doi:10.1111/evo.13160

Inclusive fitness in finite populations—effects of heterogeneity and synergy

Peter Taylor[1,2]

[1]*Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada*

[2]*E-mail: peter.taylor@queensu.ca*

Received August 14, 2015
Accepted December 14, 2016

I review recent results concerning the relationship between the inclusive fitness (IF) effect and standard measures of allele fitness in a finite-population, with attention to the effect of heterogeneity in population structure and nonadditive fitness effects. In both cases, existing theoretical work is somewhat technical and I try to provide a more transparent account. In a heterogeneous population it is known that inclusive fitness will generally fail to incorporate the effects of selection on the distribution of alleles among states unless a reproductive-value weighting is used. But even given that, recent work shows that under certain updating rules, the IF effect can fail to be equivalent to standard measures such as fixation probability. In terms of synergistic fitness effects, I review the result that in the finite population model, the IF effect can be calculated using only "additive" relatedness coefficients so that computational difficulties found in the infinite-population model do not arise. In my own work, there is an interaction here in that my 2012 work on synergy with Maciejewski made an assumption about inclusive fitness that my 2014 work on heterogeneity with Tarnita showed to be wrong. I include (Appendix C) a corrected argument for the 2012 result.

KEY WORDS: **Allele-frequency change, matrix games, population structure, reproductive value.**

Since its inception 50 years ago (Hamilton 1964) the ideas of inclusive fitness (IF) have had a significant impact on the modeling of evolutionary behavior. There are good reasons for this. It has provided a powerful heuristic for understanding how selection acts on behavior and has given us effective methods for calculating the effects of selection. At the same time there has been a vigorous discussion around the significance and generality of the inclusive fitness method in evolutionary studies. See Nowak et al. (2010) and Abbot et al. (2011) and accompanying articles for a noteworthy slice of this discussion.

This discussion seems to me to take place in two different areas, one concerns the extent to which the inclusive-fitness approach has anything significant (useful or insightful) to say about the biological world. The other concerns theoretical questions, for example the assumptions needed for the IF effect to be equivalent to the probability of gene-fixation. The first of these is inherently complex and can be hard to pin down. The second can in principle be settled unambiguously; all that is needed is a precise mathematical framework.

My interest here is the second and I will focus on two complications to the classic theory as laid out by Hamilton (1964). They have been considered for a long time, but have recently—over the past half-dozen years, received considerable discussion. They are both fairly technical and no doubt for that reason, are not as well understood or appreciated as I feel they deserve to be. They concern the impact on inclusive-fitness theory of first, population heterogeneity, and second, synergistic fitness effects. In both cases, the story is interesting; however some careful analysis is required to sort out the different factors at play. My objective in this review article is to attempt to make them more accessible.

For the first few decades inclusive fitness studies were set in what I call the infinite-population model but more recently there has been particular interest in the finite-population model. When I use the terms "infinite" and "finite" I really mean "large" and "small." In fact, as we will see, it is really about the relationship between population size and the rate of genetic renewal (e.g., mutation). The finite model makes sense when the population is small enough that, under the force of genetic drift, a mixed genetic state becomes pure (genetic fixation) more quickly than mutation is able to provide a new allele. Otherwise the infinite model makes more sense. Inclusive fitness has a somewhat different formulation in each model.

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

In the finite model there is a richer set of potential measures of allele fitness, for example, allele fixation, and that is perhaps one of the reasons for considerable recent interest in this model (Rousset and Billiard 2000; Nowak et al. 2004; Ohtsuki and Nowak 2006; Taylor et al. 2007a,b; Ohtsuki et al. 2007; Grafen and Archetti 2008; Tarnita et al. 2009; Taylor 2010; Taylor and Maciejewski 2012; Allen and Tarnita 2014; Tarnita and Taylor 2014). My objective here is to review the relationship of the inclusive fitness effect to these other measures and to see in particular how this is affected by the complications of heterogeneous population structure and synergistic effects.

The Components of Allele-frequency Change

MODEL ASSUMPTIONS

I work with a finite population of N nodes (BOX 1), each occupied by an asexual haploid individual, and any pair of nodes joined by an edge are called neighbors (Lieberman et al. 2005). These edges specify both the fitness determining interactions and the offspring dispersal probabilities and in general these might be different (Ohtsuki et al. 2007). Here, I want to use simple assumptions to obtain a clear understanding of a few recent results, and I assume that interactions occur along edges and offspring from a node disperse with equal probability to all neighbors.

Individuals carry one of two alleles, A or B and I let x_i denote the genotype at node i, $x = 1$ for A and $x = 0$ for B. I use a Moran model with BD updating and fecundity payoffs (Ohtsuki and Nowak 2006). In each time step an A-individual gives each neighbor a fecundity increment b at cost c, whereas B-individuals give nothing with no cost. Following that, one individual is chosen based on relative fecundity to give birth, and its offspring replaces a random neighbor. The offspring carries the parental allele with probability $1–u$ and with probability u is a "mutant," acquiring allele A or B with fixed probabilities q and $1–q$. I assume that mutation is rare and the results discussed are calculated to first order in the mutation rate u.

The state of the population at any time will be a specification of the allele at each node. When population size N is small and mutation is rare, the population will spend almost all of its time in a pure state (allA or allB). Once a mutation creates a mixed state, the population, under the action of drift and selection, will change state quickly (relative to the time between mutations) until it arrives again at a pure state, there to remain awaiting the next mutation to a new allele. Indeed the overall ratio of time spent in mixed to pure states is of order u. This time structure is illustrated in BOX 2.

THE EFFECTS OF SELECTION

In working with this model, our objective is typically to calculate the relative fitness of the alleles A and B. The literature provides a number of different measures of this along with a discussion of their equivalence (BOX 3). The one I will work with here is the effect of selection on the one-step change in the frequency of the allele A. Now in a small population, the one-step allele-frequency change will depend on the state and in a single time step, the state can change significantly. For this reason, we work with an average change over all states. We define $\mathrm{E}(\Delta x)$ to be the total (selection plus mutation) long-term average (over all states) one-step change in allele-frequency (Rousset and Billiard 2000). For each population state s, we calculate the one-step allele-frequency change Δx_s, and then we average these over all states using the equilibrium state frequencies π_s^* as weights:

$$\mathrm{E}(\Delta x) = \sum_s \Delta x_s \pi_s^* = 0. \tag{1}$$

As indicated, $\mathrm{E}(\Delta x)$ is of course zero as the π_s^* population is at equilibrium under the joint forces of selection and mutation.

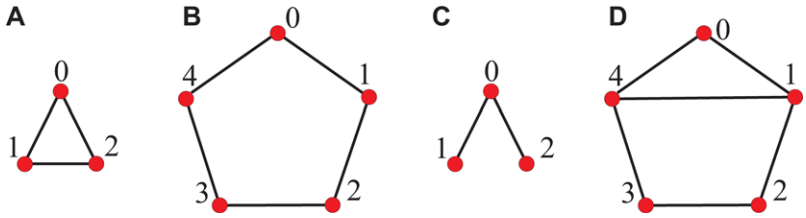
To understand this process, it helps to track through an example and in BOX 4 and BOX 5 I track two components of this process for the 3-star population (BOX 1). This analysis makes it clear that a substantial calculation is needed to obtain this average even for a population of size 3, and this makes Hamilton's concept of inclusive fitness (in particular, his use of relatedness coefficients—Appendix A) all the more remarkable. In this regard, I point out that this problem of the apparent complexity of the calculation of this average does not belong only to the finite population model. Take, for example, a classic situation in an effectively infinite population in which an individual is interacting with a number of brothers and sisters. We presume that there are many such interactions going on in many different "households" and we want to calculate the effect on allele frequency averaged over all such households and for that we will need to know the distribution of possible genotype configurations. That is the same problem and these households are the analogs of the different states.

In terms of the overall average one-step allele-frequency change, what are we to take as fitness measure (i) in BOX 3—"the effect of selection on the one-step change in allele frequency"? Tarnita and Taylor (2014) answer this question by partitioning the sum according to whether the state is mixed or pure:

$$\mathrm{E}(\Delta x) = \sum_{s\mathrm{mixed}} \Delta x_s \pi_s^* + \sum_{s\mathrm{pure}} \Delta x_s \pi_s^* \; = 0. \tag{2}$$

This decomposition of $\mathrm{E}(\Delta x)$ goes back to Nowak et al. (2010 Supporting Information A). I note right away that selection cannot act in a pure state so that the $\sum_{s\mathrm{pure}}$ term provides only the effect of mutation, and secondly, that if mutation is rare it can be ignored in the mixed states as their frequency is of order u (BOX 2). Thus if mutation is rare, both terms are of order u, the $\sum_{s\mathrm{pure}}$ term capturing the effect of mutation and the $\sum_{s\mathrm{mixed}}$ term capturing the effect of selection. Given this, we might expect that

BOX 1: Homogeneous and heterogeneous population structure



Here, the edges signify both interactions and offspring dispersal. For example, in graph (d), node 1 has three interactions in each time step, with nodes 0, 2, and 4, and sends its offspring to these same nodes each with probability 1/3. On the other hand, node 2 has only two interactions and sends offspring to nodes 1 and 3 with equal probability 1/2.

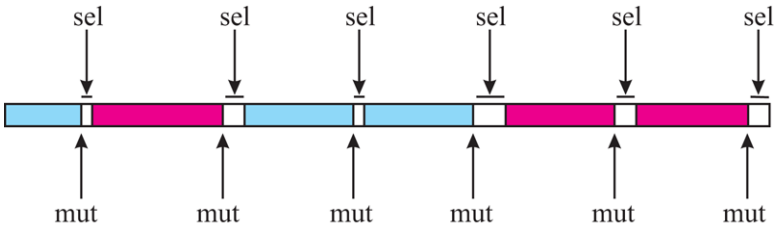
Reproductive classes. An *isomorphism* of the graph is a one-to-one mapping of the node set to itself that preserves the edges. The set of isomorphisms can be used to partition the set of nodes into reproductive classes—two nodes are in the same class if there is an isomorphism of the graph mapping one to the other. A graph is called transitive (Taylor et al. 2007b) if it has only one reproductive class—given any pair of nodes there is an isomorphism of the node set mapping to the other—roughly speaking the structure "looks the same" from every node. Transitive graphs are often called *homogeneous*; otherwise they are called *heterogeneous.*

For example, examples (a) and (b) are transitive and (c) and (d) are heterogeneous. For graphs (c) and (d) reflection in the vertical central axis is an isomorphism giving us two classes in (c): {0} and {1, 2}, and three classes in (d): {0}, {1, 4}, and {2, 3}.

