

Appendix from L. Lehmann et al., “Cultural Transmission Can Inhibit the Evolution of Altruistic Helping”

(Am. Nat., vol. 172, no. 1, p. 12)

Evaluation of the Selection Gradients

Pairwise Payoff Comparison Rule

We derive here the transmission function t_{ij} for the model where imitation occurs with a “pairwise payoff comparison” rule. Under this transmission rule, with probability $1 - m_{ct}$, a focal individual compares his/her payoff with that of a randomly chosen individual from his/her group (excluding himself/herself) and adopts the variant of the group mate with a probability given by the payoff to that individual relative to the sum of the focal individual’s and the group mate’s payoffs. With complementary probability m_{ct} , the focal individual compares his/her payoff with that of an individual chosen at random from another group (nonfocal group) and adopts the variant of that individual with a probability given by the payoff to that individual relative to the sum of the focal individual’s and the chosen individual’s payoffs. This updating rule is equivalent to the one considered by Boyd and Richerson (2002), Boyd et al. (2003), and Lehmann et al. (2007), although Boyd and Richerson (2002) considered a more general population structure where there is isolation by distance, but they assumed infinite group size. Here we consider the effect of finite group size but without isolation by distance. From the assumptions just described, the probability that individual i from group j retains its cultural variant is

$$\Pr(ij \leftarrow ij) = \sum_{k \neq i} \frac{(1 - m_{ct})(1 + f_{ij})}{(N - 1)[(1 + f_{ij}) + (1 + f_{kj})]} + \sum_{l \neq j} \sum_k \frac{m_{ct}(1 + f_{ij})}{(n_d - 1)N[(1 + f_{ij}) + (1 + f_{kl})]}, \quad (A1)$$

while the probability that individual k ($k \neq i$) from group j adopts the cultural variant of individual i from group j is

$$\Pr(kj \leftarrow ij) = \frac{(1 - m_{ct})(1 + f_{ij})}{(N - 1)[(1 + f_{ij}) + (1 + f_{kj})]}. \quad (A2)$$

Finally, the probability that individual k from group l adopts the cultural variant of individual i from group j is

$$\Pr(kl \leftarrow ij) = \frac{m_{ct}(1 + f_{ij})}{(n_d - 1)N[(1 + f_{ij}) + (1 + f_{kl})]}. \quad (A3)$$

By inserting these equations into the transmission function (eq. [9]) expanded as

$$t_{ij} = \Pr(ij \leftarrow ij) + \sum_{k \neq i} \Pr(kj \leftarrow ij) + \sum_{l \neq j} \sum_k \Pr(kl \leftarrow ij), \quad (A4)$$

we obtain

$$t_{ij} = (1 - m_{ct}) \left[\frac{1}{N - 1} \sum_{k \neq i} \frac{2(1 + f_{ij})}{(1 + f_{ij}) + (1 + f_{kj})} \right] + m_{ct} \left[\frac{1}{(n_d - 1)N} \sum_{l \neq j} \sum_k \frac{2(1 + f_{ij})}{(1 + f_{ij}) + (1 + f_{kl})} \right]. \quad (A5)$$

Under weak selection, $2(1 + f_{ij})/[(1 + f_{ij}) + (1 + f_{kl})] \approx 1 + (f_{ij} - f_{kl})/2$, and using $[1/(n_d - 1)N] \sum_{l \neq j} \sum_k f_{kl} \rightarrow f$ as $n_d \rightarrow \infty$ allows us to write

$$t_{ij} \simeq 1 + \frac{(1 - m_{ct})}{2} \left(f_{ij} - \frac{1}{N-1} \sum_{k \neq i} f_{kj} \right) + \frac{m_{ct}}{2} (f_{ij} - f). \quad (\text{A6})$$

Noting that

$$\frac{1}{N-1} \sum_{k \neq i} f_{kj} = \frac{Nf_j}{N-1} - \frac{f_{ij}}{N-1}, \quad (\text{A7})$$

we find after rearrangement that the transmission function is given by

$$t_{ij} \simeq 1 + \frac{1}{2} \left[(f_{ij} - f) - (1 - m_{ct}) \left(\frac{Nf_j}{N-1} - \frac{f_{ij}}{N-1} - f \right) \right]. \quad (\text{A8})$$

Comparing this equation with equation (13), we see that the increase in local competition is slightly lower under “pairwise payoff” comparison than under “group payoff” comparison but remains much larger than under genetic transmission (cf. eq. [8], [A8]). The factor 1/2 in front of the square bracket will not affect the direction of selection, only the speed of evolution, and when group size becomes large ($N \rightarrow \infty$), the term in square brackets in equation (A8) converges to the second term of equation (13).

