

Online Enhancements for Division of Labor, Economic Specialization and the Evolution of Social Stratification

Joseph Henrich and Robert Boyd

OUTLINE

This online appendix has five main sections.

- 1) Simulation
- 2) Derivation of model
 - a. Success biases
 - b. Migration
 - c. Five equilibrium solutions
 - d. Why economic interactions (δ) between groups is likely not near zero
 - e. Derivation of total payoffs
- 3) Discussion of the limitation of time-scales
- 4) An ethnographic example of occupational specialization, Swat Valley.
- 5) The evolution of the division of labor in other species

SIMULATION

A Visual Basic program that visually simulates this process is available at <http://www.sscnet.ucla.edu/anthro/faculty/boyd/MESB/ClassesSimulationFiles.zip>.

DERIVATION OF THE MODEL

There is now an expanding toolbox of formalizations for studying cultural evolution, learning, and strategic interaction (Gintis 2000, McElreath and Boyd 2007, Weibull 1995, Young 1998). To express the influence of success-biased cultural learning in our model we used a standard form of replicator dynamics. Equations like (1), which express the change in the frequency of individuals with strategy H in a large population, have been derived in a wide variety of ways under different assumptions (Hofbauer and Sigmund 1988, Schlag 1998, Weibull 1995), and can provide a good approximation even when strategies are continuous (Henrich and Boyd 2002). This basic form is rather robust across derivational assumptions. Here we provide a simple derivation of the replicator dynamic equation (1), and show how it can be linked to our two forms of mixing.

Success biased learning

First, we consider the case in which success biased learning occurs within each subpopulation. Assume that each time step each individual meets another randomly chosen individual from within his or her own subpopulation and compares the payoff received by this other individual with his or her own payoff. The learner then copies the other individual with a probability proportional to the difference between their payoffs. Several different “proportional learning rules” have been studied and are particularly interesting because analytical work shows that they are close to the optimal updating rule under a wide range of conditions (Schlag 1998,

1999). For concreteness, suppose that imitation is governed by the learning rule given in table A1.

Table A1: The probabilities of different encounters between H and L and the probabilities of the Imitator acquiring traits H and L after the encounter assuming that imitation occurs within subpopulations.				
Potential Imitator	Model Strategy	Probability of Pairing	Probability of H after learning	Probability of H after learning
H	H	p_i^2	1	0
H	L	$p_i(1-p_i)$	$\frac{1}{2}(1 + \beta(\pi_{Hi} - \pi_{Li}))$	$\frac{1}{2}(1 - \beta(\pi_{Hi} - \pi_{Li}))$
L	H	$p_i(1-p_i)$	$\frac{1}{2}(1 + \beta(\pi_{Hi} - \pi_{Li}))$	$\frac{1}{2}(1 - \beta(\pi_{Hi} - \pi_{Li}))$
L	L	$(1-p_i)^2$	0	1

Where p_i is the frequency of H in subpopulation i ($=1, 2$) just before learning takes place. Then the frequency of H in subpopulation i after learning takes place, $L(p_i)$ is

$$\begin{aligned}
 L(p_i) &= p_i^2 + 2p_i(1-p_i)\left(\frac{1}{2}(1 + \beta(\pi_{Hi} - \pi_{Li}))\right) \\
 &= p_i + p_i(1-p_i)\beta(\pi_{Hi} - \pi_{Li}) \\
 &= p_i + \Delta p_i^L
 \end{aligned} \tag{A1}$$

The term Δp_i^L gives the change in frequency of those with strategy H due to learning in one time step. The superscript L indicates that this is the contribution due to learning. More extensive details can be found in McElreath and Boyd (2007).

Here we note that in our early modeling efforts we also included conformist transmission and found no important qualitative differences in the results. In fact, conformist transmission made the emergence of social stratification somewhat more likely.

The assumption that people imitate the successful does not mean that vertical (parent-offspring) cultural transmission is unimportant. This model leaves open the possibility that individuals *first* engage in vertical cultural transmission, and later modify their beliefs by preferentially imitating the successful. The key is that the change in the frequency of beliefs is caused by success bias. Further discussion and modeling of such 2-stage cultural learning can be found in Henrich (2004) and Boyd & Richerson (1985).

