Supplementary Equations

1a. Evolutionary dynamics of *Shunner*, *Cooperator* and *Defector* with random group formation

We begin by writing down the payoff for each strategy, *Cooperator* (*C*), *Defector* (*D*) and *Shunner* (*S*), for a given group composition. This payoff is a sum of the payoff during the collective action stage and the payoff during the mutual aid stage. Before proceeding, we note that an individual's standing entering a particular period of the mutual aid game will influence his or her payoff in that period. It can be shown that the frequency of individuals in good standing for each strategy rapidly approaches an equilibrium value. Thus, we can approximate the expected payoff during the mutual aid game for a particular strategy by adding the payoff in the first period of the mutual aid game to the payoff in round n, by which time the frequency of individuals in good standing for that strategy has reached equilibrium, weighted by the term $\dfrac{w}{1-w}$, which represents all subsequent periods. Numerical simulation suggests that such an approximation has little effect on the model results.[1]

Returning to the payoff computations, we denote the numbers of individuals using strategies *C*, *D* and *S* amongst the $n-1$ other group members by x_1, x_2 and x_3, respectively. We denote the equilibrium proportion of individuals of strategy i in good standing and payoff to an individual playing strategy i with $G(i)$ and $V(i)$, respectively.

$$V(C) = \left\{B\frac{x_1+x_3}{n-1} - C\right\} + \left\{(1-e)\left[\frac{1}{n}b(x_1+x_3) - \frac{n-1}{n}c\right]\right\}$$
$$+ \frac{w}{1-w}\left\{(1-e)\left[\frac{1}{n}b(x_1+x_3 G(C)) - \frac{n-1}{n}c\right]\right\}$$

$$V(D) = \left\{B\frac{x_1+x_3}{n-1}\right\} + \left\{(1-e)\frac{1}{n}bx_1\right\} + \frac{w}{1-w}\left\{(1-e)\frac{1}{n}bx_1\right\}$$

$$V(S) = \left\{B\frac{x_1+x_3}{n-1} - C\right\} + \left\{(1-e)\left[\frac{1}{n}b(x_1+x_3) - \frac{n-1}{n}c\frac{x_1+x_3}{n-1}\right]\right\}$$
$$+ \frac{w}{1-w}\left\{(1-e)\left[\frac{1}{n}b(x_1+x_3 G(S)) - \frac{n-1}{n}c\left(\frac{x_1}{n-1}G(C) + \frac{x_3}{n-1}G(S)\right)\right]\right\}$$

We next derive the equilibrium fractions of *C* and *S* in good standing during the mutual aid stage for a given group composition. As the strategy *D* fails to donate during the

---

[1] For further discussion, see Panchanathan and Boyd (2003), whose indirect reciprocity game has similar dynamics to the mutual aid game modeled in this paper.

collective action stage, it immediately falls into bad standing and never recovers. We denote the fraction of strategy i in good standing during round n by $g_n(i)$. The term g_n denotes the total fraction of a social group in good standing during round n.

$$g_n(C) = g_{n-1}(C)\left[\frac{1}{n} + \frac{n-1}{n}(1 - g_{n-1}e)\right] + (1 - g_{n-1}(C))\frac{n-1}{n}(1-e)$$

$$g_n(S) = g_{n-1}(S)\left[\frac{1}{n} + \frac{n-1}{n}(1 - g_{n-1}e)\right] + (1 - g_{n-1}(S))\frac{n-1}{n}g_n(1-e)$$

To derive $G(C)$ and $G(S)$, we assume that e is small such that $e^2 \approx 0$. Further, as simulation results indicate that $G(C) \approx 1$, we can assume that $G(C)^2 \approx 2G(C) - 1$. At equilibrium, we have $G(C) \to 1 - e\left[\frac{x_1 + x_3}{n-1}\right]$ and $G(S) \to 1 - e$.