In fitness calculations, we often need to take averages over all nodes. It is generally enough to take a focal node in each class and take the average over these focal nodes, each one weighted by its class size.

Example (c) is known as the 3-star graph, with node 0 called the hub and nodes 1 and 2 called the leaves. I will use this graph as a key example in my analysis.

BOX 2: The population timeline



Population timeline displaying alternating relatively short intervals of mixed states and relatively long intervals when the state is pure, A red (black) or B blue (gray), both types of intervals of variable lengths. The transition from a pure to a mixed state is effected by a mutation when it introduces a new allele. There can also be mutation events in a fixed state that introduce a new copy of the allele already present but we ignore these as there is no change in allele frequency. The length of the fixed-state intervals has order $1/u$ (the mean time between mutations) and the length of the mixed-state intervals has order N (Crow and Kimura 1970—time to fixation. See also Wu et al. 2012). Since $u << 1/N$, the length of the mixed-state intervals is of order u and we can assume that mutation does not occur in a mixed state. An important consequence of this is that any two copies of the same allele in a mixed state can be assumed to be identical by descent.

An important observation, for our purposes, is that since selection can act only in mixed states and the relative time during which these occur is of order u, the effects of selection on allele frequency will also have order u, the same order as the effects of mutation. This is worth noting as it is *not* the case for classical large-population model in which local genetic renewal is generated by means other than mutation (e.g., migration from afar).

BOX 3: Measures of A-fitness

In the finite population model that I work with here, three measures of the "fitness" of A have been of particular interest. These are:

(i) The average one-step selective change in allele frequency at the selective equilibrium.

(ii) The difference in the allele frequency at the neutral and selective equilibria (BOX 5).

(iii) The difference $\rho_A - \rho_B$ in fixation probabilities.

Here, the fixation probability ρ_A of an allele A is the probability that a single copy of A in an otherwise B population will survive till the next fixation (and thus will have identical by descent (IBD) copies at every node). Actually while this definition works fine in a transitive population, in a heterogeneous population the fixation probability of a gene will depend on the starting node and ρ_A and ρ_B are defined as weighted averages of the node-specific fixation probabilities where the weights are the landing probabilities, through mutation, of the allele. Details can be found in Allen and Tarnita (2014) and Tarnita and Taylor (2014).

Under quite general conditions (principally rare mutation, but *not* assuming weak selection), these three measures have been shown to be sign-equivalent, that is they are all simultaneously positive or negative or zero. (Rousset and Billiard 2000; Taylor et al. 2007a; Nowak et al. 2010, Appendix A, Allen and Tarnita 2014; Tarnita and Taylor 2014). A clarification—when working with heterogeneous (nontransitive—BOX 1) structures, we often work with allele RV instead of allele frequency. But measures (i) and (ii) are about allele frequency and the equivalence holds in both transitive and heterogeneous population structures.

Our main interest here is to relate the inclusive fitness effect to these measures. The early work on inclusive fitness typically assumed a large population and measures (ii) and (iii) were not considered; thus attention was focused on allele-frequency change in a single time step. Now Price's (1970) equation will calculate this provided the genetic covariances (which provide the relatedness coefficients) are calculated at the selective equilibrium (Taylor 2009). The action of selection makes this calculation difficult, but it can be done when selection is turned off giving us the point of neutral equilibrium and the covariances we obtain here are called the neutral coefficients of relatedness. These are the coefficients used in the inclusive-fitness formulation (Appendix A) and, as pointed out by Hamilton (1964, p. 4), when selection is weak they should provide a good approximation and hence the IF effect should give a good approximation to measure (i) above. Here i discuss some interesting recent analysis giving us general conditions under which the IF effect is exactly equivalent to measure (i) and hence to measures (ii) and (iii).

the $\sum_{s\text{mixed}}$ term is equivalent to fitness measure (i) in BOX 3 and Tarnita and Taylor (2014) show that under standard general conditions (principally rare mutation, but not assuming weak selection) this is the case, and thus the $\sum_{s\text{mixed}}$ term is equivalent to fitness measures (ii) and (iii) as well.

The decomposition in equation (2) is effectively a partition of the columns of the state transition matrix M (BOX 5). For example, for the 3-star population, the mixed-state sum belongs to columns 1–4 and the pure-state sum belongs to columns 5 and 6.

A WEAK-SELECTION ANALYSIS

I use δ to measure the strength of selection such that b and c are both of order δ. Then, to first order in δ, I write $\Delta x_s = \Delta x_s^0 + \Delta x_s^\delta$ and $\pi_s^* = \pi_s^0 + \pi_s^\delta$ where Δx_s^0 is the average one-step change in allele frequency at state s and π_s^0 is the equilibrium frequency of state s (BOX 5), both in the neutral population (selection not acting). Tarnita and Taylor (2014) use this decomposition to analyze $E(\Delta x)$ into four components:

$$E(\Delta x) = \sum_{s\text{mixed}} \left(\Delta x_s^0 + \Delta x_s^\delta\right)\left(\pi_s^0 + \pi_s^\delta\right) + \sum_{s\text{pure}} \left(\Delta x_s^0 + \Delta x_s^\delta\right)\left(\pi_s^0 + \pi_s^\delta\right) \quad (3)$$

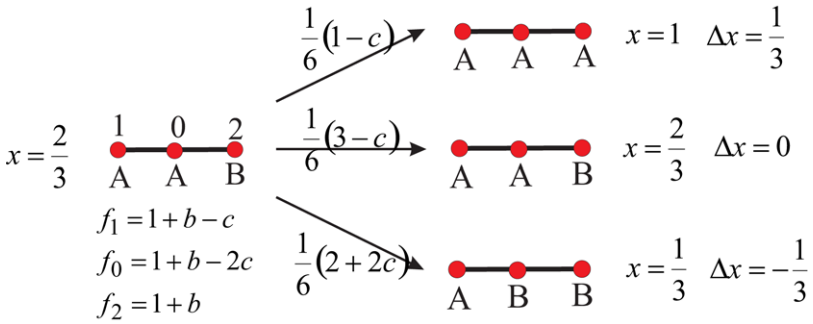
$$= \sum_{s\text{mixed}} \Delta x_s^\delta \pi_s^0 + \sum_{s\text{mixed}} \Delta x_s^0 \pi_s^\delta + \sum_{s\text{pure}} \Delta x_s^\delta \pi_s^0 + \sum_{s\text{pure}} \Delta x_s^0 \pi_s^\delta \quad (4)$$

$$= \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 = 0. \quad (5)$$

If we expand equation (3) we get a total of eight terms, four mixed, and four pure, but equation (4) only has four of these as the two δ–δ terms are of order δ^2 and can be ignored and the two 0–0 terms together provide the one-step change at the neutral equilibrium with no selection and this must be zero. Here, I am assuming that δ is small but significantly greater than the mutation rate u.

We will be working with these four terms so that in equation (5) I have given them abbreviated names Δ_i. Thus Δ_1 and Δ_2 belong to the mixed states and Δ_3 and Δ_4 belong to the pure states. Observe that selection acts on the one-step change in allele frequency in two different ways—first it changes this in each state, and second (as a result of the first), it changes the weighting used in taking the average of these different state effects. That gives us another useful way to partition the four terms—Δ_1 and Δ_3 assess the effects belonging to each state and Δ_2 and Δ_4

BOX 4: The average change in allele frequency from the state AAB



$$F_0 = \frac{f_0}{\sum f_i} = \frac{1+b-2c}{3+3b-3c} = \frac{1}{3}(1-c) + o(\delta)$$

$$\Delta x_{AAB} = \frac{1}{6}(1-c)\left(\frac{1}{3}\right) + \frac{1}{6}(3-c)(0) + \frac{1}{6}(2+2c)\left(-\frac{1}{3}\right) = -\frac{1}{18} - \frac{1}{6}c$$

I take f_i to be the initial fecundity of node i, obtained from the benefit-cost interaction. Note for example that node 0 interacts with both neighbors, and since it carries allele A, in each interaction it gives b at cost c. The probability that node i is selected to give birth is then its relative fecundity F_i calculated above for $i = 0$ to first order in the payoffs b and c taken to be of order δ.

The diagram shows that the state AAB has three possible one-step descendent states and the probabilities of each transition are derived from the fecundities F_i. For example, the F_0-calculation tells us that node 0 wins the birth with probability $\frac{1}{3}(1-c)$ and in that case sends its offspring with probability ½ to each of nodes 2 (giving us the transition to AAA) and 1 (giving us the transition to AAB). Note that there's a second way to reach AAB—if node 1 wins the birth.

For each transition the one-step change Δx in allele frequency is tabulated and the average change Δx_s from the state s = AAB is then calculated taking account of the probability of each transition. This change is calculated from each state s and the overall average allele-frequency change $E(\Delta x)$ is calculated by taking the average (eq. 1) over the equilibrium state distribution (BOX 2).

assess the effects that come from the change due to selection on the equilibrium state frequencies. A significant observation is that Δ_1 and Δ_3 can generally be calculated as they involve the neutral state distribution, while Δ_2 and Δ_4 generally cannot. To get a sense of what we will be doing with these, you can peek ahead at BOX 8.

GENERAL ANALYSIS OF THE Δ_i

Right away, to allay confusion, I observe that in a pure state $\Delta x_s^\delta = 0$ as selection cannot change the fact that the offspring produced will carry the allele belonging to that state. Thus Δ_3 is always 0. However, for comparison's sake I want to include it.

Turning to $\Delta_1 = \sum_{s\text{mixed}} \Delta x_s^\delta \pi_s^0$, this is the effect of selection acting in each state averaged over all mixed states using the neutral state distribution. Those who have worked with inclusive fitness, (abbreviated as IF–Appendix A) will recognize that this is exactly what the IF effect is meant to measure (the neutral state distribution corresponding to the use of the neutral relatedness coefficients) and indeed Tarnita and Taylor (2014) show that Δ_1 is indeed sign-equivalent to the IF effect, though further discussion of this is needed in heterogeneous populations (see below). Now I have already observed that fitness measure (i) in BOX 3 is sign-equivalent to $\Delta_1+\Delta_2$, and thus Δ_2 is what stands in the way of the sign-equivalence of the IF effect and measure (i).