Migration

To calculate the effects of mixing, suppose that a fraction m of the individuals in subpopulation 1 emigrate and are replaced by a fraction m drawn from subpopulation 2. Then, if the frequencies of H in subpopulations 1 and 2 just before migration are p_1 and p_2 respectively, then the frequency of H in subpopulation 1, after migration, $M_1(p_1, p_2)$, is (A2), which gives the frequency of strategy H in subpopulation 1 after migration. The first term on the right hand side gives the starting value of p_1 from before migration, the second term is the loss of H-individuals due to emigration, and the third term gives the gain in H-individuals due to incomers from subpopulation 2.

$$M_1(p_1, p_2) = (1 - m)p_1 + mp_2 = p_1 + m(p_2 - p_1) = p_1 + \Delta p_1^m \quad (\text{A2})$$

The term Δp_1^m gives the change in the frequency of H in subpopulation 1 due to migration in one time step. A similar derivation yields that parallel expression for the change in subpopulation 2.

Assuming selective learning processes are weak avoids assumptions about the life cycle

In general, $M_1(L(p_1), p_2) \neq L(M(p_1, p_2))$, which means that the dynamics of change depend on the order in which learning and migration occur. We regard this dependence on order as an artifact of the simple, discrete time structure of the model, not an essential feature. To eliminate this dependence, assume that the changes in the frequencies of strategies are small enough during any one time step that terms of order $m\beta$ can be ignored compared to terms of order m or β , then

$$M_1(L(p_1), p_2) = L(M(p_1, p_2)) = p_1 + \Delta p_1^L + \Delta p_1^M + O(\beta m). \quad (\text{A3})$$

This means the order of learning and migration does not matter, and the dynamics are given by equations (1) and (4) in the main body of the paper.

To make sure that the assumption of small m and β does not qualitatively affect the behavior of the model we simulated exact recursions for the three possible life cycles—migration before imitation, migration after imitation, and success-biased imitation of individuals in the other subpopulation. Figure A1 shows the equilibrium values of p_1 and p_2 for a range of m values for the two migration models cycle models and for our analytical solution which assumes weak selective forces. The qualitative behavior of these models is identical.

[Figures A1 about here]

The main text leaves open the possibility that m , migration, could capture either the physical movement of individuals or the flow of ideas. Suppose that individuals observe the behavior of a randomly chose individual from their own group with probability $1 - 2m$ and from the other group with probability $2m$, and then use the payoff biased imitation rule described

above. If they observed the behavior of an individual from their own subpopulation, the probabilities of acquiring each of the two traits is given by Table A1. If they observe the behavior of an individual from the other subpopulation, the probabilities are given in Table A2. Notice that in Table A2 if $\beta = 0$, such that imitation is not success biased, the probability of adopting the behavior of the other individual is $\frac{1}{2}$ and therefore the probability of imitating a randomly chosen individual from the other group is m , and thus a probability of $2m$ of observing a member of the other group is equivalent to a migration rate of m . Using these two tables leads the following expression for the frequency of H in subpopulation 1 after social learning:

$$\begin{aligned}
p'_1 &= (1 - 2m) \left[p_1^2 + p_1(1 - p_1)(1 + \beta(\pi_{H1} - \pi_{L1})) \right] \\
&\quad + 2m \left[p_1 p_2 + \frac{1}{2} p_1(1 - p_2)(1 + \beta(\pi_{H1} - \pi_{L2})) + \frac{1}{2} p_2(1 - p_1)(1 + \beta(\pi_{H2} - \pi_{L1})) \right] \quad (A4) \\
&= p_1 + p_1(1 - p_1)\beta(\pi_{H1} - \pi_{L1}) + m(p_2 - p_1) + O(\beta m)
\end{aligned}$$

A similar expression can be derived for the change in the frequency of H in subpopulation 2. With the βm terms assumed negligible we arrive back at the recursions used in the main text.

We compared these two models via simulation and provide an illustrative result in Figure A2, which explores the same conditions used in Figure 1. In general, as in Figure A1, there is no qualitative difference in the behavior of the two approaches.

Table A2: The probabilities of different encounters between H and L and the probabilities of the Imitator acquiring traits H and L after the encounter assuming that imitation occurs between subpopulations.				
Potential Imitator	Model	Probability of Pairing	Probability of H after learning	Probability of H after learning
H	H	$p_1 p_2$	1	0
H	L	$p_1(1 - p_2)$	$1/2(1 + \beta(\pi_{H1} - \pi_{L2}))$	$1/2(1 - \beta(\pi_{H1} - \pi_{L2}))$

L	H	$p_2(1-p_1)$	$1/2(1+\beta(\pi_{H2}-\pi_{L1}))$	$1/2(1-\beta(\pi_{H2}-\pi_{L1}))$
L	L	$(1-p_1)(1-p_2)$	0	1

To make sure that the assumption of small m and β does not qualitatively affect the behavior of the model we simulated exact recursions success-biased imitation across subpopulations. Figure A2 shows the equilibrium values for a range of m values for the model in which individuals imitate members of the other subpopulation and the analytical solution. These results indicate that the exact recursions have the same qualitative properties as the analytical solution which assumes weak forces. This program, written in Visual Basic 5, is available upon request.

