Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior

Robert Boyd [a,*], Peter J. Richerson [b]

[a] Department of Anthropology, University of California, Los Angeles, CA 90095, USA
[b] Department of Environmental Science and Policy and Ecology Graduate Group, University of California, Davis, CA 95616, USA

ARTICLE INFO

ABSTRACT

Human migration is nonrandom. In small scale societies of the past, and in the modern world, people tend to move to wealthier, safer, and more just societies from poorer, more violent, less just societies. If immigrants are assimilated, such nonrandom migration can increase the occurrence of culturally transmitted beliefs, values, and institutions that cause societies to be attractive to immigrants. Here we describe and analyze a simple model of this process. This model suggests that long run outcomes depend on the relative strength of migration and local adaptation. When local adaption is strong enough to preserve cultural variation among groups, cultural variants that make societies attractive always predominate, but never drive alternative variants to extinction. When migration predominates, outcomes depend both on the relative attractiveness of alternative variants and on the initial sizes of societies that provide and receive immigrants.

© 2009 Published by Elsevier Ltd.

With civilized nations, as far as an advanced standard of morality and an increased number of fairly good men are concerned, natural selection apparently affects little …. [T]he causes which lead to the advance of morality [are] the approbation of our fellow-men—the strengthening of our sympathies by habit—example and imitation—reason—experience and even self-interest—instruction during youth and religious feelings.
Darwin (1874, pp. 185–186), *Descent of Man*

1. Introduction

Humans are an unusual animal because we live in large, complex societies with a lot of cooperation among non-relatives. Explaining the scale of human cooperation has proven to be one of science's greatest challenges. Among sociologists and anthropologists, this pattern has given rise to an enduring explanatory paradigm, functionalism (Turner and Maryanski, 1979). It has suggested to many evolutionists, beginning with Darwin (1874, pp. 178–179), that humans must be subject to some form of group level selection (Alexander, 1987, p. 169; Bowles, 2006; Eibl-Eibesfeld, 1982; Hamilton, 1975; Wilson, 1975, pp. 561–562). Such explanations have been quite controversial and other proposals exist (Smith, 2003; Sober and Wilson, 1998), many invoking some

form of Alexander's (1987) idea of indirect reciprocity (Nowak and Sigmund, 1998).

One possibility is that human cooperation arose under the influence of group selection acting on cultural rather than genetic variation (Boyd and Richerson, 1982; Henrich, 2004) in which variation among societies is maintained by conformism, moralistic punishment, or some other process that creates multiple stable equilibria (Boyd and Richerson, 1992; Henrich and Boyd, 1998; Boyd et al., 2003; Panchanathan and Boyd, 2004). At such equilibria, the common type (or types) has higher fitness than alternative invading types. This means that in a single population, there are many possible outcomes, and which outcome is reached will then be determined by the accidents of initial conditions. However, if such population is subdivided into partially isolated subpopulations, adaptive processes can maintain variation among groups. Then if subpopulations near one equilibrium have lower extinction rates or produce more migrants, the variants that characterize that equilibrium can spread to the population as a whole.

This process is not subject to the usual criticisms directed at group selection for altruistic variants (Boyd and Richerson, 2007). To see why, it will be useful to introduce the Price equation (Price, 1970). In a population structured into groups, the change in frequency of a variant undergoing selection or an analogous cultural process (Boyd and Richerson, 1985), Δp, is given by

$$\Delta p \propto \underbrace{V_G \beta_G}_{\text{between groups}} + \underbrace{\overline{V_W \beta_W}}_{\text{within groups}}$$

* Corresponding author. Tel.: +1 310 909 3156.
  E-mail address: rboyd@anthro.ucla.edu (R. Boyd).

The first term gives the change due to selection between groups, and second term gives the change in frequency due to changes within groups. The βs give the effect of the behavior on the fitness of groups (β_G) and individuals (β_W). First, consider an altruistic behavior that is beneficial to the group and costly to the individual. Selection between groups increases the frequency of the altruistic trait, but selection within groups decreases it, so $\beta_G > 0$ and $\beta_W < 0$. The strength of these two processes depends on the relative magnitudes the variance in gene frequency between groups (V_G) and within the groups (V_W). The reason that selection among large, isolated groups does not usually lead to evolution of altruism is that when groups are large, selection is weak, and there is even a modest amount of migration among them, the variance between individuals (V_W) will be much larger than the variance between groups (V_G; Rogers, 1990). Thus unless selection within groups is much weaker than selection among groups ($\beta_G \gg \beta_W$), group selection cannot overcome opposing individual selection. Now suppose that there two stable equilibria—one at which one trait is common and a second at which the alternative is stable. One of these equilibria has higher average fitness than the other. Let p be the frequency of this group beneficial trait. This means that as before $\beta_G > 0$. If selection is strong compared to migration, each group will be near one of the two equilibria, and thus $\beta_W \approx 0$ and $V_G \gg V_W$ and, therefore, selection between groups will be strong and selection within groups will be weak. Moreover, this is true even if populations are large, and migration rates are substantial. The main requirement is that rates of adaptation within groups are high compared to rates of migration between them, and as a result this process is more likely to be important for cultural evolution than for genetic evolution. When these conditions are satisfied, group selection will lead to the spread of the most group beneficial equilibrium.

Two different kinds of intergroup competition have been studied. Early models closely reflected Darwin's (1874, p. 179) idea that tribal scale units were in conflict with one another and that the physical or social triumph of groups "possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to aid one another, and sacrifice themselves for the common good, would be victorious over most other tribes". Evidence strongly suggests that our tribal ancestors engaged in much conflict (Keeley, 1996; Otterbein, 1985) and that the details in the best understood case, highland New Guinea, fit the cultural group selection hypothesis (Soltis et al., 1995). In modern societies, competition between tribal scale organizations like business firms may function in a similar way (Nelson and Winter, 1982). Our analysis indicates that this kind of group selection can only work if successful groups expand by subdivision and the diffusion of ideas between groups is inhibited (Boyd and Richerson, 1990). The latter requirement will tend to prevent societies successful in one dimension from acquiring group beneficial cultural systems operating in other dimensions from other groups.