Selection Gradient

Darwinian Model

We derive here the gradient of selection on variant A (eq. [14]) when it affects the fitness function w . To that end, we express equation (6) as

$$1 + f_{ij} = 1 - Cp_{ij} + Bp_{-ij}, \quad (\text{A9})$$

where $p_{-ij} = [1/(N-1)] \sum_{k \neq i} p_{kj}$ designates the average frequency of variant A in group j but excluding individual i from the average. Averaging equation (A9) over all individuals within group j , we have

$$\begin{aligned} 1 + f_j &= \frac{1}{N} \sum_i (1 + f_{ij}), \\ &= 1 + (B - C)p_j, \end{aligned} \quad (\text{A10})$$

where $p_j = 1/N \sum_i p_{ij}$ is the average frequency of A in group j . Finally, the average of equation (A10) over all groups in the population is

$$\begin{aligned} 1 + f &= \frac{1}{n_d} \sum_j (1 + f_j), \\ &= 1 + (B - C)p, \end{aligned} \quad (\text{A11})$$

where p is the average frequency of A in the population and n_d is the number of demes ($n_d \rightarrow \infty$ due to our assumption of the infinite island model of dispersal). Substituting equations (A9)–(A11) into equation (8) yields to the first-order effect of selection:

$$w_{ij} = 1 - Cp_{ij} + Bp_{-ij} - (1 - m)^2(B - C)p_j - [1 - (1 - m)^2](B - C)p. \quad (\text{A12})$$

Inserting this fitness function into equation (4) and noting that the covariance of a sum can be decomposed into a sum of covariances and that the covariance between a constant and a random variable is 0, we obtain

$$\Delta p = -C \text{Cov}(p_{ij}, p_{ij}) + B \text{Cov}(p_{-ij}, p_{ij}) - (1 - m)^2 (B - C) \text{Cov}(p_j, p_{ij}) - [1 - (1 - m)^2] (B - C) \text{Cov}(p, p_{ij}). \quad (\text{A13})$$

Because of our assumption of the infinite island model of dispersal, the covariance $\text{Cov}(p, p_{ij})$ between the variant carried by individual i from group j and a homologous variant sampled in one individual chosen at random from the population is 0. Since $\text{Cov}(p_{ij}, p_{ij}) = p(1 - p)$ is the genetic variance in the population and the average frequency of variant A in group j can be written as

$$\begin{aligned} p_j &= \frac{1}{N} p_{ij} + \left(\frac{N-1}{N} \right) \frac{1}{N-1} \sum_{k \neq i} p_{kj}, \\ &= \frac{1}{N} p_{ij} + \left(\frac{N-1}{N} \right) p_{-ij}, \end{aligned} \quad (\text{A14})$$

we can factor out $p(1 - p)$ from equation (A13) and rearrange it to obtain

$$\Delta p = p(1 - p)[-C + BR - (1 - m)^2 (B - C)R^R], \quad (\text{A15})$$

where the term in square brackets is the gradient of selection on variant A ,

$$R = \frac{\text{Cov}(p_{-ij}, p_{ij})}{p(1 - p)} \quad (\text{A16})$$

is the regression of the cultural variant at the helping locus of group mates on the variant at the helping locus of a focal individual (i.e., cultural relatedness), and $R^R = [1 + (N - 1)R]/N$.

From the definition of the covariance between two random variables (i.e., $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$) and from equation (A16), the probability $E[p_{ij}p_{-ij}]$ that two homologous cultural variants sampled without replacement are both A can be written as

$$E[p_{ij}p_{-ij}] = Rp + (1 - R)p^2. \quad (\text{A17})$$

This equation suggests that R can also be interpreted as the probability that the lineages of two homologous variants sampled in two different individuals descend from the same recent common cultural ancestor (cumulative probability of coalescence in genetics; see Rousset 2002; Roze and Rousset 2003), in which case both individuals carry allele A with probability p . With complementary probability $1 - R$, the lineages of the two variants have not coalesced in the recent past (they do not descend from the same recent common ancestor), in which case the two lineages are considered as independent and the two individuals bear variant A with probability p^2 . This interpretation of the kinship coefficient R in terms of probabilities of identity between variants carried by interacting individuals is valid whenever the kinship associations reach their steady state before any significant change in allele frequency has occurred at the level of the population, which implies that selection must be weaker than the other process affecting the dynamics of kinship. This “quasi equilibrium” assumption is often used in population genetic theory (e.g., Nagylaki 1993; Bürger 2000; Kirkpatrick et al. 2002; Whitlock 2002; Cherry and Wakeley 2003; Roze and Rousset 2003; Rousset 2006) and allows us to conveniently evaluate the relatedness in terms of model parameters (see “Cultural Relatedness for the Darwinian Fitness Model”).

