Nordic Journal of Political Economy

Volume 23 1996 Pages 145-149

On the Evolution of Altruism

Oded Stark

Oded Stark *

On the Evolution of Altruism

An example is provided to illustrate how evolution can select for altruism. It is shown that evolution can sustain altruistic behavior between relatives even in a single-shot prisoner's dilemma model in which altruism benefits one's opponent at a cost to oneself, and conditions are derived under which altruism persists and flourishes to the extent that the entire population will consist of altruists. The case presented is of interest also because it illustrates how the distribution of a *population* by a trait is an outcome solely of the relative payoff to the trait in *intrafamilial* exchanges.

The game and the payoffs

Consider the following two-player, two-strategy game in which a player who cooperates gets a payoff of R if his opponent cooperates, and S if the opponent defects. A player who defects gets T if his opponent cooperates, and P if the opponent defects. In a prisoner's dilemma game, $S<P<R<T$, so that defection is a dominant strategy for each player.

We equate altruism with cooperating in a prisoner's dilemma game. To see this suppose the column player selects C.

		Column Player	
		C	D
Row Player	C	R,R	S,T
	D	T,S	P,P

If the row player selects C rather than D, he gives up T to receive the smaller R, whereas the column player gains since he receives R, which is larger than S. Suppose, alternatively, that the column player selects D. Again, if the row player selects C rather than D, his payoff declines (by $P-S$), while the column player's payoff rises (by $T-P$). This is what altruism is about: giving up something for the sake of another. Thus, throughout the rest of this

* University of Oslo and University of Vienna.
This paper evolves from ongoing work with Theodore C. Bergstrom. I am grateful to Lyudmila Belova and Tone Ognedal for enlightening discussions and helpful advice, and to Aanund Hylland for extremely constructive comments. I would also like to thank participants in the Conference on Ethical Values and Value of Ethics, the Norwegian Academy of Sciences; in a seminar at the Center for Rationality and Interactive Decision Theory at the Hebrew University of Jerusalem, and in the workshop on Intra-Household and Intra-Family Interactions at the University of Copenhagen for thought-provoking questions. I thank the National Institute on Aging (grant R01-AG13037) for partial financial support.

paper we identify altruism with playing cooperate in the one-shot prisoner's dilemma game.[1]

The rule of imitation

An individual's strategy, to play *C* or *D* against one's sibling, is determined by imitating the behavior of parents or nonparents. Note that strategy here stands for a programmed pattern of behavior, not an object of choice. Assume that with probability v a child randomly selects one parent as a role model and adopts that parent's strategy. With probability $1-v$ the child chooses a random nonparent as a role model. Each individual has a sibling with whom the individual plays a game of prisoner's dilemma. The probability that an individual survives to reproduce is proportional to the payoff in this game. For example, consider a case in which the payoff positively influences the probability of reaching maturity and of being able to procreate.

The formation of couples

Assume that mating is monogamous. Parent-couples can be one of three possible types: two-cooperator couples, "mixed couples" with one cooperator and one defector, and two-defector couples. Let x be the fraction of cooperators in the adult population. If marriage is purely random, the fraction of marriages with two cooperators is $x \cdot x = x^2$, the fraction with two defectors is $(1-x)(1-x) = (1-x)^2$, and the fraction with mixed couples is $2x(1-x)$. If marriage is purely (positively) assortative, the fractions of cooperators and defectors are, respectively, x and $1-x$. To allow mating patterns that are intermediate between the polar cases of purely random mating and purely assortative mating, we define a parameter m where $0 \leq m \leq 1$, such that when mating is purely random $m=0$, and when mating is purely assortative $m=1$. In the population at large, the proportion of two-cooperator couples is thus $x^2 + mx(1-x)$; the proportion of two-defector couples is $(1-x)^2 + mx(1-x)$; and the remaining proportion of mixed couples is $2(1-m)x(1-x)$.[2]