This kind of cultural group selection cannot be the whole story because it is too slow to account for much observed change. The data on extinction rates from New Guinea (Soltis et al., 1995) indicate that intergroup competition is likely to lead to the fairly slow spread of favored cultural variants. To some extent, the slow evolution of group-beneficial traits and the failure of different societies to benefit from the experiences of others, fit the stylized facts of human history. Human societies have grown larger and more sophisticated over the last 10,000 years, but major changes occur on millennial time scales. However, much rapid social evolution has also been documented. This is especially true today when mass literacy, mass communication, and mass migration due to cheap communication drive a very rapidly changing world. But even in tribal societies, surprisingly rapid spurts of cultural

evolution are known. For example, the several tribes of Plains Indians of North America created their novel horse-and-fur-trade adaptation in a century or so after they acquired horses from the Spanish (Oliver, 1962).

Selective borrowing can also cause the spread of group beneficial behaviors. Henrich and Gil-White (2001) review the evidence that people have a tendency to imitate successful individuals, and this bias leads to the spread of beliefs and practices that give rise outcomes that are deemed successful. When multiple equilibria exist, neighboring groups may be more variable in success than individuals within groups, and as a result the tendency to imitate successful neighbors can lead to the spread of group beneficial ideas. A simple model indicates that such selective borrowing can lead to rapid cultural evolution (Boyd and Richerson, 2002). It also allows group beneficial innovations arising in different groups to come together in the same group without independent innovations—a kind of group level recombination that greatly increases the potential for the cumulative evolution of complex cultural adaptations.

Here we investigate third process that leads to the evolution of group beneficial traits, selective migration. We are very familiar with this process in the modern world (Migration, 2005). Cheap mechanized transportation made it possible for large numbers of people to move from one place to another. Streams of migrants flow from societies that migrants perceive as offering them fewer opportunities toward ones that appear to offer them more (Martin, 2005). Immigrants often assimilate to their host culture. This is certainly true for most immigrants to the US where the process is so strong that regional differences among founding 17th century immigrants to different American colonies can still be detected despite subsequent waves of immigration (Fischer, 1989). Modern immigration and assimilation is a much studied phenomenon (e.g. Alba and Nee, 2003; Borjas, 1994) giving rise to a large literature that we cannot review in detail. The literature supports two generalizations that we use as a basis for modeling, namely (1) that migrants flow from societies where immigrants find their prospects poor to ones where they perceive them to be better, and (2) most immigrant populations assimilate to the host culture within a few generations.

Other evidence suggests that selective immigration is an ancient phenomenon. Knauft (1985) gives an interesting ethnographic example. He studied the Gebusi, a shrinking lowland New Guinea group with very high rates of witchcraft murder. The neighbors of the Gebusi, the Bedamini, had a better system for dispute resolution system resulting in a larger, more functional society. At the margins of Gebusi territory, some Gebusi had ties of marriage and friendship with the Bedamini which they mobilized to switch their ethnic identity to Bedamini. Cronk (2002) describes how the Mukogodo people of Kenya abandoned their historic language and hunting and gathering economy in favor of the language and economy of their high-prestige Maasai cattle-keeping neighbors. Barth (1969) famously argued that in traditional societies, individuals often have claims to more than one ethnicity and that therefore they can strategically deploy the ethnicity that they perceive as most advantageous. If he is correct, most ethnic boundaries are sufficiently porous to permit an appreciable amount of selective immigration. Language provides historical traces of immigration. Languages in contact always generate a flow of words and grammatical constructions across the linguistic frontier. The vehicle for linguistic borrowings is multilingual speakers, who are often people in the process of immigration (Thomason, 2001). Selective immigration and assimilation often leads to the death of the languages of disfavored societies (Abrams and Strogatz, 2003).

Ancient complex societies, such as China, Rome, and India seem likely to have grown substantially by immigration and

assimilation (United Nations, 1973, p. 225). Ancient imperial systems often expanded militarily, but the durable ones, such as Rome, succeeded by assimilating conquered peoples and by inducing a flow of migrants across their boundaries. The Roman legal system, competent administration, military protection, provision of roads and other infrastructure led to considerable wealth. Grants of citizenship and the development of vibrant civic culture in the form of provincial Roman towns in the far flung outposts of the Empire like Britain sustained Rome for several centuries (Shelton, 1998, Chapter XII). The "Confucian" administrative system of China similarly promoted expansion, peace and economic development it that region of the world (Gernet, 1996). The Hindu caste system incorporated tribal peoples into Indian civilization as self-governing *varna* (Gadgil and Malhotra, 1982; Srinivas, 1962). Islam originated as a religion bringing peace and unity to Arabian tribes (Armstrong, 1991). Islam, Roman Christianity (Stark, 1997), and Asian Buddhism formed the cores of many societies via their ability attract converts. Thus, although the Roman empires (west and east) eventually faded, their most attractive institutions were adapted by successor polities and persist in modified form to this day. Rome, India, China, and Islamic civilization stand in stark contrast to pure conquest empires like that of the Mongols, which left no institutional traces.

In this paper we analyze a simple model in which selective migration and assimilation lead to the spread of group beneficial cultural variants. No doubt the expansion of some societies at the expense of others is complex and includes immigration and assimilation among other processes. However, here we focus only on the effects of migration. We assume that cultural variation among groups is maintained by within-group processes which act against the spread of rare cultural variants. A number of such processes including conformism, moralistic punishment, and coordination games have been documented (Richerson and Boyd, 2005). If migration rates are low enough, such rare-type disadvantage processes will also lead to the assimilation of immigrants. However, a number of interesting questions remain unanswered:

1. Low migration rates will also minimize the effect of selective migration. Is there a plausible range of migration rates that allows group variation to persist and immigrants to be assimilated, but at the same time allows selective migration to play an important evolutionary role?
2. Rare type disadvantage processes create multiple cultural equilibria, each with its own basin of attraction. Mixing between groups will tend to shift groups toward equilibria with larger basins of attraction. Since equilibria with higher average payoffs need not have larger basins of attraction, this process may compete with the effects of selective migration. Are there conditions in which the equilibrium with the higher payoff, not the larger basin of attraction will spread?
3. Larger groups will produce more emigrants than smaller groups, all else equal, which means that migration will tend to lead to the spread of behaviors common in larger groups. Cumulative evolution of group beneficial traits requires that group beneficial behaviors can spread even when they are relatively rare. Are there conditions with allow this to occur?

The analysis of the model suggests that when local adaptation is strong enough to maintain between group variation, selective migration leads to the spread of group beneficial behaviors even when they have smaller basins of attraction or initially exist in smaller groups. However, unlike processes based on differential imitation or military or economic competition, selective migration

typically leads to a polymorphic outcome in which group beneficial variants predominate but alternative variants also persist.