Finally, it is worth having a brief look at $\Delta_4 = \sum_{s\text{pure}} \Delta x_s^0 \pi_s^\delta$. First, consider the neutral pure-state frequencies determined only by mutation; to zeroth order in the mutation rate, they are $\pi^0_{\text{allA}} = q$ and $\pi^0_{\text{allB}} = 1 - q$ (BOX 5), giving us a neutral allele frequency of q. The standard argument for this is found in Taylor (2007a) and Tarnita and Taylor (2014); essentially it follows from the fact that these are the frequencies that give us a zero average one-step change in allele frequency at neutrality (only mutation acting). Indeed, to calculate that, we can ignore the mixed states as they are of order u so that the net change in the pure states will be $uq(1–q) – u(1–q)q$ and this is zero as expected. The first term provides the rate uq at which A is produced in allB and the second provides the rate $u(1–q)$ at which B is produced in allA.

Now let the effect of selection on the equilibrium allele frequency be δK—this of course is measure (ii) in BOX 3 but for my purpose we need to know nothing more about it. Then

BOX 5: The transition matrix and equilibrium state vector

$$M = \frac{1}{18}\begin{array}{c}\begin{matrix}\text{BAB} & \text{AAB} & \text{ABB} & \text{ABA} & \text{BBB} & \text{AAA}\end{matrix}\\ \left[\begin{array}{cccc|cc} 0 & 0 & 0 & 0 & 12qu & 0 \\ 6-4b-8c & 9-3c & 6-2b-4c & 0 & 0 & (1-q)u(6+4b-4c) \\ 0 & 6+6c & 9+3c & 6+8b+4c & 6qu & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-q)u(12-4b+4c) \\ \hline 12+4b+8c & 0 & 3+2b+c & 0 & 18(1-qu) & 0 \\ 0 & 3-3c & 0 & 12-8b-4c & 0 & 18(1-(1-q)u) \end{array}\right]\end{array}\begin{matrix}\text{BAB}\\ \text{AAB}\\ \text{ABB}\\ \text{ABA}\\ \text{BBB}\\ \text{AAA}\end{matrix}$$

M is the one-step transition matrix of the population. Its columns and rows are indexed by the population states and the column for each state is the vector of probabilities of the transition to each state. For example, the AAB column gives us the transition probabilities diagramed in BOX 4. If $\boldsymbol{\pi}$ is a probability vector, a vector with entries ≥ 0 that sum to 1, then entries of the vector $M\boldsymbol{\pi}$ will give us the probability of being in each state one time step later.

An equilibrium state vector $\boldsymbol{\pi}$ is one that is invariant under one-step transition: $M\,\boldsymbol{\pi} = \boldsymbol{\pi}$. Such a vector is an eigenvector for the matrix M, and under mild conditions the population, starting in any state, will converge to such an equilibrium. This is the case for the 3-star population and the equilibrium state vector for this population is:

$$\pi* = \begin{bmatrix} \text{BAB}: & \frac{2}{3}q(1-q)u\left(1+q\left(2b+\frac{10}{3}c\right)\right) \\ \text{AAB}: & \frac{10}{3}q(1-q)u\left(1+q\left(2b+\frac{10}{3}c\right)-\frac{6}{5}b-\frac{7}{3}c\right) \\ \text{ABB}: & \frac{10}{3}q(1-q)u\left(1+q\left(2b+\frac{10}{3}c\right)-\frac{14}{15}b-\frac{13}{15}c\right) \\ \text{ABA}: & \frac{2}{3}q(1-q)u\left(1+q\left(2b+\frac{10}{3}c\right)-\frac{7}{3}b-3c\right) \\ \text{BBB}: & (1-q)+q(1-q)\left(2b+\frac{10}{3}c\right)+\text{O}(u) \\ \text{AAA}: & q-q(1-q)\left(2b+\frac{10}{3}c\right)+\text{O}(u) \end{bmatrix} \quad \pi_0 = \begin{bmatrix} \frac{2}{3}q(1-q)u \\ \frac{10}{3}q(1-q)u \\ \frac{10}{3}q(1-q)u \\ \frac{2}{3}q(1-q)u \\ 1-q+\text{O}(u) \\ q+\text{O}(u) \end{bmatrix}$$

Thus, if we let the population run for a long time, the entries of $\boldsymbol{\pi}*$ will give us the proportion of time the population is in that state. If we turn selection off (set $b = c = 0$) we get the neutral equilibrium $\boldsymbol{\pi}_0$ displayed at the right. Notice that the equilibrium frequency of the first four states (the mixed states) is of order the mutation rate u. This comes from the fact that the only way to escape fixation (the pure states AAA and BBB) is through mutation. See BOX 2 for a fuller discussion of this.

the pure-state frequencies at the selective equilibrium will be $\pi^{\delta}_{\text{allA}} = q + \delta K$ and $\pi^{\delta}_{\text{allB}} = 1 - q - \delta K$ and the resulting average one-step change in allele frequency will be

$$\begin{aligned}\Delta_4 &= \sum_{s\text{pure}} \Delta x_s^0 \pi_s^{\delta} = uq(1-q-\delta K) - u(1-q)(q+\delta K) \\ &= -u\delta K, \end{aligned} \tag{6}$$

and except for the factor u this is the negative of measure (ii) in BOX 3. Since $\Delta_3 = 0$, we can write the equilibrium equation (5) as

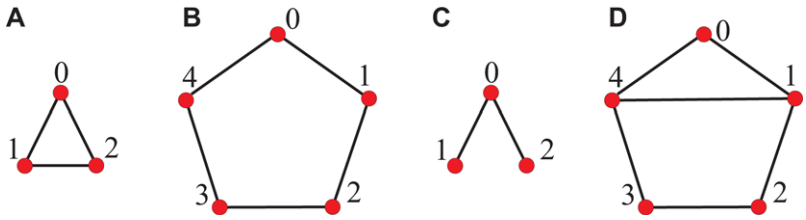
$$\Delta_1 + \Delta_2 = -\Delta_4. \tag{7}$$

Since the left-hand side is equivalent to measure (i), this is a simple argument that the two measures (i) and (ii) of BOX 3 are sign-equivalent.

TRANSITIVE POPULATIONS

It turns out that if the population is transitive (BOX 1), then $\Delta x_s^0 = 0$ for every state s and as a result $\Delta_2 = 0$. It is easy enough to see the reason. Suppose we have a gene on a node that in one time-step produces on average more than one copy of itself. There can only be two explanations for this, one has to do with the behavior of the gene, and the other with the status of its node. If we take away the difference between alleles (that is the superscript 0 of Δx_s^0) and

BOX 6: Reproductive value



In a heterogeneous graph, we assign a reproductive value (RV) v_i to each node i such that nodes in the same reproductive class have the same RV. These RVs are specified by the condition that in the neutral population (selection turned off) the RV gain of every node (through birth) is equal to its RV loss, where a death at node i counts as $-v_i$ and a birth at node i counts as v_j if the offspring colonizes node j (Taylor 1990, 2009; Leturque and Rousset 2002; Maciejewski 2014; Tarnita and Taylor 2014). These conditions determine the node RVs up to a constant.

As an example I provide the RV calculation for graph (d). I assume BD updating so that at neutrality all nodes have the same probability of giving birth, and to have simple numbers, suppose this probability is 1. Take the nodes 0, 1, and 2 as class representatives with RV v_0, v_1, and v_2. The equations are:

RV gain through birth = RV loss through death.

$$\text{node0}: \; v_1 = \left(\frac{1}{3}+\frac{1}{3}\right)v_0. \quad \text{node1}: \; \frac{v_0+v_1+v_2}{3} = \left(\frac{1}{2}+\frac{1}{3}+\frac{1}{2}\right)v_1. \quad \text{node2}: \; \frac{v_1+v_2}{2} = \left(\frac{1}{2}+\frac{1}{3}\right)v_2.$$

These solve to give (up to a constant) $v_0 = v_2 = v_3 = 3/13$ and $v_1 = v_4 = 2/13$, normalized so that $\sum v_i = 1$. In fact for examples such as those above in which the edges at each node represent offspring dispersal routes all with the same node-specific probability, a result of Maciejewski (2014) shows that the reproductive values v_i depend only on the degree of the node. Indeed under DB updating, they are proportional to the degree of the node and under BD updating to the reciprocal of the degree. Note that this makes qualitative sense. Under DB it is good for a node to have high degree as this provides many opportunities for reproduction; conversely under BD it is good for a node to have low degree as this reduces the probability of death. This result would have given us the values v_i for graph (d) without any calculation. Indeed this also gives us the node RV for the 3-star graph (c): the hub has $v_0 = 1/5$ and the leaves have $v_1 = v_2 = 2/5$ (Appendix B).

the difference between nodes (that is the transitivity) both of those explanations disappear. As a result, to first order in δ, measure (i) of BOX 3 is equivalent to Δ_1 and its calculation requires only the neutral state distribution and is therefore feasible. Indeed its sign can be obtained from the IF effect. This result goes back to Rousset and Billiard (2000), and is extended in Taylor et al. (2007a).

POPULATION HETEROGENEITY

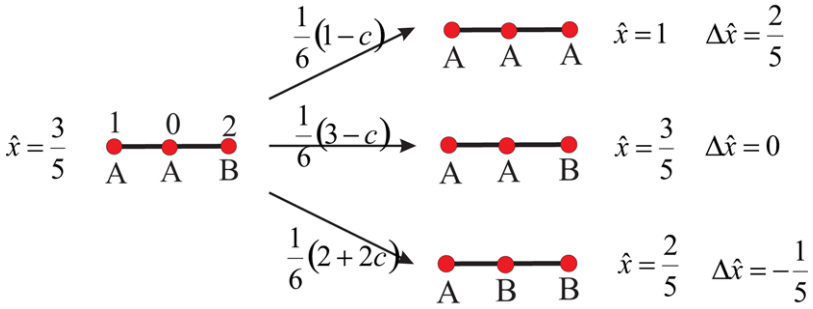
In a heterogeneous population Δx_s^0 and therefore Δ_2 can be nonzero and as a result fitness measure (i) is not in general equivalent to the IF effect and cannot generally be feasibly calculated. This is the case for the 3-star population and, for example, BOX 4 shows that $\Delta x_s^0 = -1/18$ for the state $s = \text{AAB}$. To jump ahead, BOX 8 provides a table of the Δ_i values for the 3-star population.