[Figure A2 about here]

Payoff-biased physical migration

In considering migration as the movement of people, one concern is that the difference in payoffs between subpopulations might increase the flow of people from the lower payoff subpopulations to the higher payoff subpopulations and reduce flow in the opposite direction. Intuitively, one might think that this would undermine the results presented above. This is not the case, however. To investigate this question, we modified the model described above so that the migration rate from subpopulation 1 to subpopulation 2 is $m(1+a(\bar{\pi}_2-\bar{\pi}_1))$ and the migration rate from subpopulation 2 to subpopulation 1 is $m(1+a(\bar{\pi}_1-\bar{\pi}_2))$. The parameter a controls how strongly payoffs affect migration. When $a=0$, payoffs have no effect; as a increases above zero the flow of people from the low to the high payoff subpopulation increases while the reverse direction decreases.

This model is too complex to solve analytically, however, numerical solutions indicate that increasing the parameter a **increases**—not decreases—the amount of stratification. Figure A3 shows the equilibrium values of p_1 and p_2 as a function of m for three different values of a , the parameter that controls the magnitude of the payoff bias in migration, $a = 0$, $a = 2$, and $a = 3$. When m is low enough to allow stratification, increasing a means that individuals in the low payoff group have a higher migration rate and individuals in the high payoff group have a lower migration rate. The magnitude of this effect depends on the value of m . These effects can be substantial. For example, for $a = 2$ and $m = 0.02$, the migration for the high and low payoff groups are 0.013 and 0.027 respectively. Adding the success-biased physical migration to the model actually increases the range of conditions conducive to social stratification, thus our assumption of fixed symmetric migration in generating our analytical solution was conservative.

[Figure A3 about here]

We conjecture that the reason is that the reduced migration from the high payoff subpopulation has a bigger effect on the dynamics than the increased migration from the low payoff subpopulation.. This migration was combined with the within-group learning model described above.

Since this approach creates differential migration based on payoff differences between the subpopulations, it also addresses the concern that we have not endogenized the decision to migrate based on payoff differences.

In considering physical migration it is important to understand that while economic incentives have likely long influenced migration, so did many other factors. Such movements would have been extremely costly as families were embedded in long-term communities, kinship

systems, and networks of relations and obligations. There is also reason to suspect that cues of ethnicity (McElreath, Boyd, and Richerson 2003) and our ethnic psychology (Gil-White 2001), which are likely much older than social stratification, provided an existing social formation that would have impeded such differential migration (Henrich and Henrich 2007: Chapter 9). Thus, any adjustments to the migration between subpopulations created by the economic incentives for one's offspring would have been merely small adjustments to a background rate of migration, not its primary determinant.

Five Equilibrium solutions

There are five equilibrium solutions to our set of difference equations. The first two are trivial, $\hat{p}_1 = \hat{p}_2 = 0$ and $\hat{p}_1 = \hat{p}_2 = 1$, and always unstable for any interesting parameter combinations. The third is the egalitarian equilibrium in which $\hat{p}_1 = \hat{p}_2 = \gamma$. The other two are the stratified equilibria, and are completely symmetric (the values of \hat{p}_1 and \hat{p}_2 can be switched).

The values of p_1 and p_2 at the stratified equilibrium are:

$$\hat{p}_1 = \frac{2m(G\beta\gamma - 2m)}{G\beta(1-\gamma)(G\beta\gamma - 2m) + \sqrt{-G\beta(2m - G\beta(1-\gamma))(2m - G\beta\gamma)(2m - G\beta(1-\gamma)\gamma)}} \quad (\text{A5})$$

$$\hat{p}_2 = \frac{2m(2m - G\beta\gamma)}{G\beta(1-\gamma)(2m - G\beta\gamma) + \sqrt{-G\beta(2m - G\beta(1-\gamma))(2m - G\beta\gamma)(2m - G\beta(1-\gamma)\gamma)}} \quad (\text{A6})$$