2. Model assumptions

Consider a population divided into two very large subpopulations, labeled 1 and 2. Members of each group are characterized by observable culturally transmitted marker traits. These traits have no direct effect of individual welfare. A fraction p of the population belongs to subpopulation 1 and $1-p$ to subpopulation 2.

We want the model to represent a situation in which different social norms can be stabilized in different groups, and, because different norms differ in average welfare, individuals are motivated to move from one group to the other. Suppose that there are two possible behaviors, labeled 1 and 2. Suppose that x_i is the frequency of behavior 1 in subpopulation i. Then let the expected payoffs of the two behaviors be

$$W_{i1}(x_i) = 1 - h + x_i(d + g + h)$$
$$W_{i2}(x_i) = 1 + d - x_i d \tag{1}$$

When $x_i = 1$, $W_{i1}(x_i) = 1 + d + g > 1 = W_{i2}(x_i)$ in a well mixed population behavior 1 is an ESS. Similarly, when $x_i = 0$, $W_{i1}(x_i) = 1 - h < 1 + d = W_{i2}(x_i)$ so behavior 2 is also an ESS. There is a single unstable equilibrium that marks the boundary between the two basins of attraction. Setting the two expected payoffs equal and solving yields the following expression for this equilibrium, $\hat{x}$:

$$\hat{x} = \frac{d+h}{g+h+d} \tag{2}$$

Populations in which behavior 1 is common have higher average payoff than populations in which behavior 2 is common. However, if $h > g$, behavior 1 has a smaller basin of attraction. These ideas are illustrated graphically in Fig. 1.

This model captures the basic payoff structure of a variety of models of the evolution of norms. Some norms result from coordination games. Examples include conventions like driving on the right or left or speaking English or Chinese (Sugden, 1986) and mutualistic interactions (Skryms, 2003). Other norms are maintained by systems of reciprocity, reputation and punishment. When individuals interact repeatedly, evolution can favor contingent strategies in which some behaviors are rewarded and others are punished. Such strategies can stabilize a wide variety of social norms, including ones that lead to large scale cooperation (Axelrod, 1986; Boyd and Richerson, 1992; Nowak and Sigmund, 1998; Henrich and Boyd, 1998; Panchanathan and Boyd, 2004). As
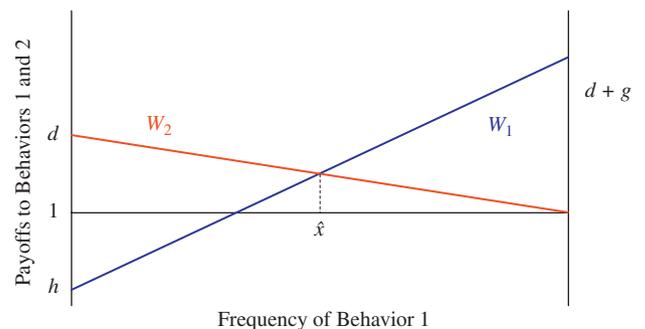

Fig. 1. Plots the payoffs of players in the coordination game as a function of the frequency of behavior 1. Notice that the strategy with the highest payoff need not have the largest basin of attraction.

is shown in Appendix A, simple models of mutualism and of reputation and punishment result in exactly the payoff structure given in (1). More complex models with two variants result in payoffs that are nonlinear functions of frequency. However, many of them have the same qualitative structure—both strategies are evolutionarily stable, there is a single unstable equilibrium separating their basins of attraction, and the basin of attraction of the group beneficial equilibrium may be smaller than that of the other equilibrium. Models with more than two variants remain to be studied, but we expect that the results presented will provide insight about more complex models.

So that payoffs shape cultural evolution within subpopulations, we assume that each time period individuals observe a randomly chosen member of their own subpopulation and switch with a probability that is proportional to the difference in payoffs between the two individuals. We do not allow imitation of members of the other subpopulation because that would create a second, competing mechanism that can also lead to the spread of group beneficial equilibria (Boyd and Richerson, 2002). Here we preclude this mechanism in order to isolate the effects of payoff biased migration.

With this assumption, the change in frequency of trait 1 in the subpopulation i ($= 1, 2$) after imitation is given by

$$x_i' = x_i + x_i(1 - x_i)\beta(W_{i1} - W_{i2}) \tag{3}$$

where β is a constant that transforms payoff units into the probability of imitation. See McElreath and Boyd (2007) for a derivation and a discussion of alternative models.

Substituting the definitions of W_{i1} and W_{i2} yields the simple form

$$x_i' = x_i + ax_i(1 - x_i)(x_i - \hat{x}) \tag{4}$$

where $\alpha = \beta/2d + g + h$.

After payoff based imitation, some individuals move between subpopulations. Assume that the overall size of the population is constant, but that the relative size of the two subpopulations is determined by the movement of peoples between them. Let m_{ij} be the fraction of individuals in subpopulation i who leave and join subpopulation j. Then the size of subpopulation 1 after migration, p'', is

$$p'' = p(1 - m_{12}) + (1 - p)m_{21} \tag{5}$$

Migration will also affect the frequencies of the traits in each subpopulation. After migration, the frequencies of traits 1 in subpopulations 1 and 2 will be

$$x_1'' = \frac{x_1'p(1 - m_{12}) + x_2'(1 - p)m_{21}}{p(1 - m_{12}) + (1 - p)m_{21}}$$
$$x_2'' = \frac{x_2'(1 - p)(1 - m_{21}) + x_1'pm_{12}}{p(1 - m_{12}) + (1 - p)m_{21}} \tag{6}$$

The numerators in (6) give the number of individuals with behavior 1 in each subpopulation after migration and the denominators give the total number of individuals in each subpopulation. To allow for payoff biased migration, we assume

$$m_{ij} = m_0(1 + \mu(\overline{W}_j - \overline{W}_i)) \tag{7}$$

where $\overline{W}_j$ is the average payoff in subpopulation j and the parameter μ gives the importance of difference in payoffs in migration decisions. The parameter μ is chosen so that the migration rates are always between zero and one. These expressions are meant to capture the idea that people's migration decisions are affected by many motives so that there is migration from high payoff subpopulations to lower payoff subpopulations as well as the reverse.