The expected fitness of a strategy is the payoff off that strategy in a given group averaged over all possible group compositions. If we let p, q and $1 - p - q$ denote the frequencies of S, C and D, respectively, in the global population, then, from the properties of the multinomial distribution, the expected values of x_1, x_2 and x_3 are $q(n-1)$, $(1 - p - q)(n-1)$ and $p(n-1)$, respectively. Substituting these expectations, we find the expected fitness for each type is

$$W(C) = W_0 + B(q + p) - C + \frac{1-e}{1-w}\frac{n-1}{n}\left[b(q + p(1 - we(p+q))) - c\right]$$

$$W(D) = W_0 + B(q + p) + \frac{1-e}{1-w}\frac{n-1}{n}bq$$

$$W(S) = W_0 + B(q + p) - C + \frac{1-e}{1-w}\frac{n-1}{n}\left[b(q + p(1 - we)) - c(q(1 - we(p+q)) + p(1 - we))\right]$$

where W_0 represents baseline fitness. When *Shunner* is common, the fitness of all three strategies will be dominated by groups in which the other $n - 1$ individuals are *Shunner*. Thus, to determine when *Shunner* can resist invasion by rare *Defectors*, we set $q = 0$ and $p \approx 1$. We then find that $W(S) > W(D)$ when

$$\frac{1-e}{1-w}\frac{n-1}{n}(b - c)(1 - we) > C$$

This conditions tells us that *Shunner* is evolutionarily stable against *Defector* when the long run benefits of reciprocity through mutual aid (the expression on the left of the inequality) is greater than the short-term benefit of free riding the collective action (the expression on the right).

Next, to find the stability criteria for *Shunner* against *Cooperator*, we begin by setting $1-p-q=0$ and $p \approx 1$. We then find the condition under which $W(S) > W(C)$ to be

$$cwe > 0$$

This tells us that so long mutual aid (1) is costly, (2) persists for more than one period, and (3) is prone to error, then *Shunner* is an ESS against *Cooperator*.

1b. Invasion analysis of *Shunner* against *Defector* assuming non-random group formation

Setting the frequency of Cooperator (q) to zero, we model this condition by assuming that the conditional probability of a type j donor matched with a type i recipient during the mutual aid game, $\Pr(i \mid j)$, is:

$$\Pr(S \mid S) = r + (1-r)p$$
$$\Pr(D \mid S) = (1-r)(1-p)$$
$$\Pr(S \mid D) = (1-r)p$$
$$\Pr(D \mid D) = r + (1-r)(1-p)$$

The parameter r is the coefficient of relatedness in a haploid sexual genetic model and, more generally, measures the degree of assortment. It then follows that the expected number of type j individuals amongst the other $n-1$ group members in a group containing one individual of type i, which we denote by $e(x_j \mid i)$, is:

$$e(x_3 \mid S) = (r + (1-r)p)(n-1)$$
$$e(x_2 \mid S) = (1-r)(1-p)(n-1)$$
$$e(x_3 \mid D) = (1-r)p(n-1)$$
$$e(x_2 \mid D) = (r + (1-r)(1-p))(n-1)$$

Substituting these expectations we have the following fitness functions:

$$W(D) = W_0 + B(1-r)p$$
$$W(S) = W_0 + B(r + (1-r)p) - C + \frac{1-e}{1-w}\frac{n-1}{n}\left[(b-c)(1-we)(r+(1-r)p)\right]$$

As *Shunner* and *Defector* are evolutionarily stable against invasion by each other, we can find the unstable equilibrium by setting the two fitness functions equal to each other. The expression for the unstable equilibrium, plotted in Figure 3a of the text, is given by

$$\hat{p} = \frac{C - r\left[B + \dfrac{1-e}{1-w}\dfrac{n-1}{n}(b-c)(1-we)\right]}{\dfrac{1-w}{1-e}\dfrac{n-1}{n}(b-c)(1-we)(1-r)}$$

To determine threshold amount of assortment (r^*) necessary for *Shunner* to invade against *Defector*, we set $p \approx 0$ and find the conditions under which $W(S) > W(D)$, solving for r. ,

$$r*\left[B + \frac{1-e}{1-w}\frac{n-1}{n}(b-c)(1-we)\right] > C$$

This critical threshold of relatedness is plotted in Figure 2a of the text.