Group Payoff Comparison

We now carry out an analysis similar to that used in “Darwinian Model” in order to derive the selection gradient on variant A , when it affects the transmission function t (eq. [21]). Substituting equations (A9)–(A11) into equation (13) gives to the first order effect of selection

$$t_{ij} = 1 - Cp_{ij} + Bp_{-ij} - (1 - m_{ct})(B - C)p_j - m_{ct}(B - C)p. \quad (\text{A18})$$

Inserting this equation into equation (5) and using the results of “Darwinian Model,” we obtain under weak selection

$$\begin{aligned}\Delta p &= -C \text{Cov}(p_{ij}, p_{ij}) + B \text{Cov}(p_{-ij}, p_{ij}) - (1 - m_{ct})(B - C) \text{Cov}(p_j, p_{ij}), \\ &= p(1 - p)[-C + BR - (1 - m_{ct})(B - C)R^R].\end{aligned}\tag{A19}$$

Cultural Relatedness for the Darwinian Fitness Model

Here, we present the equilibrium value of the cultural relatedness coefficient R under unbiased horizontal cultural transmission for the Darwinian fitness model. To that end, we define two parameters: α_s denotes the probability that two individuals sampled from a group and bearing identical cultural variants before the stage of cultural transmission (stage 4 of the life cycle) bear an identical cultural variant after cultural transmission, and α_d denotes the probability that two individuals bearing different cultural variants before the stage of cultural transmission (stage 4 of the life cycle) bear the identical cultural variant after cultural transmission. With these two parameters, the probability that two individuals sampled in a group after cultural transmission bear the same cultural variant identical in state is given at steady state by

$$R = (1 - m)^2 \left\{ \frac{1}{N} \alpha_s + \left(\frac{N - 1}{N} \right) [R \alpha_s + (1 - R) \alpha_d] \right\} + [1 - (1 - m)^2] \alpha_d,\tag{A20}$$

where $(1 - m)^2$ is the probability that two individuals stem from the same group and $1/N$ is the probability that they descend from the same parent. This equation is formally similar to the classical recursions for identity in state under a model of recurrent mutation (Crow and Aoki 1984), and its solution is

$$R = \frac{N \alpha_d + (1 - m)^2 (\alpha_s - \alpha_d)}{N - (1 - m)^2 (N - 1) (\alpha_s - \alpha_d)}.\tag{A21}$$

Spread of Group Beneficial Norms

We derive here the gradient of selection on a group beneficial norm (variant A) by assuming exactly the same payoff structure as the one used by Boyd and Richerson (2002, p. 290). Using our notation, the payoff to individual i in group j (relative to the payoff of an individual that did not give or receive benefits from the expression of the social norm) is written as

$$1 + f_{ij} = 1 + s p_{ij} (p_j - p^*) + \tilde{B} p_j,\tag{A22}$$

where \tilde{B} is a public good benefit (shared equally by everybody in the group) resulting from the expression of the social norm and s is a benefit to individual i in group j expressing the social norm, whenever the average frequency p_j of the norm in group j exceeds the threshold frequency p^* ; otherwise, it is costly to express the social norm. Averaging the payoff function over all individuals within group j , we have

$$\begin{aligned}1 + f_j &= \frac{1}{N} \sum_i (1 + f_{ij}), \\ &= 1 + s p_j (p_j - p^*) + \tilde{B} p_j,\end{aligned}\tag{A23}$$

and the average payoff in the whole population is

$$\begin{aligned}
 1 + f &= \frac{1}{n_d N} \sum_j \sum_i (1 + f_{ij}), \\
 &= \frac{1}{n_d N} \sum_j \sum_i (1 + s p_{ij} p_j - s p_{ij} p^* + \tilde{B} p_j), \\
 &= 1 + s E[p_j^2] - s p p^* + \tilde{B} p.
 \end{aligned} \tag{A24}$$

With these payoffs functions, we can evaluate the gradient of selection on the social norm under both genetic inheritance and cultural inheritance and, when cultural transmission follows the “Darwinian” model, the “pairwise payoff” and “group payoff” comparison models. We will consider here only the “group payoff” comparison model because it is qualitative similar but simpler than the “pairwise payoff” comparison model, which was originally considered by Boyd and Richerson (2002).