From (1) and (2) the difference in average payoff is

$$\overline{W}_i - \overline{W}_j = (x_i'^2 - x_j'^2)g - (x_i'(1 - x_i') - x_j'(1 - x_j'))(h + 2d) \tag{8}$$

3. Analysis of the model

This model has two kinds of stable equilibria. As long as there is some migration ($m_0 > 0$), populations in which all individuals in both subpopulations have the same behavior ($\hat{x}_1 = \hat{x}_2 = 0, 1$), and both subpopulations are the same size ($\hat{p} = 0.5$), are always locally stable equilibria. When behaviors have the same frequencies in both subpopulations, migration ceases to affect the frequencies, and since both behaviors have higher payoffs when common, a population in which everybody shares the same behavior can resist invasion by the alternative behavior. We label these *monomorphic equilibria*.

If migration rates are low enough, there also two other locally stable equilibria, in which behavior 1 is common in one subpopulation and rare in the other. We label these as *polymorphic equilibria*. We have not been able to find general expressions for subpopulation sizes or behavior frequencies at these equilibria. However, we can derive values for two special cases. First when migration is not payoff biased ($\mu = 0$), the equilibria are symmetrical so that $\hat{x}_1 = 1 - \hat{x}_2$ and $\hat{p} = 0.5$. Of more interest, when migration rates are low enough, the following approximate expressions hold:

$$\tilde{x}_1 \approx 1 - \frac{m_0(1 - \mu g)}{\beta(g + d)}$$
$$\tilde{x}_2 \approx \frac{m_0(1 + \mu g)}{\beta(h + d)}$$
$$\tilde{p} \approx \frac{(1 + \mu g)}{2} \tag{9}$$

These expressions say that when behavior 1 strongly increases average fitness, and average fitness affects migration rates, behavior 1 predominates in the population though the evolutionarily stable outcome is polymorphic.

Fig. 2 shows the equilibrium frequencies and populations sizes as a function of the relative strength of biased imitation and migration (m_0/β) for different levels of payoff bias in migration (μ) determined by numerically iterating the recursions in (4)–(8). The figure also plots the values predicted by the approximate expressions in (9). When m_0 is small compared to β, the approximate expressions are quite accurate. However, as the relative strength of migration increases, the equilibrium frequencies become more similar until a threshold is reached and the polymorphic equilibria are no longer stable. Notice that the relative sizes of the two subpopulations depends on the average payoff advantage of populations in which behavior 1 is common (g) and the importance of relative payoff on the migration rates between the two subpopulations. Also notice that the approximate expressions underestimate the subpopulation size in which the group beneficial behavior is common.

Since populations in which everybody uses behavior 1 or everybody uses behavior 2 are always stable, the evolutionary outcome always depends on initial conditions. The dynamic behavior of the model is illustrated in Figs. 3 and 4. When migration rates are high (Fig. 3) all trajectories converge to one of the monomorphic equilibria; when migration rates are lower (Fig. 4), the population converges either to monomorphic or polymorphic equilibria depending on initial conditions. The dynamics depend both on the frequencies of the two behaviors and the relative size of the two subpopulations. So, when one subpopulation is much larger than the other, the population tends to evolve toward an equilibrium in which the larger subpopulation's behavior predominates.

The spread of a new group beneficial variant is fundamentally a disequilibrium phenomenon—something changes that leads to migration when there was none before. It might be that two, culturally different subpopulations come into contact for the first time, or it might be that some change in technology or institutions in one group makes migration more attractive. In such cases, the
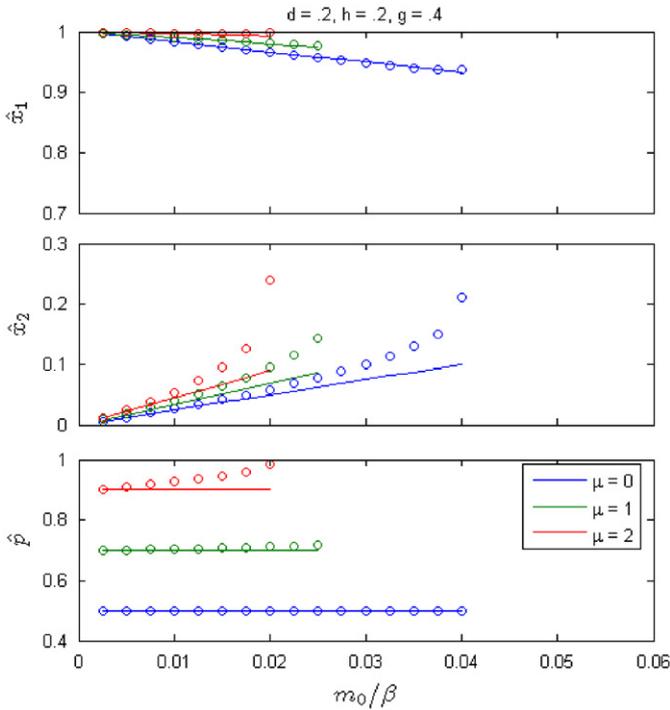
Fig. 2. The open circles in the top two panels give the equilibrium frequencies of behavior 1 in the two subpopulations at the heterogeneous equilibrium in which behavior 1 is common in subpopulation 1 as a function of the ratio of base migration rate m_0 to the base rate of cultural adaptation, β, for three values of μ. The bottom panel gives the size of subpopulation 1 at that equilibrium. The solid lines give the predicted values computed using the weak migration approximation given in (10). There is a second heterogenous equilibrium in which the labels of the two subpopulations are reversed. Other parameter values are $d = 0.02$, $h = 0.2$, $g = 0.4$.
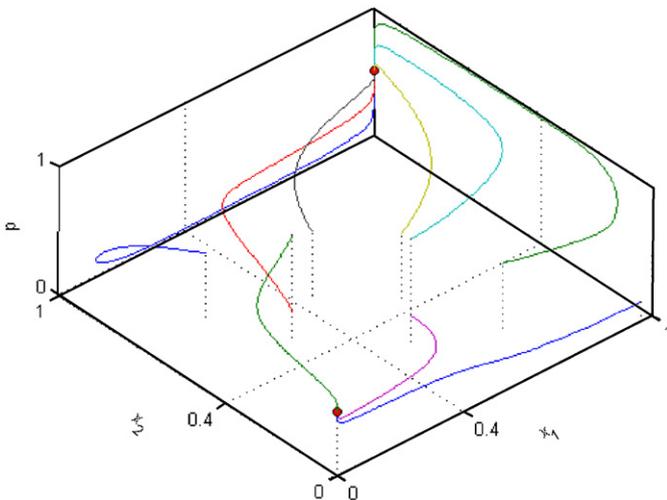