Why economic interactions (δ) between groups is likely not near zero

In considering δ some might worry that most interactions actually occur within subpopulations, not between, so δ is likely near zero. However, with smaller-scale societies in mind, we would argue that since cultural transmission occurs mostly within subpopulations,

individuals within subpopulations will tend to share similar areas of knowledge, practices and strategies so that much of the relevant variation will tend emerge between subpopulations. This means that individuals will tend to seek out members of other subpopulations with complementary sets of knowledge, practices, and strategies, thereby driving δ towards one. We have fixed δ exogenously between 0.5 and 1 in order to examine how constraints on between group interactions might influence the emergence of economic specialization and stratification. We show that constraints on economic interaction inhibit the emergence of stratification.

Derivation of Total Group Payoffs

The derivation of (15), the total population payoffs, is composed by summing up the payoffs received by each strategy in each subpopulation, as follows:

$$\hat{\pi} = p_1(\pi_{H1}) + (1 - p_1)\pi_{L1} + p_2\pi_{H2} + (1 - p_2)\pi_{L2} \quad (\text{A7})$$

We then substitute in equations (2), (3), (5) and (6) in (A7) to yield (15) in the main text.

To calculate total payoff (15) we assume that one randomly selected member from each of our large subpopulations participates in each economic exchange. We do not need to assume that populations are of equal size. Assuming they are unequal in size simply implies that members of the smaller subpopulation participate in a greater frequency of transactions than the larger subpopulation.

LIMITATIONS IN TIME SCALES

As with any formal evolutionary model we have intentionally ignored some aspects of the problem in order to focus on and understand others. Along the way we have pointed out several of our key assumptions, but here we want to highlight one additional limitation. Our

model assumes that cultural evolutionary dynamics operate on relatively shorter time scales than do demographic dynamics. Since novel cultural practices and ideas routinely sweep through populations in a single generation, even in small-scale societies, this assumption is probably not a bad first pass. This is relevant in considering longer time scales because at the stable stratified equilibrium one of the subpopulations receives a higher average payoff than the other subpopulation. If richer people have more children, the richer subpopulation will grow faster than the poorer subpopulation. Of course this need not occur as increased wealth may be associated with reduced population growth. However, if it does occur, demographic changes may destabilize the stratified equilibrium by, for example, creating a biased migration of individuals from the higher to the lower payoff subpopulation, or by reducing the bargaining power and therefore the surplus demanded by individuals from the richer subpopulation, potentially driving the whole population back to the egalitarian equilibrium. Such longer term demographic processes may or may not have time to operate, since individual stratified societies frequently have short life times (Turchin 2003). In any case, a complete analysis is required since cultural group selection acting on different stratified equilibria will be pushing toward smaller and smaller payoffs differences between subpopulations, consistently reducing the payoff—and thus reproductive—differences between subpopulations. We hope additional models can build on the cultural evolutionary dynamic we have highlighted.

INTERPRETING AN ETHNOGRAPHIC EXAMPLE OF OCCUPATION SPECIALIZATION, SWAT VALLEY

The qualitative lessons of our model can be applied to a range of empirical settings. And, while the model can certainly be interpreted as laying a foundation for the emergence of elite controlling classes of priests, warriors, or resource managers, it also applies to situations in

which social groups (perhaps ethnic groups) have evolved to occupy economic niches in a regional economy, which may or may not be ruled by a single political establishment, such as is associated with a hereditary nobility. Barth's (1965) ethnographic work among the Pashto-speaking peoples of the Swat valley, in Pakistan near the border with Afghanistan, provides an informative ethnographic example that illustrates just such a case.

The social organization of the Swat valley consists of sharply differentiated occupational *castes* that specialize as farmers, carpenters, tailors, weavers, potters, smiths, land-owners, barbers, cotton-carders, oil pressers, etc.¹ One's future occupation and likely marriage partner is largely determined by the occupation (caste) of one's father. In a census from four villages, only 16% of persons were involved in occupations different from that of their caste. Despite the intermixing of castes in villages (castes do not form localized communities), 60% of marriages were within the same caste, and an additional 17% occurred with an economically-adjacent caste—in general, when they do deviate, women tend to marry up (23.1%) more than they marry down (17.4%). Even when one does take up the occupation of another caste, the individual, his sons, and grandsons are still considered part of the individual's original caste, which strongly influences all kinds of social relationships, ritual obligations, and patterns of interaction, including one's marriage possibilities. In such cases, caste ascription only becomes ambiguous for great-grandsons.

