Division of Labor, Economic Specialization and the Evolution of Social Stratification

Joseph Henrich and Robert Boyd

Version: 20. October 21, 2007

Joseph Henrich, Department of Psychology and Department of Economics, University of British Columbia, Vancouver, B.C. Canada. (joseph.henrich@gmail.com)

Robert Boyd, Department of Anthropology, University of California, Los Angeles, CA 90254 (robert.t.boyd@gmail.com)

ABSTRACT

This paper presents a simple mathematical model that shows how economic inequality between social groups can arise and be maintained even when the only adaptive learning processes driving cultural evolution increases individual's economic gains. The key assumption is that human populations are structured into groups, and that cultural learning is more likely to occur within groups than between groups. Then, if groups are sufficiently isolated and there are potential gains from specialization and exchange, stable stratification can sometimes result. This model predicts that stratification is favored, *ceteris paribus*, by (1) greater surplus production, (2) more equitable divisions of the surplus among specialists, (3) greater cultural isolation among subpopulations within a society, and (4) more weight given to economic success by cultural learners. We also analyze how competition among societies, both egalitarian societies and those with differing degrees of stratification, influences the long-run evolution of the institutional forms that support social stratification. In our discussion, we illustrate the model using two ethnographic cases, explore the relationships between our model and other existing approaches to social stratification within anthropology, and compare our model to the emergence of heritable divisions of labor in other species.

"…explanations of the origins of institutionalized social inequality and political privilege must resolve the central paradox of political life—why people cooperate with their own subordination and exploitation in non-coercive circumstances."

J. E. Clark and M. Blake (1994: 17,
Smith and Choi 2007)

Explaining social stratification has been an important focus of social thought at least since the Enlightenment (Rousseau 1755). Anthropologists and sociologists, in particular, have defended a wide variety of theories that link economic specialization, a division of labor, and the emergence of socially stratified inequality since the birth of their discipline at the end of the 19th century (e.g., Arnold 1993, Durkheim 1933, Fried 1967, Gilman 1981, Lenski, Lenski, and Nolan 1991, Ruyle 1973, Service 1975). Here, we focus on understanding "stratification," as the emergence and persistence of institutionalized economic differences between social groups.

Inequality is ubiquitous. Within every human society, individuals of different ages, genders or abilities receive different shares of the overall economic output. In some societies these differences are glorified and exaggerated, while in others they are more subtle and often go unacknowledged, yet they are always there (Fried 1967). Our puzzle, however, is not this ubiquitous inequality among individuals, but persistent inequality among social groups, like classes, castes, ethnic groups, and guilds. Because such groups include a wide sampling of people, it is not plausible that inequality results from innate differences in size, skill, or the like among the individuals who make them up (Richerson and Boyd 2005a). Instead, these differences must result from something that individuals acquire as a consequence of group membership. This leads to the obvious question: why don't people on the wrong side of such inequalities adopt the skills, practices, behaviors, or strategies of the people in the group or groups who are getting a disproportionately large share of the economic benefits produced by a society?

Scholars have given at least three kinds of answers to this question (Durkheim 1933). Many authors deny the paradox, arguing either that people are systematically deceived about their interests, for example, as a result of elite propaganda or are coerced into submission (Cronk 1994, DeMarrais, Castillo, and Earle 1996, Kerbo 2006). Other authors (Axtell, Epstein, and Young 2001, Boone 1992, Smith and Choi 2007) have argued that exogenous differences between individuals in different groups can be amplified by a number of different social or evolutionary processes to generate persistent inequality between groups. For example, either correlated asymmetries like access to high quality resources like land or uncorrelated ones like skin color or dialect can be used to coordinate interactions that lead to systematically unequal outcomes (Axtell, Epstein, and Young 2001, Smith and Choi 2007). If investments in schooling or other forms of social capital are subject to externalities, then individual choice may lead to self-perpetuating differences between groups in investment and income (Lundberg and Startz 1988). Finally, if social inequality enhances group success, then either cultural or genetic group selection can explain the persistence of social inequality if these processes are strong enough. The genetic version of his process explains hereditary inequality in colonies of eusocial animals such as termites or naked mole rats (Oster and Wilson 1979). While each of these solutions to the puzzle has its partisans, the longevity of the debate suggests that none are completely satisfactory.