To get a better sense of what this expression means, we can simplify by assuming that (1) there are no errors to mutual aid contribution (i.e., $e = 0$) and (2) groups are large (i.e., $\dfrac{n-1}{n} \approx 1$). These assumptions do not alter the results qualitatively. We now have

$$r*\left[B + \frac{b-c}{1-w}\right] > C$$

This expression can be rewritten as

$$r*\left[\frac{B(1-w)+(b-c)}{1-w}\right] > C$$

This condition bears striking similarity to the condition for the invasion of *Tit-for-tat* against *Defector* in the model of reciprocal altruism studied by Axelrod and Hamilton (1981). This is not surprising as the mutual aid game we study here, Axelrod and Hamilton's model of reciprocal altruism, and models of indirect reciprocity (Nowak and Sigmund 1998; Panchanathan and Boyd 2003) are all systems of reciprocity and thus have similar dynamical properties. Thus, as with other models of reciprocity, relatedness and reciprocity operate synergistically, such that an increase in w results in a lowered requisite value of r^* for the evolution of cooperation. Conversely, an increase in assortment (r^*) leads to a decrease in the requisite number of reciprocity bouts (w). This relationship can be shown by taking the derivative of r^* with respect to w. When we do this, we find that

$$\frac{\partial r*}{\partial w} < 0$$

2. Evolutionary dynamics of *Shunner* and *Reciprocator* with non-random group formation

In this model, we consider two strategies: *Shunner* (S) and *Reciprocator* (R). In this model, R types do not contribute to the collective action and do not attend to whether others do while S types contribute and attend to others' contributions. We keep track of the number of S and R amongst the $n-1$ other group members with x_3 and x_4, respectively. And we keep track of the global frequency S and R with p and $1-p$, respectively. Using the binomial theorem and the previously-defined assortment parameter (r), we have:

$$
\begin{aligned}
\mathrm{Pr}_1 &= \mathrm{Pr}(S\mid S) = r + (1-r)p \\
\mathrm{Pr}_2 &= \mathrm{Pr}(R\mid S) = (1-r)(1-p) \\
\mathrm{Pr}_3 &= \mathrm{Pr}(S\mid R) = (1-r)p \\
\mathrm{Pr}_4 &= \mathrm{Pr}(R\mid R) = r + (1-r)(1-p)
\end{aligned}
$$

and

$$
\begin{aligned}
e_1 &= e(x_3\mid S) = (r+(1-r)\,p)(n-1) \\
e_2 &= e(x_4\mid S) = (1-r)(1-p)(n-1) \\
e_3 &= e(x_3\mid R) = (1-r)p\,(n-1) \\
e_4 &= e(x_4\mid R) = (r+(1-r)(1-p))(n-1)
\end{aligned}
$$

To determine the equilibrium fraction of individuals in good standing for each type, we note that while both S and R attend to standing, they each have different rules for its assignment. As such, we let $g_n(i\mid j)$ be the fraction of type i in good standing in round n as seen by type j. As *Reciprocator* does not participate in the collective action, $g_n(R\mid S)=0$ for all rounds after stage one, leaving us with the following recursions:

$$
g_n(S\mid S) = \frac{1}{n}g_{n-1}(S\mid S) + \frac{n-1}{n}\left[
\begin{array}{l}
g_{n-1}(S\mid S)(\mathrm{Pr}_1(1-g_{n-1}(S\mid S)e) + \mathrm{Pr}_2(1-g_{n-1}(R\mid S)e)) \\
+ (1-g_{n-1}(S\mid S))(\mathrm{Pr}_1\, g_{n-1}(S\mid S)(1-e) + \mathrm{Pr}_2\, g_{n-1}(R\mid S)(1-e))
\end{array}\right]
$$