Hamilton (1972, p. 204) was perhaps the first to notice that the inclusive fitness calculation might not deliver the one-step change in allele frequency when there is population heterogeneity. Using computer simulations he discovered an error in his 1967 article on extreme sex ratios in which the inclusive fitness argument led him to an incorrect value for the ESS sex allocation under haplodiploidy (p. 485). He pinned the error on his failure to weight the fitness effects with reproductive values and we will now look more closely at the connection between these and the Δ_i measures.

REPRODUCTIVE VALUE

It has been understood for a long time that in a heterogeneous population, different classes of individuals can make different future contributions to the gene pool and any analysis of fitness should take account of these differences. Fisher (1927, p. 106) referred to "a consistent system of valuation by which the potential value to the population of future generations, of each age group, may be measured" and in 1930 he formalized the notion of reproductive value (RV) of an individual as its long-term future genetic contribution to the population and argued for its use as a weighting to be applied when comparing the fitness of individuals in different age classes.

BOX 7: The average change in allele reproductive value from the state AAB



$$\Delta\hat{x}_{\mathrm{AAB}} = \frac{1}{6}(1-c)\left(\frac{2}{5}\right) + \frac{1}{6}(3-c)(0) + \frac{1}{6}(2+2c)\left(-\frac{1}{5}\right) = -\frac{2}{15}c$$

The state transition probabilities are the same as BOX 4 but now we calculate changes in allele RV where each copy of A is weighted by the RV of the node it inhabits. If we normalize so that total population RV is 1, then the hub has RV 1/5 and each leaf has RV 2/5.

For each transition the one-step change $\Delta\hat{x}$ in allele RV is tabulated and the average change $\Delta\hat{x}_s$ from the state $s = \mathrm{AAB}$ is then calculated taking account of the transition probabilities. Note that the neutral change $\Delta\hat{x}_s^0$ is now zero. This was not the case without the RV weighting. Indeed in BOX 4 we calculated that $\Delta x_s^0 = -1/18$. It might well have struck you as strange that the quantity Δx_s that is designed to measure the effect of selection, should be nonzero when selection is turned off. That was certainly one of many signals that an adjustment was needed, and that turned out to be reproductive value.

Finally, the overall average change in allele RV $\mathrm{E}(\Delta\hat{x})$ is calculated by taking the average over the equilibrium state distribution.

In the study of evolutionary graphs, each node i is assigned a reproductive value v_i determined up to a multiplicative constant by the recursive condition that at neutrality, the one-step RV-weighted fitness change of each node is zero (BOX 6). In other words, for each individual, RV-weighted fecundity equals RV-weighted mortality where offspring are weighted by the RV of the node they inhabit. We typically normalize the v_i so that total population RV is 1. Using a recursive argument it is not hard to see that these values v_i do indeed provide the long-term neutral genetic contributions of the node. If we use the concept of graph isomorphism (BOX 1) to partition the set of nodes into reproductive classes—two nodes are in the same class if there is an isomorphism of the graph mapping one to the other—then nodes in the same reproductive class have the same RV.

The condition, that at neutrality every node has zero one-step RV-weighted fitness change, leads to a significant mathematical result. Following Tarnita and Taylor (2014) we use a hat to signal the RV-weighting. Then in any state s, define the RV-weighted allele frequency to be $\hat{x}_s = \sum_i v_i x_i$ where recall that $x_i = x_{i,s}$ is the genotype of node i in state s. Then in the neutral population ($b = c = 0$) the one-step RV-weighted allele frequency change from any state s is zero $\Delta\hat{x}_s^0 = 0$ (Taylor 1990; Leturque and Rousset 2002; Taylor 2009). This is illustrated in BOX 7.

Thus, if Fisher (1930) has not already given us a good enough reason to use RV-weights when assessing fitness in heterogeneous populations, here is a mathematical reason—when we use the RV-weighted allele-frequency measure, the component Δ_2 of selective allele-frequency change belonging to selective changes in the state distribution becomes $\hat{\Delta}_2$ and vanishes. If we put hats on all the Δ_i to signal the use of the RV-weighted allele frequency, the equilibrium equation (5) becomes

$$E(\Delta\hat{x}) = \hat{\Delta}_1 + \hat{\Delta}_2 + \hat{\Delta}_3 + \hat{\Delta}_4 = 0, \tag{8}$$

and the argument above shows that $\hat{\Delta}_2 = 0$. On the other hand $\hat{\Delta}_1$ is equivalent to the RV-weighted IF effect and since it uses the neutral state distribution it can be calculated.

THE RV-WEIGHTED IF EFFECT

The story from this point is summarized in BOX 8. The use of reproductive values in our calculations of allele frequency change saves us from the intractable calculations of the selective state distribution. But there is a slight disconnect—in heterogeneous populations, the resulting measure $\hat{\Delta}_1$ of selective one-step change in allele reproductive value can fail to be equivalent to the measure $\Delta_1 + \Delta_2$ of selective one-step change in allele frequency. Since Δ_4 and $\hat{\Delta}_4$ are sign-equivalent, the reason for this disconnect has to be found in the term $\hat{\Delta}_3$.

BOX 8: The four weak-selection components of $E(\Delta x)$ and $E(\Delta \hat{x})$

Effects of selection	$E(\Delta x)$ (no RV)	$E(\Delta \hat{x})$ (RV-weighted)
Mixed states within-state effect	$\Delta_1 = \sum_{s\text{mixed}} \Delta x_s^{\delta} \pi_s^0$ IF effect (no RV)	$\hat{\Delta}_1 = \sum_{s\text{mixed}} \Delta \hat{x}_s^{\delta} \pi_s^0$ RV-weighted IF effect
Mixed states state distribution effect	$\Delta_2 = \sum_{s\text{mixed}} \Delta x_s^{0} \pi_s^{\delta}$ equals 0 in transitive pop	$\hat{\Delta}_2 = \sum_{s\text{mixed}} \Delta \hat{x}_s^{0} \pi_s^{\delta} = 0$ RV makes this zero.
Pure states within-state effect	$\Delta_3 = \sum_{s\text{pure}} \Delta x_s^{\delta} \pi_s^0 = 0$	$\hat{\Delta}_3 = \sum_{s\text{pure}} \Delta \hat{x}_s^{\delta} \pi_s^0$
Pure states state distribution effect	$\Delta_4 = \sum_{s\text{pure}} \Delta x_s^{0} \pi_s^{\delta} \equiv$ negative (ii) in BOX 3	$\hat{\Delta}_4 = \sum_{s\text{pure}} \Delta \hat{x}_s^{0} \pi_s^{\delta} \equiv$ negative (ii) in BOX 3
sum	0	0

Column 2, $E(\Delta x)$, measures changes in allele *frequency*; column 3, $E(\Delta \hat{x})$, measures changes in allele RV. In a transitive population these are the same (all $v_i = 1/N$) and rows 2 and 3 are both zero leaving us with rows 1 and 4 that are identical in magnitude and opposite in sign.

Row 1 employs the neutral state distribution and looks only at the mixed states, and that's exactly what inclusive fitness does. Thus row 1 gives the IF effect both using and not using an RV weighting.

The measures of A-fitness in BOX 3 refer to allele *frequency*. Thus they live in column 2. Measure (i) of BOX 3 is sign-equivalent to $\Delta_1 + \Delta_2$.

Row 4 is also interesting. By definition it gives us the net rate at which copies of A are produced or lost in the pure states when interactions are neutral. We calculate this in equation (6) and show that it is equivalent to the negative of measure (ii) of BOX 3.

Looking at the components in the 3-star population tabulated below we see that the RV-weighted IF effect $\hat{\Delta}_1$ can fail to be equivalent to measure (i) in BOX 3—the effect of selection on the one-step change in allele frequency $\Delta_1 + \Delta_2$. The key to understanding this lies in the interpretation of $\hat{\Delta}_3$.

The components in the 3-star population with BD updating	
$\Delta_1 = \frac{1}{81} q(1-q)\mu(-44b - 94c)$	$\hat{\Delta}_1 = \frac{2}{45} q(1-q)\mu(-11b - 21c)$
$\Delta_2 = \frac{1}{81} q(1-q)\mu(-10b + 4c)$	$\hat{\Delta}_2 = 0$
$\Delta_3 = 0$	$\hat{\Delta}_3 = \frac{2}{45} q(1-q)\mu(-b + c)$
$\Delta_4 = \frac{2}{9} q(1-q)\mu(3b + 5c)$	$\hat{\Delta}_4 = \frac{8}{45} q(1-q)\mu(3b + 5c)$
sum = 0	sum = 0

Look closely at Δ_3 and $\hat{\Delta}_3$ (BOX 8). Here, we are considering the effect of selection acting in each pure state on the one-step change in allele frequency. In the allB state selection can have no effect as there is no *b*–*c* behavior. In the allA state, selection will alter the relative probability that different nodes will be chosen to give birth. This will certainly not affect allele frequency as the offspring will always be B with probability $u(1–q)$ and otherwise will have the parental A. But the question of who is chosen to give birth does affect the node on which the offspring will establish and thus it can affect the RV of the offspring. This is the selective change in RV measured by $\hat{\Delta}_3$. This effect of selection might be thought to be negligible as it requires a mutation to take effect, but the fact is that the IF effect itself has the same order as the mutation rate, so that the two effects $\hat{\Delta}_1$ and $\hat{\Delta}_3$ work together to provide the total selective effect. The table at the bottom of BOX 8 is worth studying.

THE Δ-CALCULATION

The 3-star example tabulated in BOX 8 and analyzed in Appendix B is small enough (only six states) that we can work out its

transition matrix and from there the state frequency vector (BOX 5). For larger populations this approach quickly becomes infeasible. But if we use the neutral state distribution, recursive methods can be used to calculate average one-step allele frequency change—for example the neutral relatedness coefficients effectively calculate the average over all states, and these are calculated recursively (Taylor 2013). Now it is the odd-numbered Δ_i that use the neutral distribution, so these can all be calculated. Let us see how we can get the rest.