Here we present a novel model for the emergence of social stratification, without coercion, deception or exogenous sources of group differences. We assume that people acquire economic strategies from others via cultural learning, which includes observational learning, imitation and teaching. Of course, cultural learning is a complex process, and as a consequence cultural change need not lead to the spread of economically beneficial traits. However, to make the model as stark as possible we assume that people are predisposed to learn from economically successful people, and that this bias leads to the spread of cultural variants that increase individual economic success. We show that even when only economic success matters, stable inequality can result. The reason is that we also assume that the overall population is subdivided into social groups, and that people tend to learn more often from members of their own group, and less from other groups. This means that relative success *within* social groups, not absolute economic success is what matters. Using a simple model, we show that these processes can give rise to a stable, culturally heritable division of labor even when there is a substantial exchange of ideas or individuals among sub-groups, and despite the fact that the only force shaping cultural variation is an adaptive learning mechanism that myopically maximizes payoffs/self-interest. We also show that this can give rise to a process of cultural group selection, in which groups that establish certain forms of unequal social exchange may outcompete egalitarian societies, and those with less competitive forms of inequality.

The assumption that people imitate the successful is supported by empirical data from across the social sciences. Research shows that success and prestige biased cultural learning influences preferences, beliefs, economic strategies, technological adoptions, skills, opinions, suicide, and norms (Henrich and Henrich 2007, Richerson and Boyd 2005b). For example, recent laboratory studies in economics using monetary performance-based incentives indicate that people rely on imitating beliefs (Offerman and Sonnemans 1998), economic strategies (Alpesteguia, Huck, and Oeschssler 2003, Pingle 1995, Pingle and Day 1996, Selten and Apestegula 2002), and behaviors (Kroll and Levy 1992) of particularly *successful individuals* in social interactions. This work is consistent with experimental work in psychology and field data from sociology and anthropology showing that both children and adults have a powerful tendency to learn a wide variety of things from successful individuals (theory and evidence summarized in Henrich and Gil-White 2001).

Cultural transmission is complex and a number of processes not included in this model are undoubtedly important (Henrich and Gil-White 2001, Henrich and McElreath 2003, Richerson and Boyd 2005b). We ignore these complications in order to focus on the central puzzle: Can cultural evolution lead to stratified inequality when the *only* evolutionary process that creates cultural change is one that only leads to the spread of individually more successful beliefs and practices. This sets the bar higher than it might otherwise be.[1]

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 (2004b) and Boyd & Richerson (1985).

The assumption that people are subdivided into social groups, and that group members tend to learn from each other also has empirical support. Across the world, whether they be villages, neighborhoods, ethnic enclaves, bands, or clans, people tend to live in and preferentially associate with local aggregations. Field studies of social learning suggest that these groups are often the main locus of cultural transmission (Fiske 1998, Lancy 1996).

A MODEL OF THE EVOLUTION OF SOCIAL STRATIFICATION

Here we describe an evolutionary game-theoretical model that formalizes these assumptions. We first lay out the structure of the model and describe the main results. Then, we discuss how the model's parameters affect the evolution of inequality. Consider a large population of individuals. During each time interval, each individual interacts with one other individual in an exchange using one of two possible strategies that we have labeled High (H) and Low (L). Payoffs to players are determined jointly by strategies deployed by the two interacting individuals, as shown in Table 1. If both interacting individuals use the same strategy, each receives the baseline payoff, ω. However, if the interactants use different strategies, the individual using strategy H receives a payoff of $\omega + \gamma G$ while the individual using Strategy L receives $\omega + (1 - \gamma)G$. Thus, G can be thought of as the 'surplus' created by a division of labor, specialization, or some other kind of complimentary element in the interaction. The parameter γ gives the proportion of the surplus that goes to the individual playing the strategy H, and $1 - \gamma$ gives the proportion of the surplus going to the individual playing strategy L. We assume that γ ranges only from 0.5 (an equal split of G) to 1.0 (the H player gets it all) with no loss of generality.

[Table 1 about here]

This large population is subdivided into two subpopulations, named subpopulation 1 and subpopulation 2. This structure could result from anything that patterns social interactions, including distance, geographical barriers like rivers or mountains, or social institutions like villages, clans, or ethnic groups. The frequency of individuals in subpopulation 1 using strategy H is labeled p_1 and the frequency of individuals in subpopulation 2 using H is labeled p_2. To allow for the possibility that subpopulation membership affects patterns of social interaction we assume that with probability δ an individual is paired for an interaction with a randomly selected individual from the other population (individuals from subpopulation 1 meet those from subpopulation 2, and vice versa), and with probability $1 - \delta$ the individual meets someone randomly selected from their home subpopulation. When $\delta = 0$, individuals interact only with others from their own subpopulation; when $\delta = 1$, individuals always interact with individuals from the other subpopulation; and when $\delta = 0.5$ interaction occurs at random with regard to the overall population.