$$
g_n(R\mid R) = \frac{1}{n}g_{n-1}(R\mid R) + \frac{n-1}{n}\left[
\begin{array}{l}
g_{n-1}(R\mid R)(\mathrm{Pr}_3(1-g_{n-1}(S\mid R)e) + \mathrm{Pr}_4(1-g_{n-1}(R\mid R)e)) \\
+ (1-g_{n-1}(R\mid R))(\mathrm{Pr}_3\, g_{n-1}(S\mid R)(1-e) + \mathrm{Pr}_4\, g_{n-1}(R\mid R)(1-e))
\end{array}\right]
$$

$$
g_n(S\mid R) = \frac{1}{n}g_{n-1}(S\mid R) + \frac{n-1}{n}\left[
\begin{array}{l}
\mathrm{Pr}_1\left\{
\begin{array}{l}
g_{n-1}(S\mid R)\big[g_{n-1}(S\mid R)\,g_{n-1}(S\mid S)(1-e) + (1-g_{n-1}(S\mid R))\big] \\
+ (1-g_{n-1}(S\mid R))g_{n-1}(S\mid S)(1-e)
\end{array}\right\} \\
+ \mathrm{Pr}_2\, g_{n-1}(S\mid R)(1-g_{n-1}(R\mid R))
\end{array}\right]
$$

At equilibrium, we have

$$G(R \mid S) \to 0$$
$$G(S \mid S) \to 1-e$$
$$G(R \mid R) \to 1-e$$
$$G(S \mid R) \to \mathrm{Pr}_1(1-2e)$$

Using a similar method outlined for model one, we derive expected fitness functions for each strategy.

$$W(S) = W_0 + B\big[r + (1-r)p\big] - C + \frac{1-e}{1-w}\frac{n-1}{n}\Big[b\{1-w+w\mathrm{Pr}_1[(1-e)+\mathrm{Pr}_2(1-2e)]\} - c\,\mathrm{Pr}_1(1-we)\Big]$$

$$W(R) = W_0 + B\big[(1-r)p\big] + \frac{1-e}{1-w}\frac{n-1}{n}\Big[b\,\mathrm{Pr}_4(1-we) - c\{1-w+w[\mathrm{Pr}_3\,\mathrm{Pr}_1(1-2e)+\mathrm{Pr}_4(1-e)]\}\Big]$$

By setting these fitness functions equal to each other, we can find any internal equilibrium points. The result is a quadratic function given by

$$\hat{p}^2\,\frac{1-e}{1-w}\frac{n-1}{n}w(1-r)^2(1-2e)(b-c)$$

$$-\hat{p}\,\frac{1-e}{1-w}\frac{n-1}{n}(1-r)\big[b\{1+2w(1-2e)(1-r)\}-c\{1+w(1-2e)(1-r)\}\big]$$

$$+C - Br + \frac{1-e}{1-w}\frac{n-1}{n}\big[bw\{r^2(1-2e)-r(2-3e)+(1-e)\}-c(1-r)(1-we)\big] = 0$$

The negative root of this expression lies between [0, 1] and is an unstable equilibrium. It is graphed in Figure 3b of the text.

To determine the stability criteria for *Reciprocator*, we set $p \approx 0$ and $r = 0$ and then solve for $W(R) > W(S)$, which yields the condition

$$\frac{1-e}{1-w}\frac{n-1}{n}\big[bw(1-e)-c(1-we)\big] + C > 0$$

We next investigate the conditions under which *Shunner* can invade a population playing *Reciprocator*. We set $p \approx 0$ and find the conditions under which $W(S) > W(R)$. This results in a critical minimum threshold of relatedness (r^*) for *Shunner* to invade against *Reciprocator*, which is given by

$$(r^*)^2\left[\frac{1-e}{1-w}\frac{n-1}{n}bw(1-2e)\right] - r^*\left[B + \frac{1-e}{1-w}\frac{n-1}{n}\{bw(2-3e)-c(1-we)\}\right]$$

$$+\left[C + \frac{1-e}{1-w}\frac{n-1}{n}\{bw(1-e)-c(1-we)\}\right] < 0$$

This critical threshold of relatedness is plotted in Figure 2b of the text.