Economically, these castes are highly interdependent and their interaction depends entirely on dyadic contracts and exchanges. Success in the sophisticated, highly productive forms of agriculture practiced in this valley requires specialized skills. A single agricultural unit, which is generally integrated by a series of decentralized agreements, requires a *landowner, tenant*

farmer or laborer, carpenter, blacksmith, rope-and-thong maker and a muleteer. Each of these is obtained from a caste that specializes in that particular form of labor. Everyone involved is usually paid with some portion of the final harvest, so one's profit depends on the sufficient contributions of everyone else (as in our payoff matrix).² A single farmer could learn all of these skills, but, if the idea behind economic specialization is correct, he could never do them all as well as the specialists. It is the division of skill or knowledge, and the associated norms and relationships, that creates the 'surplus' economic production (*G*).

Despite the seeming rigidity of this system, new castes can emerge in response to novel economic opportunities for specialization. *Tailors*, for example, have emerged as a caste only recently, since sewing machines were introduced 75 years ago. Similarly, Barth tells of a potentially emerging caste based on the manufacture of a particular type of sandal that was developed only 40 years ago. The skills required in making this sandal exceed those of common leatherworkers, making this an honorable occupation that is currently pursued in several places by particularly skilled/trained leatherworkers. Barth's informants had little doubt that this would eventually develop into a sandal-making caste.

These different occupational groups receive different portions of the overall economic pie in the Swat Valley. At the dyadic level there are a variety of contract types in land-for-labor exchanges that show the dyadic inequality, but we use the *brakha-khor* type to illustrate. Under this contract, a tenant farmer supplies the seeds, labor, tools and draught animals—though usually not the manure—and in return takes a fraction of the crop. In less fertile areas, this fraction varies regionally from 3/5 to 1/3 of the total yield, while in fertile areas it is typically 1/4. At the population level, farmers and landowners form occupational castes in which one's

birth strongly influences one's occupational choices, and one's cut of the overall economic pie. Barth observes that the effect of this stratified economic inequality affects differences in average height (between higher and lower castes) and infant mortality, indicating that these differences translate into real average, and durable, group differences in health and fitness.

This example is instructive because, unlike other cases of stratification, Barth makes it clear that individuals of most occupations are not generally coerced into any dyadic contracts or social relationships that they don't like. People shift contracts all the time, and no landowners or other members of a high caste can generally compel anyone into a contract they don't like. Political leadership, control, and influence are determined by consistently shifting political alliances, complex interrelated sets of dyadic contracts, gift giving, personal strength and honor, strategic manipulation, and the tactical use of force. Political change is rapid with leaders often rising and falling over a period of a few years; influence and power are ephemeral; no contracts are binding, and a leader's best weapon is to distribute his own wealth. In fact, the ability of leaders to use physical force depends entirely on giving 'good deals' to ad-hoc assemblies of their current followers. Finally, as noted, people can change castes; they occasionally do, and it's perfectly acceptable. The fact is, however, that they just don't usually do it.

This ethnographic description fits the theoretical expectations of our model. Barth's data suggests that the division of labor between economic specialists leads to increased production, and to specialized occupational subpopulations and stratified inequality. *G* is high: This high level of specialization allows for intensive grain production based on two crops per year, fertilizer (manure), irrigation, and terracing. As noted, the benefits of increased production permit a large, dense population: Swat Valley contained 400,000 people (circa 1954), with some

single villages populated by up to 10,000 individuals. Barth estimates 800 people per square mile of productive land. Further, migration, m , is restricted by having castes that (1) strongly influence one's kin group, social obligations and marriage partners, (2) are 'sticky', such that even changing occupations does not change the caste of your kids or grandkids, and (3) are attached to notions of impurity that reduce social contact between castes of very different status.³ Meanwhile, δ is maximized by (1) having a substantial number of pure (interdependent) specialists, rather than part-time specialists who don't *have* to interact to survive (farmer-carpenters, for example) and (2) spatially interspersed populations rather than isolated mono-caste villages.