Suppose we have a heterogeneous population, and Δ_2 is nonzero. Fasten attention on the table at the end of BOX 8. Using reproductive values, we calculate $\hat{\Delta}_1$ and $\hat{\Delta}_3$. From the equilibrium equation (8), we get $\hat{\Delta}_4$. Now Δ_4 calculates the same thing as $\hat{\Delta}_4$ except in different units—Δ_4 uses frequency and $\hat{\Delta}_4$ use RV. Recall that these components live in the two pure states with selection turned off, so that the only change in allele frequency comes from mutation. Every mutation event that changes an A to B or a B to A changes the allele frequency by $\pm 1/3$. That is a Δ_4-change. What is the corresponding $\hat{\Delta}_4$-change? Well since all three nodes have birth probability 1/3, the offspring will land on the hub (node 0) with probability 2/3 and on a leaf (nodes 1 and 2) with probability 1/3. In the first case this is a change in RV of $\pm 1/5$ and in the second case this is a change in RV of $\pm 2/5$ (Appendix B). Thus the average change in RV is $(2/3)\,(\pm 1/5) + (1/3)\,(\pm 2/5) = \pm 4/15$. Thus to change from Δ_4 units to $\hat{\Delta}_4$ units, we multiply by the ratio of 4/15 to 1/3 and that is 4/5. In BOX 8, that is how 2/9 changes into 8/45.

Finally, Δ_1 can be calculated as the IF effect without using RV (though for those used to working with RV, care must be taken to count all the different kinds of fitness effects), and Δ_2 is obtained from equation (5).

Inclusive Fitness with Synergistic Fitness Effects

MATRIX GAMES

Evolutionary game theory goes back more than 40 years. In binary interactions, each player's payoff depends on the strategy employed by both partners, and the possible payoffs are recorded in a payoff matrix (Maynard Smith 1974; Grafen 1979; Maynard Smith 1982). Much recent work concerns the case of two strategies, treated here, giving us a 2×2 payoff matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \tag{9}$$

where the first row gives the payoffs to an A-player (with an A or B partner) and the second row belongs similarly to B. I assume that the payoffs represent small increments in fitness, small enough that we can ignore second-order effects in the a_{ij}. For each player, the payoff is added to its baseline fitness of 1. This weak-selection assumption, that the a_{ij} are small, allows us to write individual fitness as a linear combination of the payoffs (the a_{ij}) belonging to different interactions, but this is not enough to give us another form of additivity—additivity in the genotypes of the interacting individuals. That is the additivity that I am interested in here as that is what is required for the standard relatedness coefficients to work.

SYNERGY

Queller (1985) was the first to discuss this question, and to note that it required an extension of the standard notion of relatedness. He worked with the payoff matrix

$$\begin{bmatrix} b-c+d & -c \\ b & 0 \end{bmatrix}, \tag{10}$$

which has the "cost-benefit" form found in Hamilton's Rule, in that an A-individual gives benefit b to its partner at cost c, but if the partner is also A, it gives a "synergistic" bonus d. As an example, suppose nodes i and j play the matrix game and node j sends offspring to node k with probability m_{jk}. Then the fitness effect on k through this potential offspring is $w_K = -m_{jk}(bx_i - cx_j + dx_ix_j)$ and this is linear in the genotypes only if $d = 0$. The general 2×2 matrix A (eq. 9) can be put into the *b-c-d* form (eq. 10) by subtracting a_{22} from all terms. This gives an equivalent game as it changes the payoffs but not the final fitness effects. In this equivalence, the synergistic term d will be zero exactly when $a_{11} + a_{22} = a_{21} + a_{12}$. The advantage of the form (eq. 10) is that it situates the quadratic fitness effect entirely in the parameter d and if $d = 0$, the matrix (eq. 10) simply delivers the *b–c* interaction.

Having said that, I should point out that when a payoff matrix is put into this form, b need not have a benefit interpretation, nor need c be interpreted as a cost. As an example, the classic hawk-dove game (Maynard Smith 1978; Grafen 1979) has matrix $\left[\begin{smallmatrix} v & 0 \\ 2v & v-k \end{smallmatrix}\right]$ where row one belongs to the dove, v is half the value of the resource being contested and k is the cost of a fight for each contestant. This is put into the synergistic form (eq. 10) by subtracting v–k from all entries, giving us $\begin{bmatrix} k & k-v \\ 2vk+v & 0 \end{bmatrix} = \begin{bmatrix} b-c+d & -c \\ b & 0 \end{bmatrix}$ where $b = v + k$, $c = v - k$, and $d = -k$ (Taylor and Maciejewski 2012). But clearly b and c are not easily interpretable as benefit and cost.

Following Queller's (1985)article, a number of studies found different ways to handle these synergistic effects. The earliest of these employs regression coefficients (Queller 1992; Lehmann and Keller 2006; Gardner et al. 2011); this approach does indeed deliver fitness as a linear function of individual genotypes, but an important disadvantage for me is that the interaction

payoffs b, c, and d are buried in the regression coefficients and we lose sight of them. A second approach introduces higher order relatedness coefficients (Ohtsuki 2010; Taylor and Maciejewski 2012; Taylor 2013); these preserve the integrity of the payoffs and capture the quadratic behavior by the relatedness coefficients that now need to accommodate relatedness among more than two (for the model here, three) individuals. I will return to this in the discussion.

But now for a surprise. It turns out, quite unexpectedly, that in the standard finite population model with rare mutation discussed here, the mathematical difficulties posed by this form of synergy disappear and the standard inclusive fitness effect is able to calculate both Δ_1 and $\hat{\Delta}_1$—neither regression coefficients nor higher order relatedness coefficients are needed.

SYNERGY IN THE FINITE-POPULATION MODEL

Tarnita et al. (2009) work with the general payoff matrix (eq. 9) and show that under the assumptions of weak selection and that the only difference between A and B is found in the matrix payoffs (see Appendix C for details), there exists a parameter σ, dependent on the population structure but independent of the payoffs, for which the condition that allele A be fitter than B can be written:

$$\sigma(a_{11} - a_{22}) + (a_{12} - a_{21}) > 0. \tag{11}$$

Here, they define A-fitness using measure (ii) of BOX 3, but then it must hold for measures (i) and (iii) as well. If we replace the a_{ij} in condition (eq. 11) by the entries of the matrix (eq. 10) and rearrange, we get $(\sigma - 1)b - (\sigma + 1)c + \sigma d > 0$. Reparametrizing, taking β and γ to be the coefficients of b and $-c$, this tells us that the coefficient of d is halfway between β and γ, giving us the form

$$\beta b - \gamma c + \left(\frac{\beta + \gamma}{2}\right) d > 0. \tag{12}$$

We have seen above that in a transitive population, the IF effect is equivalent to measure (i) of BOX 3 giving us the sign equivalence:

$$W_{IF} \equiv \beta b - \gamma c + \left(\frac{\beta + \gamma}{2}\right) d \tag{13}$$

(Taylor and Maciejewski 2012). If we set $d = 0$ we have the standard IF effect $W_{IF} = \beta b - \gamma c$ (with the same β and γ, since these do not depend on the payoffs) and that tells us that β and γ can be calculated using the standard relatedness coefficients.

Equation (13) can be rewritten

$$W_{IF} \equiv \beta\left(b + \frac{d}{2}\right) - \gamma\left(c - \frac{d}{2}\right), \tag{14}$$

and we can say that the synergistic matrix (10) is "equivalent" to the additive benefit-cost matrix

$$\begin{bmatrix} B - C & -C \\ B & 0 \end{bmatrix}, \tag{15}$$

with $B = b + \frac{d}{2}$ and $C = c - \frac{d}{2}$. To be more precise, a finite population in which every A individual gives benefit $b + d/2$ to its partner (regardless of partner genotype) at cost $c - d/2$, will experience exactly the same overall conditions (averaged over all states) for the frequency of A to increase as for the synergistic interactions of the matrix (eq. 10) (Taylor 2016).

These inclusive-fitness remarks hold in transitive populations. To clarify this situation, note that the results of Tarnita et al. (2009) hold in all finite population structures, transitive or not. In extending these results to the IF analysis, Taylor and Maciejewski (2012) assumed that the IF effect was always equivalent to the effect of selection on the one-step change in allele frequency (measure (i) of BOX 3) provided reproductive value was used in heterogeneous populations. As described above, the subsequent analysis of Tarnita and Taylor (2014) showed that this may not be the case (BOX 8). This makes the analysis of Taylor and Maciejewski (2012) incomplete and I correct this in Appendix C.

If the synergistic interactions in the matrix (eq. 10) pose no problem for inclusive fitness, what was all the early fuss about? The point is that we are working here with the finite population model whereas, for example, Queller's (1985) analysis was set in what is often called the infinite-population model. So what is it that makes these two population models behave differently? I now show that this is not about population size per se, but is about its relation to the rate of what might be called genetic renewal.

MUTATION RATE AND POPULATION SIZE

We say that two individuals are identical by descent (IBD) if they descend from the same mutation event with no intervening mutation. Now our assumption of rare mutation together with small population size allows us to assume (to first order in u) that mutation never occurs when there is more than one IBD class in the population. The point is that drift will clear the population of more than one IBD class much more quickly than mutation is able to provide a new such class (BOX 2). In particular, in a mixed state (which is where selection acts) any two copies of the same allele must be IBD. This turns out to be the critical condition behind the Tarnita et al. (2009) result and the IF formulation (eq. 13). To illustrate that, I provide here a direct heuristic argument for the equivalence of the synergistic matrix (eq. 10) with the additive matrix (eq. 15). A rigorous argument is found in Appendix C.

HEURISTIC ARGUMENT FOR THE EQUIVALENCE OF THE MATRICES (EQ. 10) AND (EQ. 15)

To begin, I remark that one of the main assumptions of the Tarnita et al. (2009) article is that the interaction payoffs provide the only difference between the alleles A and B (Appendix C, assumption AS2). Thus, in the neutral population (all payoffs equal to 0) the probability G that a focal individual is IBD to an interactant is the same whether the focal genotype is A or B. Thus, using the payoff matrix (eq. 10) the average payoff to a focal A or B player is

$$\begin{aligned} W_{\mathrm{A}} &= G(b-c+d)+(1-G)(-c) \\ W_{\mathrm{B}} &= (1-G)b. \end{aligned} \tag{16}$$

Here, we have used the fact that in a mixed state any two copies of the same allele must be IBD. If we make the same calculation for the payoff matrix (eq. 15), we get

$$\begin{aligned} W_{\mathrm{A}} &= G(b-c+d)+(1-G)(-c+d/2) \\ W_{\mathrm{B}} &= (1-G)(b+d/2). \end{aligned} \tag{17}$$

Now observe that the difference $W_{\mathrm{A}} - W_{\mathrm{B}}$ is the same for both (eq. 16) and (eq. 17). Alternatively if we add $(1–G)d/2$ to both payoffs in (eq. 16) we get the payoffs in (eq. 17). Since increasing all payoffs by the same amount has no effect on any fitness effect, the two matrices give equivalent conditions.