As we are interested in invasion conditions for which $r*$ is small, we can assume that $(r*)^2 \approx 0$. We now have

$$r*\left[B + \frac{1-e}{1-w}\frac{n-1}{n}\{bw(2-3e)-c(1-we)\}\right] - \left[C + \frac{1-e}{1-w}\frac{n-1}{n}\{bw(1-e)-c(1-we)\}\right] > 0$$

As in the previous section, we set $e = 0$ and $\frac{n-1}{n} \approx 1$, which leaves us with the condition

$$r*\left[B + \frac{2bw-c}{1-w}\right] > \left[C + \frac{bw-c}{1-w}\right]$$

From this inequality, it is not obvious how assortment $(r*)$ and reciprocity (w) interact. Thus, we take the derivative of $r*$ with respect to w, which yields

$$\frac{\partial r*}{\partial w} = \frac{B(b-c) - C(2b-c) + bc}{[B(1-w) + 2bw - c]^2}$$

As the denominator is positive, we can write

$$\frac{\partial r*}{\partial w} \propto B(b-c) - C(2b-c) + bc$$

It is not clear whether this expression is positive or negative. Let us, therefore, momentarily assume that there is no cost to contributing to the collective action. This seems a safe assumption as it represents the most conducive condition for the invasion of the *Shunner* strategy. In a moment, we will relax this assumption. By setting $C = 0$, we now have

$$\frac{\partial r*}{\partial w} \propto B(b-c) + bc$$

As this expression is positive, we observe that relatedness and reciprocity operate antagonistically. This result is in stark contrast to the one derived in section **1b** and that of Axelrod and Hamilton's model of reciprocal altruism. If the *Shunner* type is to invade, it can only be driven by inclusive benefits from the collective action. As the importance of reciprocity (w) increases, the requisite assortment level $(r*)$ increases.

How safe is it to analyze the relationship between assortment $(r*)$ and reciprocity (w) when the collective action is costless ($C = 0$)? Above we found that

$$\frac{\partial r *}{\partial w} \propto B(b-c) - C(2b-c) + bc$$

When this derivative is positive ($\frac{\partial r *}{\partial w} > 0$), reciprocity and inclusive fitness operate antagonistically. This derivative will be positive when the collective action benefit (B) is sufficiently high. Precisely,

$$B > \frac{C(2b-c) + bc}{b-c}$$

Under most model parameter values, $\frac{\partial r *}{\partial w} > 0$ when $B > 3C$ (Figure 4).
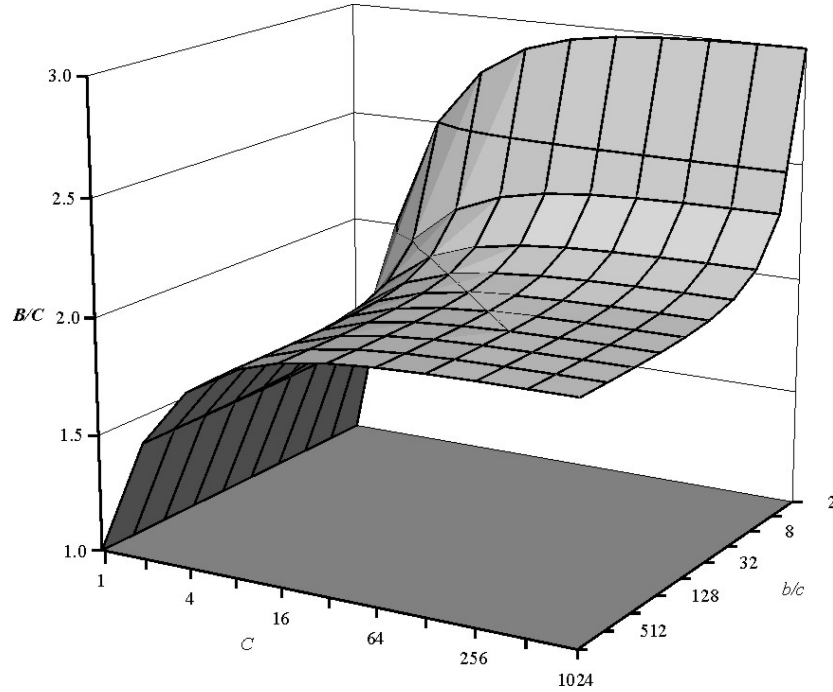