THE EVOLUTION OF DIVISION OF LABOR IN OTHER SPECIES

While the division of labor, economic specialization, and exchange among members of the same animal species is relatively common, it rarely results from heritable differences. For example, males and females commonly play different roles in the production of offspring, but sex is not heritable. In some species, individuals take on different morphs depending on non-heritable differences. For example, young salmon who happen to develop rapidly transform into smolts, a large type that moves out into the ocean a year ahead of those who remain in their natal streams for another year, maintaining a slower growth rate and smaller body size (Mangel 1994). The only highly developed systems of division of labor occur in eusocial species like ants and termites in which individuals belonging to different "castes" perform different functions, like guarding the colony, tending the brood, foraging, and so on. There are a few examples of distinct heritable types. For example in the marine isopod *P. scupta* there are three types of males, large males who defend aggregations of females from other males, medium sized males who insinuate

themselves into these aggregations by mimicking female morphology and behavior, and tiny males who attempt to hide amongst the females (Shuster and Wade 1991).

In stark contrast, the division of labor and exchange between members of different species is fairly common. Famous examples include ant species that guard acacia trees that provide them with shelter and nourishment, fungi that supply plants with nitrogen in return for carbohydrates, and insects that transfer pollen in return for an energetic reward.

It seems likely that heritable division of labor within a single species is rare because it can only persist when all types have the same fitness. The logic behind this requirement is enshrined in the Bishop-Cannings Theorem (Bishop and Cannings 1978): if a strategy has an expected payoff less than another strategy in the population, the other strategy ought to be in the process of replacing it. Thus all three morphs in *P. sculpta* have same average mating success. This requirement means that any specialization that increases the fitness of one type relative to the others cannot persist within a species, and thus strongly constrains the kinds of within species specializations that can evolve. In contrast, because genes carried in members of different species do not compete, it does not even make sense to compare their fitnesses. Thus, between species exchange can persist as long as it is beneficial given the behavior of the other species.

Human cultural evolution is intermediate between these two extremes. While individuals from different human subpopulations frequently interact, the transmission of culture may occur predominately within each subpopulation. The model analyzed here indicates that if the amount of mixing is substantial, all types have to have the same payoff and human cultural evolution parallels genetic evolution within other species. However, if the amount of cultural mixing is lower than the *stratification threshold*, human societies are more like an ecosystem in which

different, partially isolated cultural groups evolve mutualism as different species do. Thus, human sociocultural systems, at least under some conditions, can generate subpopulations of strategies in which one strategy is maintained at equilibrium with another strategy that receives a substantially higher payoff.⁴ This means that humans have stratified inequality *because* we are a cultural species.

Ethnographically, this difference can be seen starkly in rural India, where different occupational castes filled an enormous variety of economic and ecological niches (Gadgil and Malhotra 1983). These castes specialized in such things as carpentry, pottery, leatherwork, buffalo-keeping, sheep-keeping, indigenous medicine, tool making (4 different castes), entertainment (12 different castes), religious functions (14 different castes), landownership, and foraging (just to name a few). Among the foraging castes alone, some specialized in some hunting techniques and some species, while other relied on quite separate skills and emphasized different species. Through their interactions with other castes, these castes effectively occupied specialized, mutually beneficial, economic/ecological niches.