---

1. Consider the following quite general formulation: $U(C_1, C_2) = (1-\alpha) V_1(C_1) + \alpha V_2(C_2)$ where U is agent 1's utility, C_i is consumption of agent i, $i=1,2$, $0<\alpha<1$ is the weight that agent 1 places on the felicity of agent 2 relative to his own felicity, and V_i is the direct pleasure of agent i from consumption (felicity). Suppose the total supply of the consumption good is fixed at $C_1 + C_2 = \bar{C}$, and that initially all this quantity is in the hands of agent 1. Take the case $V_i(C_i) = ln(C_i)$. Agent 1 maximizes his utility. This requires that
$$\frac{\partial U(C_1,C_2)}{\partial C_1} = \frac{1-\alpha}{C_1} - \frac{\alpha}{\bar{C}-C_1} = 0 \text{ which implies that } \frac{C_2}{C_1} = \frac{\alpha}{1-\alpha}. \text{ Since } \frac{\partial(C_2/C_1)}{\partial\alpha} = \frac{1}{(1-\alpha)^2} > 0,$$
we see that altruism entails agent 1 giving up some consumption for the sake of agent 2 receiving more consumption (as long as $\alpha > 0$, $\frac{C_2}{C_1} > 0$ which is equivalent to $C_1 < \bar{C}$), and that a stronger altruism results in a larger transfer. Thus, the nature of altruism is giving up some for the sake of another receiving more.
2. The matching process can be characterized in the following way. Interpret m as the fraction of each of the two types who systematically marry members of their own type, and interpret $1-m$ as the fraction of each of the two types who marry randomly (that is, independently of type). We refer to cooperators as type *C* and to defectors as type *D*. Thus, a fraction mx of the population are individuals of type *C* who systematically marry individuals of type *C*, whereas a fraction $m(1-x)$ of the population are individuals of type *D* who systematically marry individuals of type *D*. Of the $1-m$ who marry randomly, x^2 are of type *CC* and $(1-x)^2$ are of type *DD*. Therefore, the total fraction of marriages that are of type *CC* is $mx + (1-m)x^2 = x^2 + mx(1-x)$, and the total fraction of marriages that are of type *DD* is $m(1-x) + (1-m)(1-x)^2 = (1-x)^2 + mx(1-x)$. Finally, of the fraction $1-m$ of type *C* who marry randomly, $x(1-x)$ are of type *CD* and of the fraction $1-m$ of type *D* who marry randomly, $(1-x)x$ are of type *CD*. Therefore, the total fraction of marriages that are of type *CD* is $(1-m)x(1-x) + (1-m)(1-x)x = 2(1-m)x(1-x)$.

The outcome

Given the assumptions about the rule of imitation and the formation of couples, what happens to the share of cooperators in the population, x? We specify a case where the unique and stable equilibrium is one in which the entire population will consist of cooperators.[3] For this monomorphic outcome to occur, two conditions must be satisfied. First, that a population of defectors would be "invaded" by cooperators. Second, that a population of cooperators could not be "invaded" by defectors.

The proportion of cooperators in the population will increase or decrease depending on whether the average payoff to cooperators is higher or lower than that of defectors. If defectors were as likely as cooperators to have cooperative siblings, then defectors would get higher expected payoffs than cooperators. However, siblings are more likely to be similar than random pairs of individuals.

Claim 1: As the proportion of one type in the population becomes rare, the *probability* that an individual of the rare type is married to an individual of the rare type approaches m.

Proof: Consider, for example, the case of rare cooperators. If an individual is a married cooperator, what is the probability that he will be married to a cooperator, when cooperators are rare in a population consisting of, say, N couples? This conditional probability is the total number of cooperators married to cooperators, divided by the total number of cooperators who are married at all, that is:

$$\frac{2[x^2+mx(1-x)]N}{2[x^2+mx(1-x)]N+2(1-m)x(1-x)N}=x+m-mx,$$

which, when $x \to 0$, is equal to m. □

Thus, when cooperators are rare, the *probability* of a cooperator-cooperator match is m.

When the proportion of one type in the population approaches zero, what is the probability that an individual of the rare type has a sibling of the rare type?

Claim 2: The probability that an individual of the rare type has a sibling of the rare type approaches $(1+m)v^2/2$.

Proof: When cooperators are rare, a child can be a cooperator only if the child imitates a parent, provided the parent is a cooperator. (Clearly, if the child imitates a nonparent, the child most surely will be a defector.) In order for both a child and his sibling to be cooperators, both children need to imitate either parent when both parents are cooperators, and the cooperating parent when one parent is a cooperator and one is a defector. The probability of the first of these events is mv^2; the probability of the second event is $(1-m)v^2/2$.[4] The probability then that a cooperating child will have a cooperating sibling is $mv^2+(1-m)v^2/2=(1+m)v^2/2$. □ [5]

3. By "stable equilibrium" we mean an equilibrium that is dynamically stable. This should not be confused with the notion of Nash equilibrium in "evolutionary stable strategies" discussed in evolutionary game theory.
4. The probability that "both children are cooperators" is equal to the probability that "both children imitate a parent ∩ the parent is a cooperator." This probability is equal to $v^2 \cdot 1$ when both parents are cooperators – which in turn occurs with probability m, and to $v^2/2$ when one parent is a cooperator and the other parent is a defector – which in turn occurs with probability $1-m$.
5. Equations (1) through (9) in the appendix provide an alternative proof of Claim 2.