Figure 4. The minimum threshold of collective action benefit ($B\!/\!C$) necessary for Relatedness and reciprocity to operate antagonistically ($\frac{\partial r *}{\partial w} > 0$) as a function collective action cost (C) and cost benefit ratio of mutual aid ($b\!/\!c$). The cost of cooperating during one bout of mutual aid (c) is set to 1.

Surprisingly, this says that when the collective action is sufficiently group beneficial ($B > 3C$), increasing the strength of reciprocity (w) increases the requisite relatedness threshold for *Shunners* to invade against *Reciprocity* (r^*). Why would a less group beneficial collective action make it easier for *Shunners* to invade? From Figure 4, we see that $\dfrac{\partial r^*}{\partial w} < 0$, when $B < C$. In Figure 5, we plot the threshold relatedness (r^*) when $B = C$.
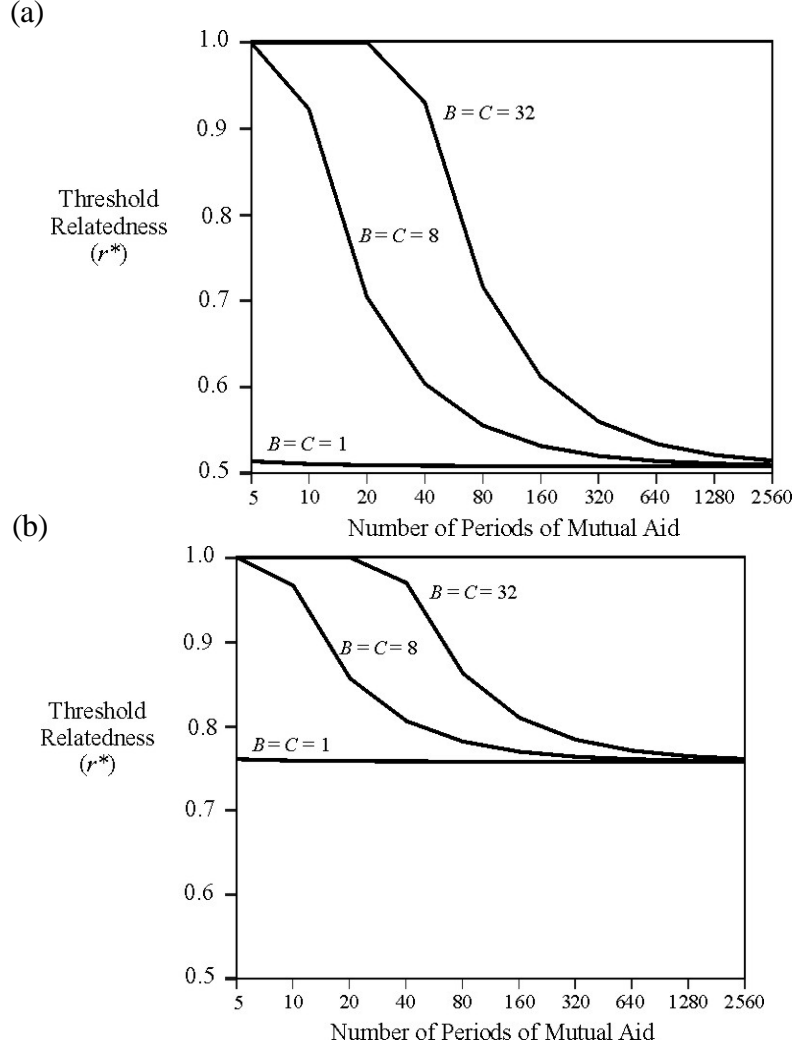
(a)

(b)

Figure 5. The threshold degree of assortment (r^*) necessary for rare *Shunners* to invade a population in which *Reciprocators* are common as a function of the number of mutual aid periods ($\dfrac{1}{1-w}$) for different costs and benefits of collective action (B, C). In (**a**), $b = 2$ and in (**b**), $b = 4$. In both (**a**) and (**b**), $c = 1$, $n = 30$ and $e = 0.01$.