Next, cultural learning and mixing between subpopulations takes place. Mixing and learning can happen in two different ways. First, individuals could learn their strategy from someone in their home (natal) subpopulation with a probability that is proportional to the difference between the learner's payoff and the model's payoff (see the appendix for details), and then mix by physically moving between subpopulations, carrying their ideas along. This is modeled by assuming that there is a probability *m* that people migrate from one subpopulation to the other. Alternatively, one can ignore physical movement, and assume that individuals usually learn their strategy from someone in their home population but sometimes use a model from the other subpopulation, in either case acquiring the strategy with a probability proportional to the difference in their payoffs (as above). In this case, mixing occurs as ideas flow from between subpopulations. In the appendix, we show that this is equivalent to assuming that learners observe and learn from a model in the other subpopulation with probability $2m$, and learn from someone in their home subpopulation with probability $1 - 2m$. These different life histories lead to the same model, so we will label this parameter the *mixing* rate (m).

Using standard tools from cultural transmission and evolutionary game theory (Gintis 2000, McElreath and Boyd 2007) we can express the *change* in the frequency of individuals playing the *High* strategy (H) in subpopulation 1 in one time step, Δp_1, as:

$$\Delta p_1 = \underbrace{p_1(1-p_1)\beta(\pi_{H1}-\pi_{L1})}_{\text{Success-biased transmission}} + \underbrace{m(p_2-p_1)}_{\text{Migration}} \tag{1}$$

The derivation of (1) in the appendix assumes that changes in trait frequencies during learning and migration are sufficiently small that the order of learning and migration do not matter. We use simulations to show that relaxing this assumption does not influence our qualitative findings.

Recursion (1) contains two parts: (a) the effects of success-biased transmission, and (b) cross-subpopulation transmission and/or the movement of people between the subpopulations. The symbol π_{H1} gives the expected payoff received by individuals playing the *H*-strategy from subpopulation 1, while π_{L1} gives the expected payoff to individuals from subpopulation 1 playing the *L*-strategy. These payoffs, π_{H1} and π_{L1}, depend on the production from exchange/specialization (G), the division of this 'surplus' production (γ), the current frequency of strategies in each subpopulation (p_1 and p_2), and on the probability of interacting with an individual from the other subpopulation (δ):

$$\pi_{H1} = \delta\overbrace{\left[(1-p_2)(\omega+G\gamma)+p_2\omega\right]}^{\text{Payoff playing subpopulation 2}} + (1-\delta)\overbrace{\left[p_1\omega+(1-p_1)(\omega+G\gamma)\right]}^{\text{Payoff playing subpopultion 1}} \tag{2}$$

$$\pi_{L1} = \delta\left[p_2(\omega + G(1-\gamma)) + (1-p_2)\omega\right] + \quad (1-\delta)\left[p_1(\omega + G(1-\gamma)) + (1-p_1)\omega\right] \qquad (3)$$

These expressions have units of payoff. The parameter β in (1) is a constant that scales differences in payoffs into changes in the frequency of cultural variants and can be thought of as the degree that a learner is influenced by any particular learning event.

The first terms on the right hand side of equations (2) and (3) are the probabilities of interacting with an individual from the other subpopulation (δ) multiplied by the expected payoff received by an individual playing the *H* strategy, given the probability of meeting *H* (p_2) and *L*-strategies ($1 - p_2$) from subpopulation 2. The second term is the probability of interacting with an individual from the players' own subpopulation (subpopulation 1, in this case, $1 - \delta$) multiplied by the expected payoff to an *H* given the chance of meeting either an *H*- or and *L*- strategy (p_1 and $1 - p_1$, respectively). The expected payoff to an individual playing the *L*-strategy from subpopulation 1, π_{L1}, can be explained similarly.

Following a parallel development, the change in the frequency of *High* strategies in subpopulation 2, Δp_2, can be expressed as:

$$\Delta p_2 = p_2(1-p_2)\beta(\pi_{H2} - \pi_{L2}) + m(p_1 - p_2) \qquad (4)$$

where, as above,

$$\pi_{H2} = \delta\left[(1-p_1)(\omega + G\gamma) + p_1\omega\right] + (1-\delta)\left[p_2\omega + (1-p_2)(\omega + G\gamma))\right] \qquad (5)$$

$$\pi_{L2} = \delta\left[p_1(\omega + G(1-\gamma)) + (1-p_1)\omega\right] + (1-\delta)\left[p_2(\omega + G(1-\gamma)) + (1-p_2)\omega\right] \qquad (6)$$

Determining the Equilibrium Behavior of the Model

Equations (1) and (4) describe how social behavior, population movement, and social learning affect the frequency of the two behaviors in each subpopulation over one time step. By iterating this pair of difference equations, we can determine how the modeled processes shape behavior in the long run. Of particular interest are the stable equilibria. The equilibria are combinations of p_1 and p_2 which according to (1) and (4) lead to no further change in behavior. An equilibrium is locally stable if the population will return to that equilibrium even if perturbed a bit. It is unstable if small shocks cause the population to evolve off to some other configuration.