Fig. 3. Trajectories for different initial conditions plotted in phase space for a high enough migration rate that only the monomorphic equilibria are stable ($m_0/\beta = 0.04$). The trajectories begin at the points marked by the dotted lines that drop to the $p = 0$ plane and end in the red circles. The other parameter values are $d = 0.02$, $h = 0.2$, $g = 0.4$ and $\mu = 2$. The unstable equilibrium separating the basins of attraction in an isolated population is at the value 0.4. Notice that basins of attraction depend on the initial size of populations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

two subpopulations are initially different, and the question is then what happens? When does the behavior that leads to the highest group payoff come to predominate, and when do behaviors in the
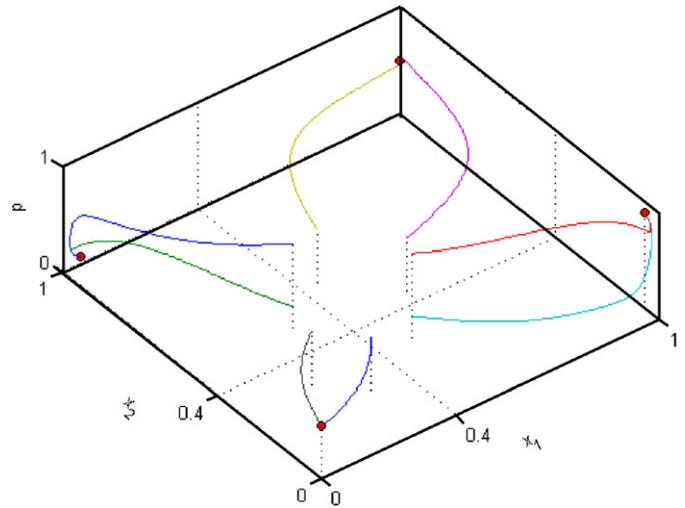

Fig. 4. Trajectories for different initial conditions plotted in phase space for low enough migration rate that both monomorphic and polymorphic equilibria are stable ($m_0/\beta = 0.01$). The trajectories begin at the points marked by the dotted lines that drop to the $p = 0$ plane and end in the red circles. The other parameter values are $d = 0.2$, $h = 0.2$, $g = 0.4$ and $\mu = 2$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

larger subpopulation predominate despite lower payoffs? And, how are outcomes affected by the relative size of the basins of attraction of the two equilibria?

To answer these questions, we studied the behavior of the model with an initial state in which one behavior is common in one subpopulation and the alternative behavior is common in the other subpopulation. As is illustrated in Fig. 5, the outcome depends on both the initial sizes of the two subpopulations, and the migration rate. When migration rates are high, the population evolves to the one of the two monomorphic equilibria depending on the initial size of the two subpopulations. When migration rates are lower, there are three possible outcomes. If one subpopulation is much larger than the other, the population may evolve to the monomorphic equilibrium in which the entire population is characterized by the behavior that predominated in the initially larger subpopulation. If the two subpopulations are closer to the same size, the population evolves to the polymorphic equilibrium.

Figs. 6 and 7 show how equilibrium outcomes are affected by the initial sizes the two subpopulations and by model parameters. In all cases, behavior 1 is initially common in subpopulation 1 ($x_1 = 1$) and rare in subpopulation 2 ($x_2 = 0$). The initial size of subpopulation 1 is given by the vertical axis, and the relative strength of migration is plotted on the horizontal axis. The labeled regions refer to the three possible equilibrium outcomes. First, consider Fig. 6 which plots equilibrium outcomes when the group beneficial equilibrium also has a larger basin of attraction. When migration rates are high enough, the population behaves as if it were well mixed. If the initial frequency of behavior 1 in the population as a whole is in the basin of attraction of behavior 1 ($p_0 > \hat{x}$), the population evolves to the monomorphic equilibrium at which behavior one is common. If it is not, the population evolves toward the monomorphic equilibrium at which behavior 2 is common. Decreasing the migration rate increases the range of initial conditions that lead to an equilibrium at which the group beneficial behavior is common. The fact that this occurs even when migration is not payoff biased ($\mu = 0$) suggests that the increase is due to the fact that behavior 1 has a larger basin of
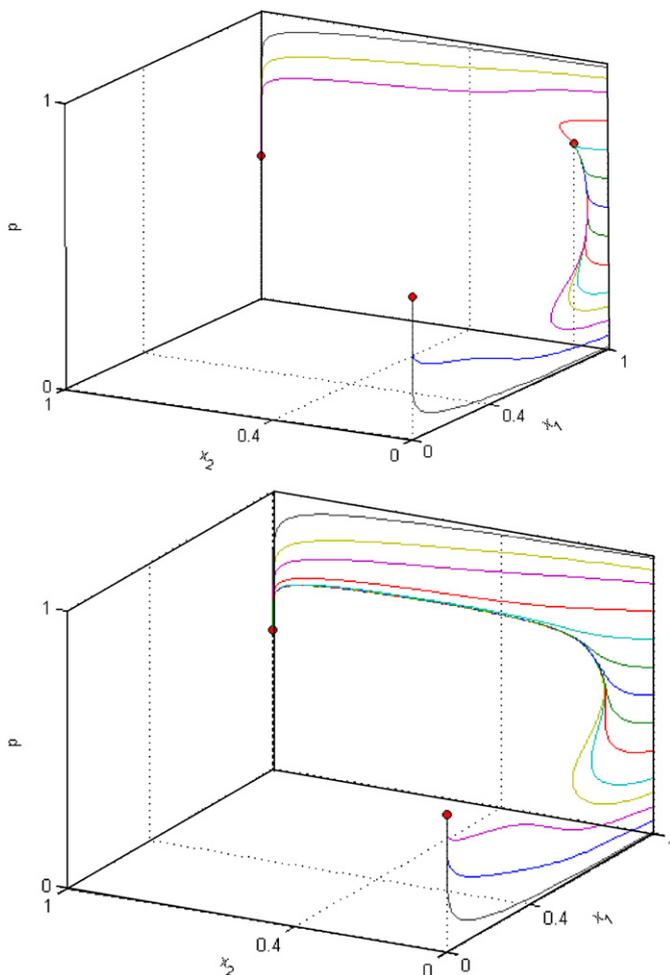
Fig. 5. Trajectories with the initial conditions $x_1 = 1$ and $x_2 = 0$ for different initial values of p, the size of subpopulation 1 for two values of m_0. In the top panel, the migration rate is low ($m_0/\beta = 0.02$), and as a result the polymorphic equilibrium is stable. In the lower panel, the migration rate is higher ($m_0/\beta = 0.04$), and only the monomorphic equilibria are stable. The other parameter values are $d = 0.02$, $h = 0.2$, $g = 0.4$ and $\mu = 1$. Notice that in both cases there is a threshold initial size above which the population evolves to the equilibrium at which the group beneficial behavior is more common than the alternative.