From this figure we see that the threshold relatedness does in fact decrease as the strength of reciprocity increases ($\frac{\partial r*}{\partial w} < 0$). This implies that relatedness and reciprocity operate synergistically (*i.e.*, Axelrod and Hamilton's "ratchet"). Notice, however, that the threshold relatedness in Figure 5 approaches an asymptote which never falls below 0.5. If we are interested in a model of human social behavior, these regions of the parameter space are biologically uninteresting. Thus, we can say that under relevant parameter values our assertion that $\frac{\partial r*}{\partial w} > 0$ seems to hold.

3. Evolutionary dynamics of two types of *Shunner* with non-random group formation

In this section, which was not presented in the text, we investigate how social norms change in our social model. Thus, we have two strategies: *Shunner Type 1* (*S1*) and *Shunner Type 2* (*S2*). *S1* contributes to collective action 1 at a private net cost of C_1, which produces a benefit B_1 which is shared amongst the $n-1$ other group members. This strategy assigns a lifetime of bad standing to those that do not contribute to this collective action. *S2* contributes to collective action 2 with associated costs and benefits C_2 and B_2, respectively. *S2* assigns a lifetime of bad standing to those that do not contribute to this collective action. We keep track of the number of *S1* and *S2* amongst the $n-1$ other group members with $n-1-x$ and x, respectively. And we keep track of the global frequency *S1* and *S2* with $1-p$ and p, respectively. Using the binomial theorem and the previously-defined assortment parameter (r), we have:

$$\begin{aligned}
\Pr_1 &= \Pr(S2 \mid S2) = r + (1-r)p \\
\Pr_2 &= \Pr(S1 \mid S2) = (1-r)(1-p) \\
\Pr_3 &= \Pr(S2 \mid S1) = (1-r)p \\
\Pr_4 &= \Pr(S1 \mid S1) = r + (1-r)(1-p)
\end{aligned}$$

and

$$\begin{aligned}
e_1 &= e(x \mid S2) = (r + (1-r)p)(n-1) \\
e_2 &= e(n-1-x \mid S2) = (1-r)(1-p)(n-1) \\
e_3 &= e(x \mid S1) = (1-r)p(n-1) \\
e_4 &= e(n-1-x \mid S1) = (r + (1-r)(1-p))(n-1)
\end{aligned}$$

As in the previous section, *S1* and *S2* both attend to standing but have different assignment rules. As before, we let $g_n(i \mid j)$ be the fraction of type i in good standing in round n as seen by type j. As *S1* and *S2* do not contribute to the collective action which

the other strategy expects, we have $g_n(S1 \mid S2) = 0$ and $g_n(S2 \mid S1) = 0$ for all rounds after stage one, leaving us with the following two recursions:

$$g_n(S1 \mid S1) = \frac{1}{n} g_{n-1}(S1 \mid S1) + \frac{n-1}{n} \left[\begin{array}{l} g_{n-1}(S1 \mid S1)(\Pr_4(1 - g_{n-1}(S1 \mid S1)e) + \Pr_3(1 - g_{n-1}(S2 \mid S1)e)) \\ + (1 - g_{n-1}(S1 \mid S1))(\Pr_4 g_{n-1}(S1 \mid S1)(1-e) + \Pr_3 g_{n-1}(S2 \mid S1)(1-e)) \end{array} \right]$$

$$g_n(S2 \mid S2) = \frac{1}{n} g_{n-1}(S2 \mid S2) + \frac{n-1}{n} \left[\begin{array}{l} g_{n-1}(S2 \mid S2)(\Pr_1(1 - g_{n-1}(S2 \mid S2)e) + \Pr_2(1 - g_{n-1}(S1 \mid S2)e)) \\ + (1 - g_{n-1}(S2 \mid S2))(\Pr_1 g_{n-1}(S2 \mid S2)(1-e) + \Pr_2 g_{n-1}(S1 \mid S2)(1-e)) \end{array} \right]$$