This argument emphasizes that the rarity of mutation is a critical reason the synergistic matrix (eq. 10) is effectively additive, in that it causes genetic identity in state (having the same allele) to be the same as identity by descent in any mixed state interaction. Said another way, effective nonadditivity in the matrix (eq. 10) requires interactions between A-individuals who are not IBD. Such interactions happen routinely in the infinite-population model when the force of genetic renewal (e.g., long-range migration) is not rare (as it typically the case).

To be more general, this result is actually about the interaction between population size and genetic renewal (BOX 2). If genetic renewal is rare enough that the time required for genetic drift to clear multiple IBD classes is much shorter than the time between renewal events, then the matrix (eq. 10) will behave additively, that is, it will be equivalent to (eq. 15).

THE INFINITE POPULATION MODEL

To emphasize this I look briefly at the infinite population model for which Taylor and Maciejewski (2012) obtained the following generalization of equation (13):

$$\mathrm{W}_{\mathrm{IF}} = \left(\beta b - \gamma c + \frac{\beta+\gamma}{2}d\right) + \alpha\left(p-\frac{1}{2}\right)d, \tag{18}$$

where α, β, and γ are independent of both the matrix payoffs and the equilibrium allele frequency p. Here, genetic renewal is generated by periodic migration "from afar" and p is the frequency of A among these immigrants and thus it is the analogue of q in the finite population model. The first bracket has the form of equation (13) and β and γ involve only the standard linear relatedness coefficients, whereas the coefficient α typically requires the calculation of more complex higher order relatedness coefficients. However, Taylor and Maciejewski (2012) show that, relative to β and γ, the coefficient α has the same order as the rate of genetic renewal—so if this is rare, frequency dependence disappears to first order in this rate. For example, for an infinite island model (Wright 1931) with randomly mixing demes of size n and BD updating (Moran 1958), Taylor and Maciejewski (2012) obtain

$$W_{\mathrm{IF}} \sim \left(-c+\frac{1}{2}d\right) + \left(p-\frac{1}{2}\right)\left(\frac{nm}{nm+2(1-m)}\right)d, \tag{19}$$

and we can clearly see the relative dependence of the final term on the migration rate m. If migration was rare, we could ignore the final term and, as for the finite population model, the synergistic effects could be handled with the standard relatedness coefficients.

Discussion

I have worked here with what is often called the finite population model. Over the 50-year span of our study of inclusive fitness, it has been a focus of interest only for the last third. This interest began with the work of Rousset and Billiard (2000) and continued from there with the growing interest in evolutionary graph theory (Lieberman et al. 2005). The ideas that have emerged have greatly enriched our understanding of strategy fitness and allowed us to compare the inclusive fitness effect with an array of other fitness measures, for example those in BOX 3.

My interest here has been to gain a better understanding of recent studies of two theoretical aspects of inclusive fitness, one that seeks to understand its interaction with heterogeneous population structure, and the other that looks at the way it handles nonadditive (synergistic) fitness effects. Both of these have a long-standing history in the shaping of inclusive-fitness theory but recently special considerations have arisen in the finite-population model. Ideas of class structure and the significance of reproductive value certainly go back to Fisher (1930), and Hamilton (1972) was the first to incorporate these into an inclusive-fitness model. Subsequent discussions of the role of reproductive value are found in Charlesworth and Charnov (1981), Taylor (1990), Rousset (1999), Pen and Weissing (2000), and Taylor (2009). The special status of synergistic fitness effects in matrix games was first discussed by Queller (1985) and subsequently analyzed by Queller (1992), Lehmann and Keller (2006), Ohtsuki (2010), Gardner et al. (2011), and Taylor and Maciejewski (2012).

HETEROGENEITY AND REPRODUCTIVE VALUE

We have seen that in a finite transitive (BOX 1) population structure the inclusive fitness effect gives a measure of allele fitness equivalent to those obtained by other standard approaches (BOX 3), though as always, one needs to look carefully at the assumptions being made. But in a heterogeneous or class-structured population, this is not in general the case and I have illustrated this in some detail using a small ($N = 3$) population.

To discuss the reason for this, we must first understand that selection acts on allele frequency in two different ways, first in each population state, it alters the fitness of different individuals, and secondly (effectively as a result of the first), it alters the state distribution, the frequency with which certain states occur. Standard inclusive fitness, in working with the neutral state distribution, accounts only for the first of these. Of course one must ask at this point how allele fitness should be measured and it has been argued (Grafen 2015) that fitness effects arising from changes in allele distribution, for example across reproductive classes (which will be the case here), are a separate category and, following Fisher (1930), should perhaps not be counted as part of the effects of selection on allele frequency. That may well have merit, but my interest here is in comparing the IF effect with other standard measures of allele fitness and the fact of the matter is that in finite heterogeneous populations the effect of selection on the one-step change in allele frequency, measure (i) of BOX 3, has to include this effect of allele distribution to be generally equivalent to other standard measures.

Under population transitivity, and this was often implicitly assumed in the early models, the effect of the altered state distribution on its own has no effect on allele frequency change, at least to first order in the strength of selection. The reason for this is that when no selection is acting, there is no average change of allele frequency from any particular state, and thus the average change over all states will be zero no matter how we weight the states.

But in a heterogeneous population, this may no longer be the case and the altered state distribution can have a significant effect on allele frequency. There is a fix for this and that is to use reproductive values as a weight on the nodes when calculating the IF effect (e.g., compare Δ_2 and $\hat{\Delta}_2$ in BOX 8). However, the RV-weighted IF effect ($\hat{\Delta}_1$ in BOX 8) might no longer provide the entire effect of selection on allele-frequency change ($\Delta_1 + \Delta_2$ in BOX 8). The reason is that once we start to work with RV-weighted allele frequency, another factor can enter, and that is the effect of selection, not in the mixed states (which is where inclusive fitness operates) but in the pure states where a node inhomogeneity can cause a change in the RV of a new mutant—a change that acts differently on different alleles ($\hat{\Delta}_3$ in BOX 8). That is an effect that IF does not capture but the standard measures of selective allele frequency change (BOX 3) do.

This result is quite sensitive to the assumptions of the model. Here, I have assumed BD updating but with DB updating (and fecundity payoffs) $\hat{\Delta}_3 = 0$ and the effect disappears. The reason is simple enough; DB updating together with fecundity payoffs produces no effect on the node occupied by a new mutant. Another assumption concerns the timing of mutation or any other form of genetic renewal.

This IF anomaly in a heterogeneous population does not arise in the infinite population model. Here, the RV-weighted IF effect does capture the full effect of selection on the one-step change in RV-weighted allele frequency. The reason is that in this model, the population is virtually always in a mixed state unlike the finite model in which the mixed-state frequencies are of order u. And in a mixed state the IF effect does captures everything, provided of course that under heterogeneity, we use RV weights (Taylor 1990). Most of the 50-year history of inclusive fitness has worked in a large or infinite population and that is one reason that this exception has not been discussed until recently.

Adaptive dynamics (Geritz et al. 1998) is an active area that works with a large population in which fixation probability is a prominent feature. But the population model it works with is quite different from the finite population model here. Our fitness measure here is an average over all population states but adaptive dynamics works largely with states in which one or the other of the alleles is rare; thus it requires the notion of frequency-dependent fitness. The populations it works with are typically transitive (e.g., randomly mixed) but it might be interesting to apply the ideas here to adaptive dynamics in a heterogeneous population.

SYNERGISTIC INTERACTIONS

Queller (1985) first drew attention to the problems experienced by the inclusive fitness approach when there are synergistic interactions, specifically that generalized "quadratic" covariances are needed to play the role of relatedness and these are difficult to calculate. Not long afterwards he proposed a solution, that a focus on genes (as is the case here) rather than on phenotypes, together with the use of partial regression coefficients (of fitness on genotypes of interactants), can give a completely general formulation of Hamiton's Rule (Queller 1992). An excellent review of this approach is found in Gardner et al. (2011). Queller (1992) did point out that while this approach was theoretically significant, it did lose sight of the different phenotypic effects that one can more easily measure and that we are certainly more interested in here. In terms of the analysis here, these "phenotypic effects" correspond to the entries b, c, and d of the payoff matrix (eq. 10) that are typically buried in the regression coefficients, and thus my preferred solution to the synergism dilemma is to assign the nonlinearity to the relatedness coefficients and consider the relatedness of "a pair of actors" to a recipient (Ohtsuki 2010; Taylor 2013).

Having said all this, it is a surprise to discover that, in the finite population model, synergistic fitness effects do not actually require the use of the generalized (quadratic) relatedness coefficients (Tarnita et al. 2009; Taylor and Maciejewski 2012; Taylor 2016). Indeed equation (13) shows that the coefficient of d in the W_{IF} formula is simply the average of the coefficients of b and c, and these can be calculated in terms of the standard relatedness coefficients.

It is worth saying a bit about this result. Equation (18) belongs to an infinite population model in which genetic renewal (e.g., through migration from afar in a population with allele frequency p) is not necessarily rare. Alan Grafen, in comments on an earlier draft, suggested that the heuristic argument for the equivalence of equations (16) and (17) might be made to work in such a case if G were replaced by the probability of having an identical allele (call this $G^{\#}$) that might or might not be identical by descent (IBD). These will not be the same when genetic renewal is not rare. The problem with this is that $G^{\#}$ will be different for A- and B-individuals unless the population allele frequency p is ½. But in that case, the alternative argument does work and we get the IF form (eq. 13). This can of course be directly obtained from equations (18) and (19) setting $p = ½$.

I have a few reasons for deciding to discuss synergistic effects here. First of all I regard Tarnita et al. (2009) as a significant piece of work leading to an important feature of the finite population model. In thinking about this result I am struck by the fact that my own work with nonadditive interactions, using only the matrix form (eq. 10), would have had great difficulty finding that result as the argument (Appendix C) relies heavily on the more symmetric structure of the matrix (eq. 9). That is a nice example of the importance of the right notation in facilitating discovery.

The second reason has to do with the "understanding" of this result, that it really is not about population size at all but about the relationship between population size and the rate of genetic renewal. What we need is a population timeline, such as found in BOX 2, such that there is unlikely to be a mutation in any "mixed-state" interval between fixations—for example this will be the case if $u << 1/N$. What this gives us, and this is in fact the critical condition for equation (13), is that in a mixed state two individuals with the same genotype will be IBD.