attraction. When migration rates are low enough so that cultural variation can persist, almost all initial conditions converge to the polymorphic equilibrium. Since the frequency of the group beneficial behavior at this equilibrium is sensitive to the strength of payoff bias, and thus under conditions in which cultural variation among subpopulations is stable, payoff bias strongly increases the frequency of group beneficial behavior. Fig. 7 shows what happens when the group beneficial equilibrium has a smaller basin of attraction. As before, when migration is strong enough, the population behaves as if it were well mixed. But now the effect of decreasing migration does depend on the extent to which it is payoff biased. When there is no payoff bias, decreasing migration rates *decreases* the range of initial conditions which lead to an equilibrium at which the group beneficial trait is common because it has a smaller basin of attraction. When payoff bias is strong enough, decreasing migration rates increases the range of initial conditions which lead to an equilibrium at which the group beneficial trait is common. As before, once migration rates are low enough that cultural variation between subpopulations can be maintained, the polymorphic equilibrium is reached from a wide range of initial subpopulation sizes.

4. Discussion and conclusions

The simple model analyzed above indicates that the effect of selective migration depends strongly on the rate of migration among subpopulations. When local adaptation is strong enough compared to migration to maintain cultural variation among subpopulations and when such variation exists and affects average welfare, populations evolve toward a polymorphic equilibrium at which cultural variation persists, but the group beneficial behavior is more common. Initial subpopulation sizes and the sizes of the basins of attraction play relatively minor roles. When migration is stronger, initial population sizes and sizes of the basins of attraction predominate.

Thus, selective voluntary migration from disfavored to favored societies followed by assimilation can act as a strong evolutionary force. Note that it is an interesting force mainly when the disfavored society can persist. If the trait that makes the disfavored society disfavored is declining of its own accord, voting with your feet can only hasten the extirpation of the disfavored trait. In the interesting case, the disfavored trait will never go extinct in the disfavored society by the mechanism we model here. But its relative size will shrink in proportion the absolute advantage of the favored trait (g) and the strength of selective migration (μ).

Of course the real world is more complex. Take the case of the Iron Curtain in the days of the Cold War. The Warsaw Pact countries took steps to prevent residents from voting with their feet that involved barbed wire and gun towers, and when the will to continue such measures broke down, the Pact dissolved and each country undertook reforms aimed at reducing the incentives to leave. Mostly, they attempted with greater or lesser success to remodel their economies along market lines. Donor countries today are seldom content to lose masses of people without imitating the policies that confer the advantage on the receiving societies.

The receiving societies are often ambivalent about large flows of immigrants. The short term costs of absorbing immigrants may be perceived to be high. People worry that immigrants may not assimilate even though they take advantage of the benefits produced by the receiving economies. Note that we have only considered one character in this model. Immigrants may indeed bring durable influences to their host cultures. For example, the sales of salsa in the US is said to have passed that of catsup. Although nativists often worry about the corruption of their culture by foreign ideas, the "recombination" of desirable features from different cultures overcomes a big problem with simple contest group selection which breaks down if much migration between societies occurs. From the broader perspective of the advance of the human welfare more generally, it is important to note that the immigration model is not hostile to recombination.

Anthropologists and linguists worry that immigration is absorbing vital small societies into the underclass of modern ones. In the case of very large societies impinging upon very small societies the model suggests indeed that a large society with a low payoff equilibrium can overwhelm a small society with a higher payoff equilibrium. Size dependent migration effects are liable to hasten the demise of such small societies. Given that the attractions of modern societies may be illusory (Easterlin, 2001), we cannot necessarily assume that immigration is benign, especially if the current attractiveness of modernity is being purchased at the expense of future sustainability. We have elsewhere modeled the process by which symbolic group boundaries can arise to protect cultures from the effects of migration (McElreath et al., 2003). At least some small societies do not assimilate, such as the Anabaptist groups in the US and Canada. Those with high payoff equilibria are more likely to evolve such boundaries than ones with low payoff equilibria.
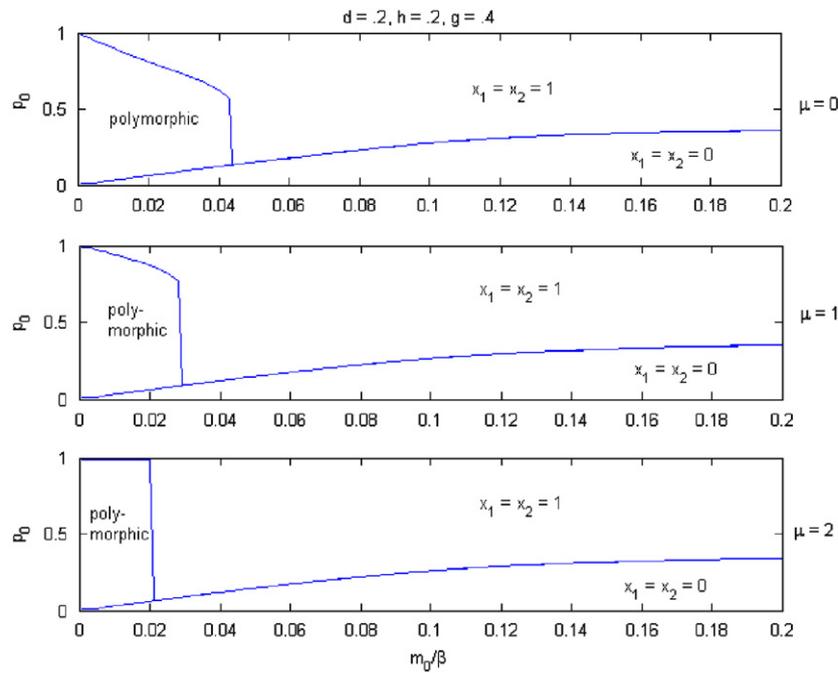
Fig. 6. The range of initial sizes of subpopulation 1 (p_0) that lead to each of the three possible stable equilibria as a function of the migration rate (m_0/β) and the strength of group benefit for behavior 1 (μ) are shown. In all cases the behavior 1 is initially common in subpopulation 1 ($x_1 = 1$) and behavior 2 in subpopulation 2 ($x_2 = 0$). The boundary between the basins of attraction, $\hat{x}$, is 0.4. When migration is strong, the populations behaves as if it were well mixed, and behavior 1 can increase only if the average frequency in the population as a whole is greater than 0.4. However, when migration is weak, behavior 1 can increase even if subpopulation 1 is quite small.
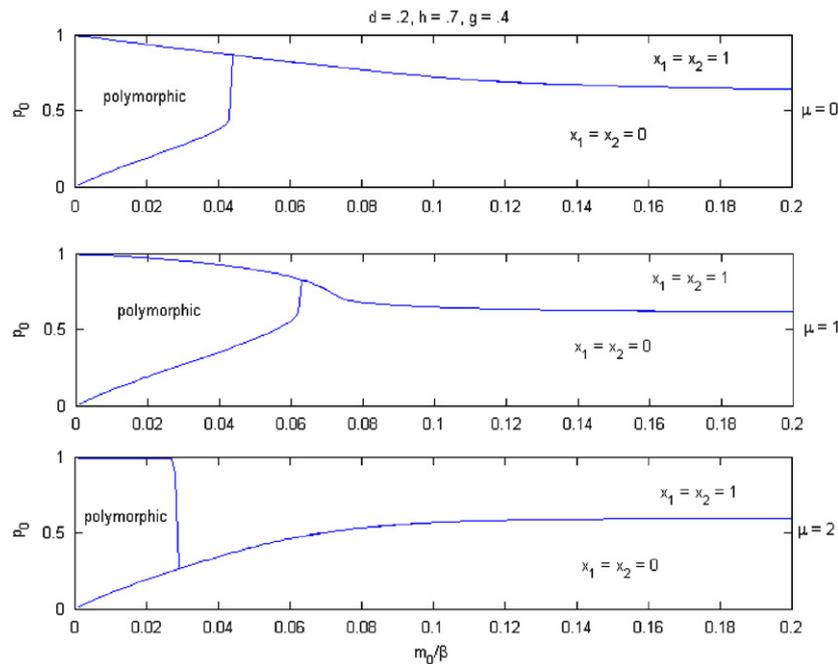