Claim 3: When cooperators are rare, the difference between the expected payoff of a rare cooperator and that of a normal defector (that is, a defector child born to a two-defector couple) is
$\beta=(1+m)(v^2/2)(R-S)-(P-S)$.

Proof: When cooperators are rare, the expected payoff to a cooperator from the game played with a sibling is determined by the probability that the cooperator has a cooperator sibling, which is $(1+m)v^2/2$, by the probability that the cooperator has a defector sibling, which is $[1-(1+m)v^2/2]$, and by the respective payoffs. The expected payoff is therefore
$(1+m)(v^2/2)R+[1-(1+m)v^2/2]S$.
When cooperators are rare, the expected payoff to a normal defector from the game the defector plays with a sibling is P.[6] The difference between the expected payoff of a rare cooperator and that of a normal defector is

$$(1+m)(v^2/2)R+[1-(1+m)v^2/2]S-P =$$
$$(1+m)(v^2/2)(R-S)-(P-S)=\beta. \quad \square$$

A similar procedure shows that when defectors are rare, the difference between the expected payoff of a cooperator and the expected payoff of a defector is
$\alpha=(1+m)(v^2/2)(T-P)-(T-R)$.

Claim 4: When β and α are both positive, the population will consist entirely of cooperators.[7]

We cannot, of course, say that β and α must be positive. But we can find prisoner's dilemma games with payoff parameters S,P,R,T such that both $\beta>0$ and $\alpha>0$.[8]

Explaining the outcome

The likelihood that cooperative behavior will prevail depends on $(1+m)v^2/2$. If children are likely to imitate their parents rather than a random role model, v is high; and parents are likely to be cooperators when m is high. The higher is $(1+m)v^2/2$, the greater the set of payoff parameters for which both β and α are positive, in which case the population will consist of cooperators only. That is, the

6. We ignore the possible case in which a defector child interacts with a cooperator sibling because when cooperators are rare, that is, $x\to 0$, the conditional probability that a sibling of a child of type D is of type C, which is $(1-k)x$ where $k=(1+m)v^2/2$, approaches zero. Conversely, when cooperators are rare, the conditional probability that a sibling of a child of type D is of type D, which is $k+(1-x)$, approaches 1. (These conditional probabilities are derived in the appendix.)
7. The assumption that a small group of cooperators will continue to grow when it has already gotten bigger and will, in the end, take over the entire population requires examination of intermediate cases, that is, of cases other than the ones in which cooperators are rare ($x\to 0$) or defectors are rare ($x\to 1$). However, when the structure of the model is linear in such a way that we can infer about the intermediate cases from the extreme cases, a study of the intermediate cases is not necessary. A proof that this applies in the case of the current model is provided in the appendix. (An alternative proof is provided in Stark (1995), chapter 6.) In the appendix we show that the model has a simple linear structure: the difference between the expected payoff to a cooperator child from interacting with a sibling and the expected payoff to a defector child from interacting with a sibling is $\alpha x+\beta(1-x)$. This expression is positive for any x (that is, not only for $x\to 0$ or $x\to 1$) if and only if α and β are both positive.
8. Note that $(1+m)v^2/2$ lies in the closed interval $[0,1]$. It is an increasing function of both m and v. We have $(1+m)v^2/2=0$ if and only if $v=0$, and $(1+m)v^2/2=1$ if and only if $m=v=1$. We have that $\beta>0$ if and only if $(1+m)v^2/2>(P-S)/(R-S)=k_1$, and that $\alpha>0$ if and only if $(1+m)v^2/2>(T-R)/(T-P)=k_2$. The numbers k_1 and k_2 lie strictly between zero and 1 since $S<P<R<T$.

greater is $(1+m)v^2/2$, the more likely it is that cooperative behavior will prevail. In particular, in the extreme case $m=v=1$, we get $\beta=\alpha=R-P>0$ and the population will consist of only cooperators for any set of payoff parameters.