At equilibrium, we have

$$G(S1 \mid S1) \rightarrow 1-e$$
$$G(S2 \mid S1) \rightarrow 0$$
$$G(S1 \mid S2) \rightarrow 0$$
$$G(S2 \mid S2) \rightarrow 1-e$$

Using a similar method outlined for model one, we derive expected fitness functions for each strategy.

$$W(S1) = W_0 + B_1 \left[r + (1-r)(1-p) \right] - C_1 + B_2 \left[(1-r)p \right] + \frac{1-e}{1-w} \frac{n-1}{n} (b-c)(1-we) \left[r + (1-r)(1-p) \right]$$

$$W(S2) = W_0 + B_2 \left[r + (1-r)p \right] - C_2 + B_1 \left[(1-r)(1-p) \right] + \frac{1-e}{1-w} \frac{n-1}{n} (b-c)(1-we) \left[r + (1-r)p \right]$$

To determine the stability criteria for *S1*, we set $p \approx 0$ and $r = 0$ and then solve for $W(S1) > W(S2)$, which yields the condition

$$\frac{1-e}{1-w} \frac{n-1}{n} (b-c)(1-we) + C_2 > C_1$$

We next investigate the conditions under which *S2* can invade a population playing *S1*. We set $p \approx 0$ and find the condition under which $W(S2) > W(S1)$, which results in a critical minimum threshold of relatedness (*r**) for *S2* to invade against *S1*.

$$r * \left[B_2 - B_1 + \frac{1-e}{1-w} \frac{n-1}{n} (b-c)(1-we) \right] > \left[C_2 - C_1 + \frac{1-e}{1-w} \frac{n-1}{n} (b-c)(1-we) \right]$$

As in the previous section, we set $e = 0$ and $\frac{n-1}{n} \approx 1$, which leaves us with the condition

$$r * \left[B_2 - B_1 + \frac{b-c}{1-w} \right] > \left[C_2 - C_1 + \frac{b-c}{1-w} \right]$$

From this inequality, it is not obvious how assortment and reciprocity interact. Thus, we take the derivative of $r*$ with respect to w, which yields

$$\frac{\partial r *}{\partial w} = \frac{(b-c)\left[(B_2 - C_2) - (B_1 - C_1)\right]}{\left[(B_2 - B_1)(1-w) - (b-c)\right]^2}$$

As the denominator is positive, we can write

$$\frac{\partial r *}{\partial w} \propto (b-c)\left[(B_2 - C_2) - (B_1 - C_1)\right]$$

This says that relatedness and reciprocity operate antagonistically when the new social norm increases mean fitness compared to the old social norms (i.e., $B_2 - C_2 > B_1 - C_1$). That is, under such conditions, the invasion of *S2* is only determined by the inclusive benefits from the collective action. As the effect of reciprocity increases (i.e., w increases), the requisite assortment (i.e., $r*$) increases.

Note that above condition also implies that relatedness and reciprocity operate synergistically, which occurs when the new social norm decreases mean fitness (i.e., $B_1 - C_1 > B_2 - C_2$). Simulation results indicate, however, that this result is misleading; the synergy between assortment and reciprocity only occurs under conditions not otherwise conducive to the evolution of cooperation.

REFERENCES CITED

Axelrod, R. and W. D. Hamilton (1981). "The evolution of cooperation." Science **211**: 1390-1396.
Nowak, M. A. and K. Sigmund (1998). "The dynamics of indirect reciprocity." Journal of Theoretical Biology **194**(4): 561-574.
Panchanathan, K. and R. Boyd (2003). "A tale of two defectors: the importance of standing for the evolution of indirect reciprocity." Journal of Theoretical Biology **224**: 115-126.