Fig. 7. The range of initial sizes of subpopulation 1 (p) that lead to each of the three possible stable equilibria as a function of the migration rate (m_0/β) and the strength of group benefit for behavior 1 (μ) are shown. In all cases the behavior 1 is initially common in subpopulation 1 ($x_1 = 1$) and behavior 2 in subpopulation 2 ($x_2 = 0$). The boundary between the basins of attraction, $\hat{x}$, is 0.7. When migration is strong, the populations behaves as if it were well mixed, and behavior 1 can increase only if the average frequency in the population as a whole is greater than 0.7. However, when migration is weak, behavior 1 can increase even if subpopulation 1 is quiet small.

Thus, we are inclined to count migration as a relatively benign force generally favoring improvements in the human condition, especially compared to group-competitive struggles for dominance and power. Group selection by violent physical and social extirpation of other societies has no doubt been an important force in human history, bringing in its train frightful slaughter and nightmarish weapons. Some today seem to think such processes are inevitable, and seem to relish the prospect of future clashes. In the past, immigrants have been attracted institutions like the great world religions and the tolerant civilizations they were often associated with. The ideas that millennium in and millennium out have contributed most to civilization in the best sense of that word have increased mainly by attracting immigrants or inspiring foreign imitation or both. As Darwin's remarks in our epigraph, natural selection in the form of violent conflict is not the main means of our moral progress.

Table A1
Payoffs in the two person Stag Hunt game.

	Mutualistic	Selfish
Mutualistic	$1+d+g$, $1+d+g$	$1-h$, 1
Selfish	1, $1-h$	$1+d$, $1+d$

Appendix A

Mutualism is often modeled as a Stag Hunt game. The payoff matrix for a two person Stag Hunt is given in Table A1. There are two strategies, mutualistic and selfish. If both players pursue mutualistic behavior, they both achieve a higher payoff. However, a mutualist paired with a selfish type does worse than the solitary individual. If players are paired at random, this yields exactly the payoffs given in (1). An *n*-person version of the game in which payoffs are linear functions of the number of mutualists yields the same payoff structure (McElreath and Boyd, 2007, p. 373).

The following simple model captures many features of the evolution of norms that are maintained by reputation and punishment. Groups of *n* individuals interact. The game has two stages: During the first stage, individuals cooperate or defect. Each cooperative act produces a benefit of b/n for every member of the group, but reduces the payoff of the cooperator an amount *c*. During the second stage, individuals can punish others an amount *p* at a cost *k*. There are two strategies. Defectors defect and do not punish. Cooperators cooperate and punish non-cooperators. Let *i* be the number of cooperators among the other $n-1$ individuals in a group. Then the payoffs of the two types are

$$V_D = \frac{ib}{n} - pi$$

$$V_c = \frac{i+1}{n}b - c - k(n-1-i)$$

Let *x* be the frequency of in the population as a whole. Then, assuming that $b>c>b/n$ and $b/n<p$, taking the expectations yields a linear, bistable payoff structure like that given in (1). This simple model avoids the second-order free rider problem by assuming that there are no individuals that cooperate but do not punish. However, punishment can be stabilized by repeated interaction (Boyd and Richerson, 1992), conformism (Henrich and Boyd, 2001), group selection (Boyd et al., 2003), and when punishment takes the form of denied indirect reciprocity (Panchanathan and Boyd, 2004). More complex multiperiod models generate nonlinear payoff functions. Nonetheless, they frequently have two stable equilibria at which only one behavior is present, and a single unstable equilibrium dividing the two basins of attraction.

References

Abrams, D.M., Strogatz, S.H., 2003. Modeling the dynamics of language death. Nature 424, 900.

Alba, R., Nee, V., 2003. Remaking the American Mainstream: Assimilation and the New Immigration. Harvard University Press, Cambridge, MA.

Alexander, R.D., 1987. The Biology of Moral Systems. Aldine de Gruyter, Hawthorne, NY.