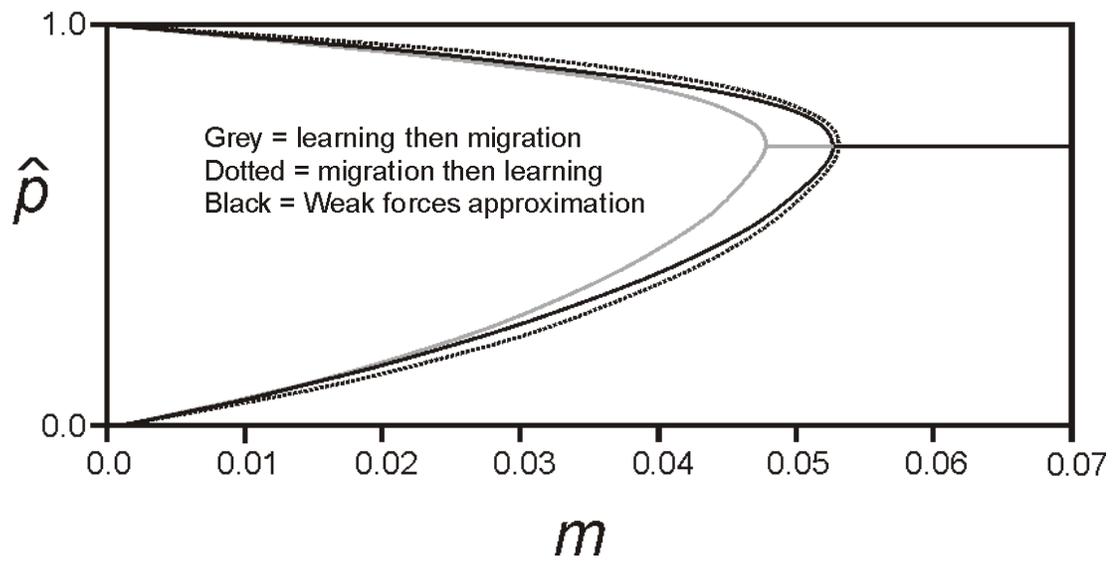


Figure A1. Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, m , for three different models. The black line is the analytical solution assuming weak forces. The grey line assumes learning occurs first followed by migration, and the dashed line assumes that migration occurs first. The parameters are $\gamma = 0.7$ and $\square G = 0.5$.

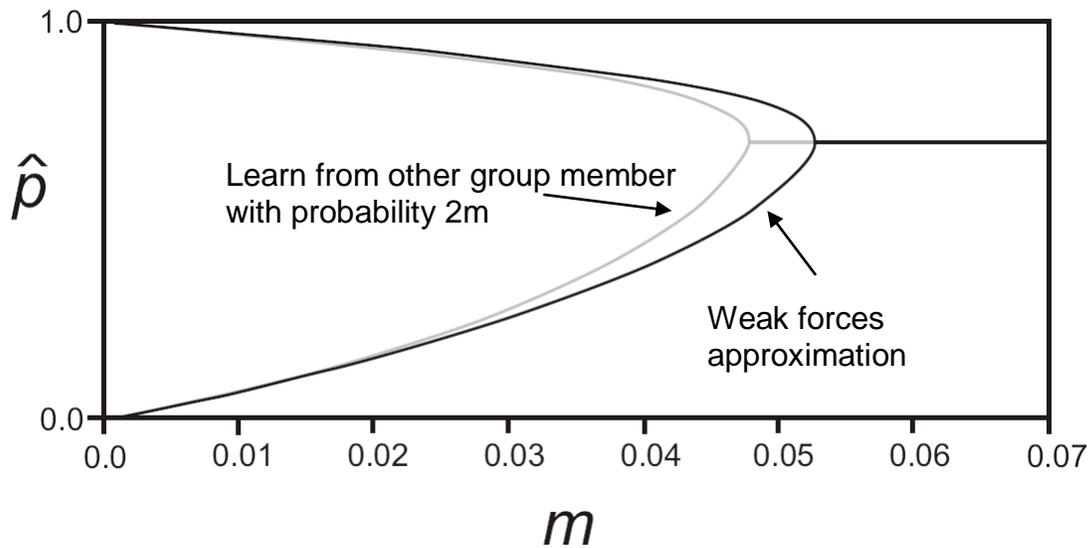


Figure A2. Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, m , for two models. The black line is the analytical solution assuming weak forces. The grey line assumes that individual imitate an individual from their own subpopulation with probability $1 - 2m$ and imitate an individual from the other subpopulation with probability $2m$. The parameters are $\gamma = 0.7$ and $\beta G = 0.5$.

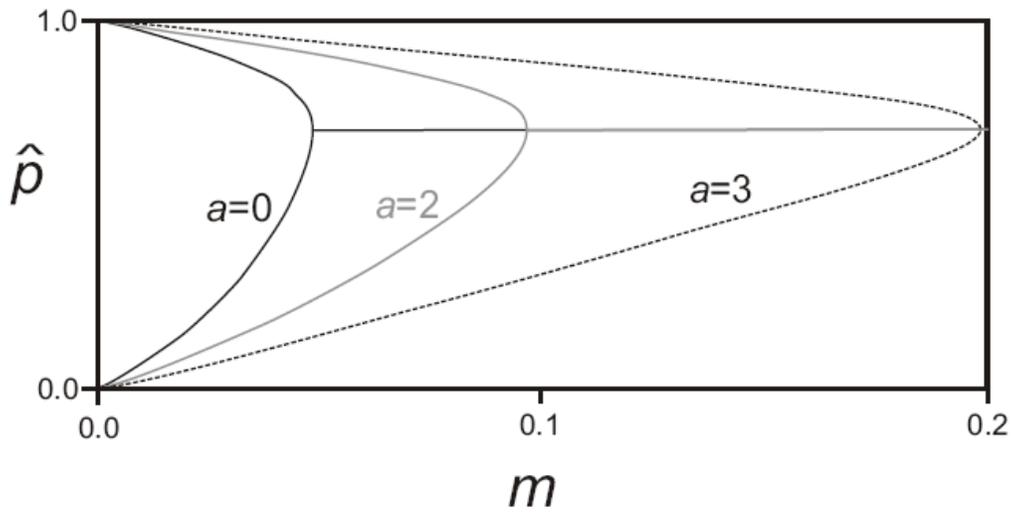


Figure A3: Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, m , for three amounts of success biased migration. In each case imitation occurs first, then migration. The black line is assumes no payoff bias ($a = 0$). The black and dashed lines show greater amounts of success bias in migration. The parameters are $\gamma = 0.7$ and $\beta G = 0.5$.