Reference

Stark, Oded. 1995. *Altruism and Beyond, An Economic Analysis of Transfers and Exchanges Within Families and Groups.* Cambridge: Cambridge University Press.

Appendix

Denote the fractions of marriages that are of the three types, that is, two cooperators, cooperator-defector, and two defectors, by, respectively, γ_{CC}, γ_{CD}, and γ_{DD}. Then,

(1) $\gamma_{CC}=x^2+mx(1-x)$,

(2) $\gamma_{CD}=2(1-m)x(1-x)$,

(3) $\gamma_{DD}=(1-x)^2+mx(1-x)$.

The probabilities that a child in a marriage of each type is of type C are

(4) $\delta_{CC}=(1-v)x+v$,

(5) $\delta_{CD}=(1-v)x+\frac{v}{2}$,

(6) $\delta_{DD}=(1-v)x$.

Assuming that the number of children on average is the same in all types of marriages, the total fraction of children who are of type C is

(7) $\Delta=\gamma_{CC}\delta_{CC}+\gamma_{CD}\delta_{CD}+\gamma_{DD}\delta_{DD}=x.$[9]

Given that a child is of type C, denote by ε_{CC}, ε_{CD}, and ε_{DD} the conditional probabilities that the marriage into which the child was born is of a particular type. Thus,

(8) $\varepsilon_{CC}=\dfrac{\gamma_{CC}\delta_{CC}}{\Delta}$.

Similar equations can be written for ε_{CD} and ε_{DD}. For a given pair of siblings, G and H, the conditional probability that H is of type C given that G is of type C is

(9) $A(x,m,v)=\gamma_{CC}\delta_{CC}+\gamma_{CD}\delta_{CD}+\gamma_{DD}\delta_{DD}\underset{x\to 0}{=}(1+m)v^2/2$

or, for any arbitrary value of x,

(10) $A(x,m,v)=x+(1-x)k=k+(1-k)x$
where $k=(1+m)v^2/2$.

Therefore, the conditional probability that a sibling of a child of type C is of type D is

(11) $1-[k+(1-k)x]=(1-k)(1-x)$.

The conditional probability that a sibling of a child of type D is of type D is given simply by substituting $1-x$ for x in (10). We thus get,

(12) $k+(1-k)(1-x)$
and similarly, the conditional probability that a sibling of a child of type D is of type C is given by substituting x for $1-x$ in (11). We thus get,

(13) $(1-k)x$.

(We can also derive this last probability from (12) by writing $1-[k+(1-k)(1-x)]=(1-k)x$.)
Hence, when $x\to 0$, the conditional probabilities (12) and (13) approach, respectively, 1 and zero.

For the general case (any arbitrary value of x), by using (10) and (11) we calculate first the expected payoff of a child of type C from interacting with a sibling. This payoff is $[k+(1-k)x]R+(1-k)(1-x)S$. Next, by using (12) and (13) we calculate the expected payoff of a child of type D from interacting with a sibling. This payoff is $[k+(1-k)(1-x)]P+(1-k)xT$. Therefore, the difference between these two numbers is

(14) $x[R-[kP+(1-k)T]]+(1-x)[[kR+(1-k)S]-P]=x\alpha+(1-x)\beta$

where α and β are defined in the last but one section of the paper. The expression $\alpha x+\beta(1-x)$ is positive for any x if and only if α and β are both positive. In *Claim 4* we take $x\to 0$, in which case (14) reduces to
$kR+(1-k)S-P=(1+m)(v^2/2)R+[1-(1+m)v^2/2]S-P$.

9. Exploiting the similarity between the δs, this result can be obtained as follows:
$\gamma_{CC}+\gamma_{CD}+\gamma_{DD}=x^2+2mx(1-x)+2(1-m)x(1-x)+(1-x)^2=x^2+2x(1-x)+(1-x)^2=1$, and

$\gamma_{CC}+\frac{1}{2}\gamma_{CD}=x^2+x(1-x)=x$. Thus,

$\Delta=\gamma_{CC}\delta_{CC}+\gamma_{CD}\delta_{CD}+\gamma_{DD}\delta_{DD}=\gamma_{CC}((1-v)x+v)+\gamma_{CD}((1-v)x+\frac{v}{2})+\gamma_{DD}(1-v)x$

$=(\gamma_{CC}+\gamma_{CD}+\gamma_{DD})(1-v)x+(\gamma_{CC}+\frac{1}{2}\gamma_{CD})v=(1-v)x+xv=x.$