Armstrong, K., 1991. Muhammad: A Western Attempt to Understand Islam. Victor Gollancz, London.

Axelrod, R., 1986. An evolutionary approach to norms. American Political Science Review 80, 1095–1111.

Barth, F., 1969. Ethnic Groups and Boundaries: The Social Organization of Culture Difference. Little Brown and Company, Boston.

Borjas, G.J., 1994. The economics of immigration. Journal of Economic Literature 32, 1667–1717.

Bowles, S., 2006. Group competition, reproductive leveling, and the evolution of human altruism. Science 314, 1569–1572.

Boyd, R., Richerson, P.J., 1982. Cultural transmission and the evolution of cooperative behavior. Human Ecology 10, 325–351.

Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. University of Chicago Press, Chicago, IL.

Boyd, R., Richerson, P.J., 1990. Culture and Cooperation. In: Mansbridge, J.J. (Ed.), Beyond Self-interest. Chicago University Press, Chicago, pp. 111–132.

Boyd, R., Richerson, P.J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology 13, 171–195.

Boyd, R., Richerson, P.J., 2002. Group beneficial norms can spread rapidly in a structured population. Journal of Theoretical Biology 215, 287–296.

Boyd, R., Richerson, P.J., 2007. In: Gangestad, S., Simpson, J. (Eds.), Group Selection: A Tale of Two Controversies, Evolution of Mind, Fundamental Questions and Contorversies. Guilford Press, New York, pp. 221–225.

Boyd, R., Gintis, H., Bowles, S., Richerson, P.J., 2003. The evolution of altruistic punishment. Proceeding of the National Academy of Sciences USA 100, 3531–3535.

Cronk, L., 2002. From true Dorobo to Mukogodo Maasai: contested ethnicity in Kenya. Ethnology 41, 27–49.

Darwin, C., 1874. The Descent of Man and Selection in Relation to Sex. American Home Library, New York.

Easterlin, R.A., 2001. Subjective well-being and economic analysis: a brief introduction. Journal of Economic Behavior & Organization 45, 225–226.

Eibl-Eibesfeld, I., 1982. Warfare, man's indoctrinability, and group selection. Zeitschrift fur Tierpsychologie 67, 177–198.

Fischer, D.H., 1989. Albion's Seed: four British Folkways in America. America, a cultural history, vol. 1. Oxford University Press, New York.

Gadgil, M., Malhotra, K.C., 1982. Ecology of a pastoral casts: Gavli Dhangars of peninsular India. In Human Ecology 10, 107–143.

Gernet, J., 1996. A History of Chinese Civilization. Cambridge University Press, Cambridge.

Hamilton, W.D., 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: Fox, R. (Ed.), Biosocial Anthropology. Wiley, New York, pp. 133–155.

Henrich, J., 2004. Cultural group selection, coevolutionary processes and large-scale cooperation. A target article with commentary. Journal of Economic Behavior and Organization 53, 3–143.

Henrich, J., Boyd, R., 1998. The evolution of conformist transmission and the emergence of between-group differences. Evolution and Human Behavior 19, 215–242.

Henrich, J., Boyd, R., 2001. Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. Journal of Theoretical Biology 208, 79–89.

Henrich, J., Gil-White, F.J., 2001. The evolution of prestige-Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior 22, 165–196.

Keeley, L.H., 1996. War before Civilization. Oxford University Press, New York.

Knauft, B.M., 1985. Good Company and Violence: Sorcery and Social Action in a Lowland New Guinea Society. Studies in Melanesian Anthropology. University of California Press, Berkeley.

Martin, P., 2005. Migrants in the global labor market, Global Commission on International Migration, pp. 1–57.

McElreath, R., Boyd, R., 2007. Mathematical Models of Social Evolution: A guide for the Perplexed. University of Chicago Press, Chicago.

McElreath, R., Boyd, R., Richerson, P.J., 2003. Shared norms and the evolution of ethnic markers. Current Anthropology 44, 122–129.

Migration, I.O.f., 2005. World Migration 2005: Costs and Benefits of International Migration, International Organization for Migration, 494pp.

Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Belknap Press of Harvard University Press, Cambridge, MA.

Nowak, M.A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. Nature 393, 573–577.

Oliver, S.C., 1962. Ecology and Cultural Continuity as Contributing Factors in the Social Organization of the Plains Indians. Publications in American Archaeology and Ethnology, vol. 48, no. 1. University of California Press, Berkeley.

Otterbein, K.F., 1985. The Evolution of War: A Cross-cultural Study. Human Relations Area Files Press, New Haven, CT.

Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432, 499–502.

Price, G., 1970. Selection and covariance. Nature 227, 520–521.

Richerson, P.J., Boyd, R., 2005. Not by Genes Alone: How Culture Transformed Human Evolution. University of Chicago Press, Chicago.

Shelton, J.-A., 1998. As the Romans did. Oxford University Press, Oxford.

Skryms, B., 2003. The Stag Hunt and the Evolution of Social Structure. Cambridge University Press, Cambridge.

Smith, E.A., 2003. Human cooperation: perspectives from behavioral ecology. In: Hammerstein, P. (Ed.), Genetic and Cultural Evolution of Cooperation. MIT Press, Cambridge, MA, pp. 401–428.

Sober, E., Wilson, D.S., 1998. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge, MA.

Soltis, J., Boyd, R., Richerson, P.J., 1995. Can group-functional behaviors evolve by cultural group election? An empirical test. Current Anthropology 36, 473–494.

Srinivas, M.N., 1962. Caste in Modern India, and Other Essays. Asia Publishing House, Bombay.

Stark, R., 1997. The Rise of Christianity: How the Obscure, Marginal Jesus Movement became the Dominant Religious Force in the Western World in a Few Centuries. HarperCollins, San Francisco.

Sugden, R., 1986. The Economics of Rights, Co-operation, and Welfare. Blackwell, London.

Thomason, S.G., 2001. Language Contact. Georgetown University Press, Washington, DC.

Turner, J.H., Maryanski, A., 1979. Functionalism. The Benjamin/Cummings Publishing Company, Menlo Park, CA.

United Nations, 1973. The Determinants and Consequences of Population Trends: New Summary of Findings on Interaction of Demographic, Economic, and Social Factors, Department of Economic and Social Affairs, New York, pp. xvi+661.

Wilson, E.O., 1975. Sociobiology: The New Synthesis. Harvard University Press, Cambridge, MA.