It is worth reinterpreting this in an infinite-population model. Take for example the infinite (or large) island population belonging to equation (19) with demes of size n and migration between demes at rate m. If the product mn is small ($<< 1$) then the last term in equation (19) can be ignored and we are left with the additive form of equation (13). In this case, each *deme* would follow the timeline of BOX 2 and since interactions are within each deme, we could again say that in a mixed deme two individuals with the same genotype will be IBD.

The third reason is that in my own theoretical work, I made the assumption that (under weak selection and rare mutation) the IF effect is equivalent to the selective one-step change in allele frequency (measure (i) BOX 3) in both transitive and nontransitive populations but that turned out to be incorrect requiring the argument in Appendix C.

ACKNOWLEDGMENTS

This work was supported by Grant 8145 from the Natural Sciences and Engineering Council of Canada. I am hugely grateful for Alan Grafen and one anonymous reviewer who went carefully through two drafts of this manuscript with substantial changes leading finally to the present form. I must also cite and thank the Associate Editor Eva Kisdi for an incisively careful final reading and for much wise and generous commentary helping me find along the way the right form for this submission that started out halfway between a research article and a review article.

LITERATURE CITED

Abbot P, Abe J, Alcock J, Alizon S, Alpedrinha JAC, Andersson M, Andre J-B, van Baalen M, Balloux F, Balshine S. et al. 2011. Inclusive fitness theory and eusociality. Nature 471. Part of a collection of brief communications E1–E10.

Allen, B., and C. E. Tarnita. 2014. Measures of success in a class of evolutionary models with fixed population size and structure. J. Math. Biol. 68:109–543.

Charlesworth, D., and E. L. Charnov.1981. Kin selection in age-structured populations. J. Theor. Biol. 88:103–119.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper and Row, New York.

Fisher, R. A. 1927. The actuarial treatment of official birth records. Eugenics Rev. 19:103–108.

———. 1930. The genetical theory of natural selection. Clarendon Press, Oxford (reprinted and revised, 1958, 1981. Dover, New York).

Geritz, S. A. H., É. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evol. Ecol. 12:35–57.

Gardner, A., S. A. West, and G. Wild. 2011. The genetical theory of kin selection. J. Evol. Biol. 24:1020–1043.

Grafen, A. 1979. The hawk-dove game played between relatives. Anim. Behav. 27:905–907.

———. 1985. Hamilton's rule OK. Nature 318:310–311.

———. 2015. Biological fitness and the fundamental theorem of natural selection. Am. Nat. 186:1–14.

Grafen, A., and M. Archetti. 2008. Natural selection of altruism in inelastic homogeneous viscous populations. J. Theor. Biol. 252:694–710.

Hamilton, W. D. 1964. The genetical evolution of social behaviour I. J. Theor. Biol. 7:1–16.

———. 1972. Altruism and related phenomena, mainly in social insects. Ann. Rev. Ecol. Syst. 3:192–232.

Lehmann, L., and L. Keller. 2006. Synergy, partner choice and frequency dependence: their integration into inclusive fitness theory and their interpretation in terms of direct and indirect fitness effects. J. Evol. Biol. 19:1426–1436.

Leturque, H., and F. Rousset. 2002. Dispersal, kin competition, and the ideal free distribution in a spatially heterogeneous population. Theoret. Popul. Biol. 62:169–180.

Lieberman, E., C. Hauert, M. A. Nowak. 2005. Evolutionary dynamics on graphs. Nature 433:312–316.

Maciejewski, W. 2014. Reproductive value in graph-structured populations. J. Theor. Biol. 340:285–293.
Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. J. Theor. Biol. 47:209–221.
———. 1978. Optimization theory in evolution. Ann. Rev. Ecol. Syst. 9:31–56.
———. 1982. Evolution and the theory of games. Cambridge Univ. Press, Cambridge.
Maynard Smith, J., and G. R. Price. 1973. The logic of animal conflict. Nature 246:15–18.
Michod, R. E., and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. Nature 288:694–697.
Moran, P. A. P. 1958. Random processes in genetics. Proc. Cambridge Phil. Soc. 54:60–71.
Nowak, M. A., A. Sasaki, C. Taylor, and D. Fudenberg. 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428:646–650.
Nowak, M. A., C. E. Tarnita, and E. O. Wilson. 2010. The evolution of eusociality. Nature 466:1057–1062.
Ohtsuki, H. 2010. Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. Evolution 64:3344–3353.
Ohtsuki, H., and M. A. Nowak. 2006. Evolutionary games on cycles. Proc. R Soc. B 273:2249–2256.
Ohtsuki, H., M. A. Nowak, and J. Pacheco. 2007. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. Phys. Rev. Lett. 98:108106.
Pen, I, and F. J. Weissing. 2000. Towards a unified theory of cooperative breeding: the role of ecology and life history re-examined. Proc. R Soc. B 267:2411–2418.
Price, G. R. 1970. Selection and covariance. Nature 227:520–521.
Queller, D. C. 1985. Kinship, reciprocity and synergism in the evolution of social behaviour: a synthetic model. Nature 318:366–367.
———. 1992. A general model for kin selection. Evolution 46:376–380.
Rousset, F. 1999. Reproductive value vs sources and sinks. Oikos 86:591–596.
Rousset, F., and S. Billiard. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J. Evol. Biol. 13:814–825.
Tarnita, C. E., H. Ohtsuki, T. Antal, F. Fu, and M. A. Nowak. 2009. Strategy selection in structured populations. J. Theor. Biol. 259:570–581.
Tarnita, C. E., and P. Taylor. 2014. Measure of relative fitness of social behaviors in finite structured populations. Am. Nat. 184:477–488.
Taylor, P. D. 1990. Allele frequency change in a class-structured population. Am. Nat. 135:95–106.
———. 1996. Inclusive fitness arguments in genetic models of behaviour. J. Math. Biol. 34:654–674.
———. 2009. Decompositions of Price's formula in an inhomogeneous population structure. J. Evol. Biol. 22:201–213.
———. 2010. Birth-death symmetry in the evolution of a social trait. J. Evol. Biol. 23:2569–2578.
———. 2013. Inclusive and personal fitness in synergistic evolutionary games on graphs. J. Theor. Biol. 21:76–82.
———. 2016. Hamilton's Rule in finite populations with synergistic interactions. J. Theoret. Biol. 397:151–157.
Taylor, P. D., T. Day, and G. Wild. 2007a. From inclusive fitness to fixation probability in homogeneous structured populations. J. Theor. Biol. 249:101–110.
———. 2007b. Evolution of cooperation in a finite homogeneous graph. Nature 447:469–472.
Taylor, P. D., and W. Maciejewski. 2012. An inclusive fitness analysis of synergistic interactions in structured populations. Proc. R Soc. B 279:4596–4603.
———. 2014. Hamilton's inclusive fitness in finite-structured populations. Phil. Trans. R Soc. B 369:20130360.
West, S. A. and A. Gardner. 2010. Altruism, spite and greenbeards. Science 327:1341–1344.
Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.
Wu, B., C. S. Gokhale, L. Wang, and A. Traulsen. 2012. How small are small mutation rates? J. Math. Biol. 803–827.

Associate Editor: E. Kisdi
Handling Editor: M. Servedio

Appendix A

Inclusive fitness and the one-step change

In the finite-population model, one measure of the fitness of the allele A (measure (i) of BOX 3) calculates the one-step selective increase in allele frequency at each state and then takes the average of this over all states. But in practice, rather than handle each state separately and then take the average over all states (since state frequencies are hard to get hold of), we take each *node* separately (in fact it is enough to take one focal node in each reproductive class) and calculate its average fitness (fecundity and mortality) over all possible states—and then finally average this over all nodes. Of course the fitness effects at each node depend on the genotypes of the focal neighbors, but if we know the probability that each neighbor is A or B, that should be enough to allow us to take the average. Those probabilities are effectively what Hamilton (1964) called the coefficients of relatedness.

To avoid having to take account of the genotype distribution of pairs and triples of neighbors, etc. we make an assumption of additivity of effects. If the focal individual receives help from several neighbors we do not expect that in general the total benefit will be the sum of the benefits from each individual action. If these effects are small, then this additive assumption should be a reasonable approximation and subsequent analyses have shown that this "weak selection" assumption can deliver the needed additivity (Grafen 1985; Taylor 1996).

A transitive population

In this case, the "personal fitness" effect of the A-behavior on a focal A-individual has the mathematical form:

$$W_{PF} = \left(\sum\nolimits_i \beta_i R_i\right) b - \left(\sum\nolimits_i \gamma_i R_i\right) c, \qquad \text{(A.1)}$$

where the sum runs over all individuals i whose behavior affects focal fitness and the coefficients β_i and γ_i provide the fractional effects on the focal individual of the benefit b and cost c. The coefficient R_i is the relatedness of the focal individual to individual i and it effectively provides the probability that individual i carries the allele A (and therefore exhibits the A-behavior) given that the focal individual carries A (Michod and Hamilton 1980).

Hamilton (1964) turned this cause-effect relationship around, taking a single focal A-actor and adding up all the fitness effects of its A-behavior on the fitness of others, each effect again weighted by the focal relatedness R_i to the recipient. The resulting expression looks the same but it is now called the inclusive fitness effect of the allele A:

$$W_{IF} = \left(\sum_i \beta_i R_i\right) b - \left(\sum_i \gamma_i R_i\right) c, \qquad (A.2)$$

where the sum runs over all individuals affected by the focal A-behavior and the R_i provide the probability that these individuals carry the focal allele. Thus this is the effect of the action of a single copy of A on the fitness of all neighboring copies. The conceptual advantage of this formulation is that it puts the bearer of the allele A in the driver's seat, so to speak, such that it can be viewed as selecting its behavior with regard to all the various effects it might have on the fitness of many neighbors, each weighted by its relatedness to the focal actor.

As an example, suppose that the focal A-individual provides fecundity benefit b to node i that sends its offspring to nodes j and k with probabilities 3/4 and 1/4, respectively. In that case the IF effect of this single benefit will be

$$W_{IF} = \left(\frac{3}{4}(R_i - R_j) + \frac{1}{4}(R_i - R_k)\right) b. \qquad (A.3)$$

Note that I have chosen to break the R_i up into its two components, each corresponding to a node replacement—on nodes j and k. This is my preferred way to do the accounting and it is useful when we use reproductive values. The fecundity gift of b to i is called a *primary* effect, and the resulting mortality effect on j and k is called secondary (West and Gardner 2010).

