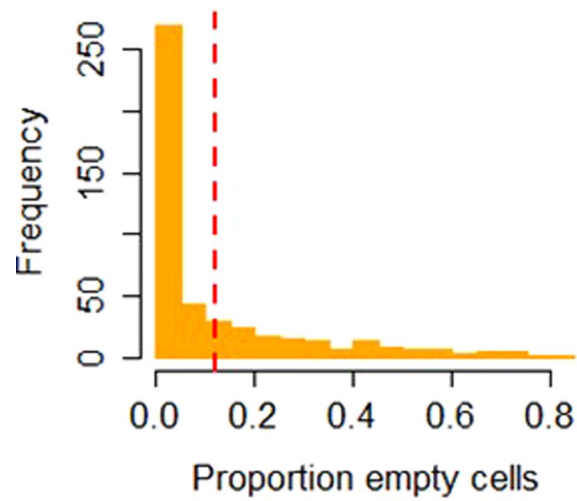
Extended Data Fig. 6 | Residual deviance against worker number. Residual deviance contributions by each 5-day colony-observation, plotted against the predictor 'worker number' (x_{workers}).



Extended Data Fig. 7 | Residual deviance against worker-to-brood ratio. Residual deviance contributions by each 5-day colony-observation, plotted against the predictor 'worker-to-brood ratio' (x_{ratio}).



Extended Data Fig. 8 | Residual deviance against brood number. Residual deviance contributions by each 5-day colony-observation, plotted against the predictor 'brood number' (x_{brood}).



Extended Data Fig. 9 | Proportion of cells that are empty. Predictions for between-colony effects plotted in Fig. 3 in the main text take the mean proportion emptiness across colony-observations (dashed red line).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Longitudinal observational field data was collected on the paper wasp <i>Polistes canadensis</i> at 5-day intervals. Night-censused worker numbers were obtained at points over the 56-day study period, and interpolated to generate estimated worker number, which was used as a predictor of brood-rearing rates in a Markov model.
Research sample	A total sample of 257,867 observations of nest cells were made on 91 post-emergence colonies of the wasp <i>Polistes canadensis</i> over a 56-day period using 1,027 colony-observations. Of these, 168,811 were transitions involving live brood; the main analysis ('Model 2') focuses on a core dataset of 123,116 transitions involving live brood, by discarding observations for which interpolated worker number is not available and any colony observation with fewer than 10 A-to-L-stage larvae.
Sampling strategy	The maximum possible number of colonies were observed during fieldwork subject to constraints of time and logistics.
Data collection	<p>Extract from Methods:</p> <p>To measure the impact of workers at different developmental stages, we split brood into stereotypical categories in a sequence. Each category, and its notation, is listed in Extended Data Fig. S2. We examined each brood cell at 5-day intervals, using a ladder to access colonies and a flashlight to illuminate each cell. A small number of observations were made on the following morning due to issues with field site access. The current classification of the brood cell was dictated to a second observer, who recorded it on a hexagonal grid of the nest. Accordingly, brood classification was done by a single observer blind to the previous state of the cell.</p> <p>We censused adult group size by recording total adult numbers at night (8 pm–11 pm, 6–7 censuses per colony across the monitoring period). We used a red light (which wasps are unable to see) to avoid disturbance. Nests that were difficult to observe were counted multiply on a given night and averages taken. A small number of males emerged and stayed on nests. Adult males observed in daylight surveys during the brood counts on 5-day intervals were used to interpolate male number through time; otherwise, males were assumed to be absent. Female number (henceforth, 'group size') during each 5-day interval was defined as interpolated total number of adults minus interpolated male number. We assumed that any changes between night-censuses were smooth, to avoid imposing artificial step changes in the model.</p>
Timing and spatial scale	Colony observations were made from 14th June 2016 to 8th August 2016, using six aggregations on the Panamanian north coast near Colon (15.2 ± 7.7 colonies per aggregation, mean \pm S.D.). Five aggregations were in clearings between lowland tropical forest and the Panama Canal (former US Army Base Fort Sherman, San Lorenzo National Park, Colón Province) and one aggregation was in a clearing in a mangrove swamp (Galeta Point, Colón Province).
Data exclusions	Data exclusions were as described above (see 'Research sample'): brood transitions for which no interpolated worker number was available or for which the sample size would be below 10 A-to-L-stage larvae in a colony-observation were excluded.
Reproducibility	n/a, as the data were non-experimental.
Randomization	n/a, as subjects were not allocated to experimental treatments; instead, brood transitions occur in colonies, and colony was included as a covariate in the analysis.
Blinding	Field data collection was conducted with full blinding as to the previous state of each nest cell. An observer dictated the category of each cell to a recorder, who entered the updated brood cell state on a hexagonal grid of the nest.
Did the study involve field work?	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

Field work, collection and transport

Field conditions	Most data collection occurred in July 2016. Average conditions for Colon in July 2016 were 27°C and 85% humidity (https://www.timeanddate.com/weather/panama/colon/historic?month=7&year=2016). July average rainfall in Colon is 209mm (https://weather-and-climate.com/colon-colon-pa-July-averages).
Location	Colony aggregations were in the area administered by the Panamanian Aeronaval Police and the Ministry of the Economy (9.359092,-79.950568) and in the grounds of the Smithsonian Tropical Research Institute's Punta Galeta Field Station (9.402662,-79.864165).
Access & import/export	Research and export permits were obtained with MiAmbiente (Ministerio de Ambiente) with the project represented at the Smithsonian Tropical Research Institute (STRI). Field site access permission was obtained at the Ministerio de Economía y Finanzas de Panamá (MEF) in Colon.

Disturbance

Minimal disturbance was caused by the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

n/a

Wild animals

The study focused on workers and brood of the Neotropical paper wasp *Polistes canadensis*, using wild populations.

Field-collected samples

n/a

Ethics oversight

No ethical guidance was required for this study; research and export permission for research involving Panamanian wildlife was obtained through Ministerio de Ambiente (MiAmbiente).

Note that full information on the approval of the study protocol must also be provided in the manuscript.