References Cited

- Barth, F. 1965. *Political Leadership among Swat Pathans*. Toronto: Oxford University Press.
- Bishop, D. T., and C. Cannings. 1978. A generalized war of attrition. *Journal of Theoretical Biology* 70.
- Boyd, R., and P. J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Gadgil, M., and K. C. Malhotra. 1983. Adaptive significance of the Indian caste system: an ecological perspective. *Annals of Human Biology* 10:465-478.
- Gil-White, F. 2001. Are ethnic groups biological 'species' to the human brain? Essentialism in our cognition of some social categories. *Current Anthropology* 42:515-554.
- Gintis, H. 2000. *Game Theory Evolving*. Princeton: Princeton University Press.
- Henrich, J. 2004. Demography and Cultural Evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity* 69:197-214.
- Henrich, J., and R. Boyd. 2002. On Modeling Cultural Evolution: Why replicators are not necessary for cultural evolution. *Journal of Cognition and Culture* 2:87-112.
- Henrich, N. S., and J. Henrich. 2007. *Why Humans Cooperate: A Cultural and Evolutionary Explanation* Oxford: Oxford University Press.
- Hofbauer, J., and K. Sigmund. 1988. *The theory of evolution and dynamical systems : mathematical aspects of selection. London Mathematical Society student texts ; 7*. Cambridge [England] ; New York: Cambridge University Press.
- Mangel, M. 1994. Climate change and Salmonid life history variation. *Deep Sea Research II (Tropical Studies in Oceanography)* 41:75-106.
- McElreath, R., and R. Boyd. 2007. *Modeling the Evolution of Social Behavior*. Princeton Princeton University Press.
- McElreath, R., R. Boyd, and P. J. Richerson. 2003. Shared Norms and the Evolution of Ethnic Markers. *Current Anthropology* 44:122-129.
- Schlag, K. H. 1998. Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-Armed Bandits. *Journal of Economic Theory* 78:130-156.
- . 1999. Which one should I imitate? *Journal of Mathematical Economics* 31:493-527.
- Shuster, S. M., and M. J. Wade, Nature 350, 608–610. . 1991. Equal mating success among male reproductive strategies in a marine isopod. *Nature* 350:608-610.
- Turchin, P. 2003. *Historical Dynamics: Why States Rise and Fall*. Princeton: Princeton University Press.
- Weibull, J. W. 1995. *Evolutionary Game Theory*. Cambridge, Mass: MIT Press.
- Young, H. P. 1998. *Individual strategy and social structure : an evolutionary theory of institutions*. Princeton, N.J: Princeton University Press.

NOTES

¹ We use the word *castes* to remain consistent with Barth's description. However, as Barth points out, these should not be confused with Hindu castes. These occupational castes lack the ritual ascription and assumptions of impurity found in the Hindu systems. Yet, while people readily recognize that some people do change castes, there remains some notions of 'caste impurity,' which are based on Islamic, rather than Hindu, prescriptions and prohibitions. Degrees of impurity depend on the occupation's handling of feces, manure, and dead animals.

² Farmers need not remain with the same landowners, and do not live on the land they farm.

³ This nuance fits our model. The value of m to maintain stratification depends on the payoff differences observed between competing strategies. When alternative strategies are similar in payoffs, there is less incentive to imitate, so high values of m still permit stratification to be maintained. However, if payoffs between strategies are very different, m must be much lower to still maintain stratification. This suggests that stratification is unlikely to exist if, for example, the lowest caste was only permitted to socialize with and marry the highest caste, and vice versa.

⁴ By restricting reproduction to certain types of individuals, social insects have achieved something that parallels social stratification. Both humans and social insects seem to have solved structurally similar problems by 'tricking' their way around the Bishop-Canning theorem. Eusocial insects use kinship, highly restricted reproduction, and often particular genetic transmission systems. Humans used their second system of inheritance, culture.