A heterogeneous or class-structured population

In this case nodes in different classes may have different reproductive value v_i and both the primary and the secondary effects of each payoff should be weighted by the node RV. For the *i-j-k* example above, the RV-weighted IF effect will be

$$\hat{W}_{IF} = \left(\frac{3}{4}(R_i - R_j)v_j + \frac{1}{4}(R_i - R_k)v_k\right) b. \qquad (A.4)$$

Note when a behavior affects the probability of a birth event, the relatedness belongs to the node of the parent but the offspring is valued according to the node it inhabits.

Appendix B

Inclusive fitness analysis of the star population with $N = 3$ nodes (Fig. A1)

Node 0 is called the hub and is connected to nodes 1 and 2, called the leaves. The hub has two interactions and each leaf has one.

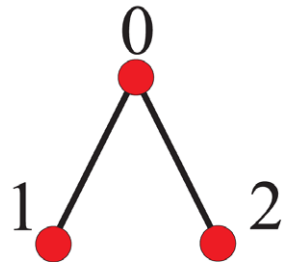


Figure 1. The 3-star graph

Offspring dispersal from a leaf is only to the hub and from the hub it is to each leaf with probability ½. This is a small population but it still has all the complexities that concern us and it has a manageable number of states. I use BD updating with fecundity payoffs.

RV calculation

There are two reproductive classes, the hub class $H = \{0\}$ and the leaf class $L = \{1, 2\}$. To use simple numbers for the calculation, suppose all nodes are given one offspring:

RV gain through birth = RV loss through death.

$$\text{Hub} : v_L = 2v_H$$

$$\text{Leaf} : v_H = 1/2v_L$$

Note that the equations are identical. If there are C reproductive classes, the RVs will be determined by C–1 of the C possible equations. The solution that sums to one ($v_0 + v_1 + v_2 = 1$) is

$$v_H = 1/5 \text{ and } v_L = 2/5.$$

Relatedness calculation

As relatedness coefficients between nodes, I use the probabilities G that the nodes are identical by descent (IBD) (Michod and Hamilton 1980; Rousset and Billiard 2000; Taylor 2013). Let these be between:

$$\text{Individual and self} : \quad G_0 = 1$$

$$\text{Hub and leaf} : \quad G_1$$

$$\text{Leaf and leaf} : \quad G_2$$

The one-step recursive equations between these are:

$$G_1 = (1-u)\left(\frac{(1/2)G_0 + (1)G_0 + (1)G_2}{5/2}\right)$$

$$= (1-u)\left(\frac{3G_0 + 2G_2}{5}\right)$$

$$G_2 = (1-u)G_1$$

For the first, consider the hub-leaf pair. The most recent replacement came at the leaf from the hub at rate ½ ($G = 1$) or at the hub from the same leaf at rate 1 ($G = 1$) or at the hub from the other leaf at rate 1 ($G = G_2$). To form an average I have divided by the total rate 5/2. The second belongs to the leaf-leaf pair. Replacement is from the hub giving G_1. These equations solve to give, to first order in u:

$$G_1 = 1 - \frac{7}{3}u$$

$$G_2 = 1 - \frac{10}{3}u.$$

Inclusive fitness calculation

We calculate the fitness of the allele A by adding up the fitness effects of all interactions of a focal A-individual. There are two ways of organizing this calculation: working with a focal recipient (personal fitness) or with a focal actor (inclusive fitness). Here, I use the latter. For this I need to take a focal individual in each actor class. In this case, both classes act (sometimes only one class acts but both classes feel the effects) so I consider a focal hub actor and a focal leaf actor. Finally, I add up the effects weighting each actor by its number of primary interactions of each type.

Hub actor (*interacting with one leaf*) Gives b to leaf $R = G_1$. The extra offspring go to the hub with RV $= v_H$ and increase hub mortality—relatedness to actor (self) is $R = G_0$ and again RV $= v_H$.

$$\text{Effect}: \quad b(G_1 - G_0)v_H$$

Gives $-c$ to self: $R = G_0$. The extra offspring go to the leaf RV $= v_L$ and increase leaf mortality $R = G_1$ and again RV $= v_L$.

$$\text{Effect}: \quad -c(G_0 - G_1)v_L$$

$$\text{Total effect}: \quad \hat{W}_{IF}(H) = b(G_1 - G_0)v_H - c(G_0 - G_1)v_L = (bv_H + cv_L)(G_1 - G_0)v_L = \tfrac{1}{15}(-7b - 14c).$$

Leaf actor (*interacting with hub*) Gives fecundity b to hub: $R = G_1$. The extra offspring goes to a leaf RV $= v_L$ and increases leaf mortality $R = (G_0+G_2)/2$ (leaf to self and other leaf) and again RV $= v_L$.

$$\text{Effect}: \quad b\left(G_1 - \frac{G_0 + G_2}{2}\right)v_L.$$

Gives $-c$ to self: $R = G_0$. The extra offspring goes to hub RV $= v_H$ and increases hub mortality $R = G_1$ and again RV $= v_H$.

$$\text{Effect}: \quad -c(G_0 - G_1)v_H$$

$$\text{Total effect}: \quad \hat{W}_{IF}(L) = b\left(G_1 - \tfrac{G_0+G_2}{2}\right)v_L - c(G_0 - G_1)v_H = \tfrac{1}{15}(-4b - 7c)$$

Sum over both focal actors The focal hub actor has two interactions (with both leaves) and the leaf actor has two interactions (as there are two leaves).

$$\hat{W}_{IF} = 2\hat{W}_{IF}(H) + 2\hat{W}_{IF}(L) = u\frac{2}{15}(-11b - 21c)$$

to first order in the mutation rate u. We often omit the u-multiplier, though it's a useful reminder that in this model, the IF effect is of order u. Comparing this result with BOX 8 we see it is 3 times the measure $\hat{\Delta}_1$ tabulated there. The reason is that our analysis here counted changes in allele *number* rather than allele *frequency*, and with $N = 3$, the latter is 1/3 of the former.

Appendix C

The relationship between the result of Tarnita et al. (2009) and the IF effect in a heterogeneous population

Our argument here draws heavily on the proof found in Tarnita et al. (2009). Suppose we have a finite graph-structured population of asexual haploid individuals occupying all nodes, with rare mutation and fitness determining pairwise interactions such that an A-individual gives benefit b to its partner at cost c, but if the partner is also A, it gives a "synergistic" bonus d. We make the following two assumptions:

AS1: Selection is weak. More precisely the conclusions are valid to first order in the strength of selection. As part of this assumption we need the fitness components (e.g., the fecundities, mortalities, and the state transition probabilities) to depend differentiably on the strength of selection (otherwise "first order" has little meaning).

AS2: The payoffs provide the only difference between the alleles A and B. For example, there are no differences connected with the update rules, partner choice or offspring dispersal (e.g., in the neutral population, the probability an offspring born on node i will disperse to node j is independent of the allele it carries).

Then the IF effect can be calculated with the standard "additive" inclusive-fitness analysis and has the mathematical form:

$$W_{IF} = \beta b - \gamma c + \left(\frac{\beta + \gamma}{2}\right)d,$$

where the coefficients β and γ are dependent on the population structure but independent of the payoffs. Here it is understood that in a heterogeneous population (Taylor 2009; Tarnita and Taylor 2014), the IF effect $\hat{W}_{IF}$ will be formulated with reproductive values.

Proof. In a heterogeneous population, the IF effect is equivalent to what we have called $\hat{\Delta}_1 = \sum_{s\text{mixed}} \Delta\hat{x}_s^{\delta}\pi_s^0$, the effect of selection on the average one-step change in allele RV calculated with the neutral state distribution. With a slight notational modification, I suppress the strength of selection δ, and in terms of the

general matrix (eq. 9), I let $\Delta\hat{x}_s^{\mathrm{A}} = \Delta\hat{x}_s^{\mathrm{A}}(a_{11}, a_{12}, a_{21}, a_{22})$ denote the average one-step RV change for the allele A at the state s, and let $\Delta\hat{x}_s^{\mathrm{B}}$ be the corresponding change for allele B. For example for the state $s = \mathrm{AAB}$ in BOX 7, we have $\Delta\hat{x}_s^{\mathrm{A}} = -\frac{2}{15}c = \frac{2}{15}a_{12}$. Then

$$\Delta\hat{x}_s^{\mathrm{A}}(a_{11}, a_{12}, a_{21}, a_{22}) = -\Delta\hat{x}_s^{\mathrm{B}}(a_{11}, a_{12}, a_{21}, a_{22})$$
$$= -\Delta\hat{x}_s^{\mathrm{A}}(a_{22}, a_{21}, a_{12}, a_{11}) \quad \text{(C1)}$$

The first equation holds because the RV sum over all nodes is always 1 so what A gains, B will lose and the second equation follows from AS2 above so that if we both interchange the alleles and permute the corresponding matrix entries the result should not change.

Now take the average of the first and last expressions over all states s using the neutral state distribution noting that π_s^0 is independent of the payoffs:

$$\hat{\Delta}_1(a_{11}, a_{12}, a_{21}, a_{22}) = -\hat{\Delta}_1(a_{22}, a_{21}, a_{12}, a_{11}). \quad \text{(C2)}$$

Using the differentiability of $\hat{\Delta}_1$ (AS1) and the fact that it is zero at neutrality, the first-order multivariable Taylor expansion of $\hat{\Delta}_1$ will have the form:

$$\hat{\Delta}_1 = k_1 a_{11} + k_2 a_{12} + k_3 a_{21} + k_4 a_{22}, \quad \text{(C3)}$$

where the k_i are independent of the payoffs. From (B2):

$$k_1 a_{11} + k_2 a_{12} + k_3 a_{21} + k_4 a_{22}$$
$$= -\left(k_1 a_{22} + k_2 a_{21} + k_3 a_{12} + k_4 a_{11}\right). \quad \text{(C4)}$$

Now since (B4) holds for all a_{ij}, we deduce that $k_4 = -k_1$ and $k_3 = -k_2$. This allows us to write (B3) as

$$\hat{\Delta}_1 = k_1(a_{11} - a_{22}) + k_2(a_{12} - a_{21})$$
$$\sim \sigma(a_{11} - a_{22}) + (a_{12} - a_{21}),$$

where I have taken $\sigma = k_1/k_2$ and used $\sim$ to denote sign equivalence. This gives us equation (11).