Under the “group payoff” comparison model, the transmission function (eq. [13]) for the social norm is obtained by using the payoff functions (eqq. [A21]–[A23]), which gives to first order

$$t_{ij} = 1 + s(p_{ij} p_j - p_{ij} p^*) + \tilde{B} p_j - (1 - m_{ct})[s(p_j^2 - p_j p^*) + \tilde{B} p_j] - m_{ct}\{s(E[p_j^2] - p p^*) + \tilde{B} p\}. \tag{A25}$$

Substituting equation (A25) into equation (5) and using $\text{Cov}(p, p_{ij}) = 0$ and $\text{Cov}(E[p_j^2], p_{ij}) = 0$, we obtain

$$\begin{aligned}
 \Delta p &= s[\text{Cov}(p_{ij} p_j, p_{ij}) - p^* \text{Cov}(p_{ij}, p_{ij})] + \tilde{B} \text{Cov}(p_j, p_{ij}) \\
 &\quad - (1 - m_{ct})\{s[\text{Cov}(p_j^2, p_{ij}) - p^* \text{Cov}(p_j, p_{ij})] + \tilde{B} \text{Cov}(p_j, p_{ij})\}.
 \end{aligned} \tag{A26}$$

Since p_{ij} is an indicator variable equal to 1 or 0, we have $p_{ij} = p_{ij}^2$, so that $E[p_j p_{ij} p_{ij}] = E[p_j p_{ij}]$. Then, from $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$, we have

$$\begin{aligned}
 \text{Cov}(p_j p_{ij}, p_{ij}) &= E[p_j p_{ij} p_{ij}] - E[p_j p_{ij}]E[p_{ij}], \\
 &= E[p_j p_{ij}](1 - p),
 \end{aligned} \tag{A27}$$

because $E[p_{ij}] = p$. The probability $E[p_j p_{ij}]$ that two homologous cultural variants randomly sampled with replacement from the same group are both A is equal to $E[p_j^2]$ ($E[p_j^2] = \sum_j p_j^2/n_d$ and $E[p_j p_{ij}] = \sum_{i,j} p_j p_{ij}/(n_d N)$) and can be written in the infinite island model as

$$E[p_j^2] = R^R p + (1 - R^R) p^2. \tag{A28}$$

Using this equation, we have

$$\text{Cov}(p_j p_{ij}, p_{ij}) = [R^R + (1 - R^R) p] p (1 - p). \tag{A29}$$

The definition of the covariance also gives

$$\text{Cov}(p_j^2, p_{ij}) = E[p_j^2 p_{ij}] - E[p_j^2]E[p_{ij}], \tag{A30}$$

where $E[p_j^2 p_{ij}]$ is the probability that three homologous cultural variants randomly sampled with replacement from the same group are all of type A ($E[p_j^2 p_{ij}] = E[p_j^3]$). This is

$$E[p_j^3] = R_3^R p + R^R p^2 + (1 - R^R - R_3^R) p^3, \tag{A31}$$

where R_3^R can be interpreted as the probability that the lineages of three homologous variants sampled in three different individuals all descend from the same recent cultural ancestor (see Roze and Rousset 2003). Combining the last two expressions, we obtain

$$\frac{\text{Cov}(p_i^2, p_{ij})}{p(1-p)} = (1+p)R_3^R. \quad (\text{A32})$$

Taking into account all terms, the change in frequency of the social norm under weak selection is

$$\Delta p = p(1-p)\{s[R^R + (1-R^R)p - p^*] + m_{ct}\tilde{B}R^R - (1-m_{ct})s[(1+p)R_3^R - p^*R^R]\}, \quad (\text{A33})$$

where the gradient of selection is now frequency dependent. In order to obtain an expression for Δp in closed form, it now remains to evaluate the probabilities of identity R^R and R_3^R . The probability R^R is obtained by substituting the solution of equation (25) into equation (15), while the recursion for R_3^R can be written down (e.g., Roze and Rousset 2003), but its solution is a bit complicated. In order to get around this problem, we now assume large group size. In that case, we can neglect R_3^R because $R_3^R = O(1/N^2)$, where $O(1/N^2)$ is a remainder of order $1/N^2$, which represents the probability that two individuals have copied a cultural variant from the same individual and gives the rate at which R_3^R builds up per generation. To the first order in group size (neglecting all terms of order $1/N^2$ and of higher order), the change in variant frequency is given by

$$\Delta p = p(1-p)\{s[R^R + (1-R^R)p - p^*] + [m_{ct}\tilde{B} + (1-m_{ct})sp^*]R^R\}. \quad (\text{A34})$$

With this assumption of large group size, the social norm can invade the population when rare ($p \rightarrow 0$) if the threshold frequency p^* satisfies the inequality

$$p^* < \frac{R^R(s + m_{ct}\tilde{B})}{[1 - (1 - m_{ct})R^R]s}, \quad (\text{A35})$$

where $R^R = 1/[Nm_{ct}(2 - m_{ct})]$, which is obtained by substituting equation (25) into equation (15) and assuming large group size. If $\tilde{B} = 0$, the inequality simplifies to $p^* < R^R/[1 - (1 - m_{ct})R^R]$.