The system can be characterized by one of two types of stable equilibrium conditions. There are *egalitarian equilibria* in which each of the two subpopulations has the same mix of individuals using the *High* and *Low* strategies. This situation could be interpreted as each individual using a mixed strategy of H and L (individual lack task specialized skills). At such egalitarian equilibria, the *average* payoff of all individuals is the same no matter what subpopulation they are from. There are also *stratified equilibria* in which a majority of individuals in one subpopulation play the *High* strategy while most of the individuals in the other subpopulation use the *Low* strategy. In this case, the average payoff to the subpopulation that consists mostly of *H*s is higher than the average payoff of individuals in the subpopulation consisting of mostly *L*s. Our work below indicates that depending on the parameters—i.e., the values of G, m, γ, β, and δ—either one *or* the other type of stable equilibrium exist, but never both.

We begin by assuming that individuals always interact with someone from the other subpopulation ($\delta = 1$), which allows us to derive some instructive analytical results, and then we use a combination of computer simulations and more complex, and therefore less transparent, analytical solutions to show that the simpler analytical results are robust. Initially assuming $\delta = 1$ makes sense because success-biased learning, or self-interested decision-making, will favor higher values δ by the majority of the subpopulation anytime there are differences in the relative frequencies of H and L in the subpopulations. Lower values of δ are never favored. However, it is important to explore values of δ less than one, since real world obstacles may prevent δ from reaching one.

To find the equilibrium values we set Δp_1 and Δp_2 equal to zero and solve. This yields two interesting solutions:[2] An egalitarian equilibrium where:

$$\hat{p}_1 = \hat{p}_2 = \gamma \tag{7}$$

This tells us that the frequency of *H*-strategies in each subpopulation at equilibrium will be equal to the fraction of the surplus received by *H*-strategies during an interaction (this result holds for $0.5 \leq \delta \leq 1$).

And, a stratified equilibrium at which:

$$\hat{p}_1 = \frac{2m(G\beta\gamma - 2m)}{G\beta(1-\gamma)(G\beta\gamma - 2m) + \sqrt{-G\beta(2m - G\beta(1-\gamma))(2m - G\beta\gamma)(2m - G\beta(1-\gamma)\gamma)}} \tag{8}$$

$$\hat{p}_2 = \frac{2m(2m - G\beta\gamma)}{G\beta(1-\gamma)(2m - G\beta\gamma) + \sqrt{-G\beta(2m - G\beta(1-\gamma))(2m - G\beta\gamma)(2m - G\beta(1-\gamma)\gamma)}} \tag{9}$$

The stratified equilibria are locally stable and the egalitarian equilibrium is unstable if

$$G\beta\gamma(1-\gamma) > 2m \tag{10}$$

If (10) is not satisfied, only the egalitarian equilibrium exists and it is stable.

These expressions tell us that the system has two qualitatively different types of equilibrium states: either the egalitarian equilibrium is stable, and all cultural evolutionary roads lead to that equilibrium; or, it is unstable and all evolutionary pathways lead to stratification. We refer to the point at which $G\beta\gamma(1-\gamma) = 2m$ as the *stratification threshold*. To help clarify this, Figure 1 shows the location of the two equilibria points as functions of the mixing rate, *m.* The flat horizontal line denotes the location of the egalitarian equilibrium ($\hat{p}_1 = \hat{p}_2 = \gamma$). This equilibrium always exists, but is not always stable. For the stratified equilibrium, the curves labeled $\hat{p}_1$ and $\hat{p}_2$ give the equilibrium frequency of *H*-strategies in subpopulations 1 and 2. When mixing rates are low, the stratified equilibrium exists and is stable and the egalitarian exists but is unstable. As the flow of strategies or people between subpopulations (*m*) increases the frequencies of *H*-strategies in the two subpopulations converge toward each other. Stratification disappears at exactly the point at which the frequency of individuals adopting the *H*-strategy becomes equal in both populations. At higher rates of mixing only the egalitarian equilibrium exists and it is stable.

[Figure 1 here]

When the stratified equilibria are stable, the existence of the unstable egalitarian equilibrium does not influence the final location of the evolving system. However, this unstable equilibrium does influence the system's dynamics by acting as an unstable attractor. For example, under conditions in which only the stratified equilibria are stable and both subpopulations begin with only a few H's in each one, p_1 and p_2 will initially race toward the egalitarian equilibrium, only to veer off at the last minute and head for their final destination—stratification. [3]

The average payoff in each of the subpopulations at the stratified equilibrium is:

$$\hat{\bar{\pi}}_1 = \hat{p}_1\hat{\pi}_{H1} + (1 - \hat{p}_1)\hat{\pi}_{L1} \tag{11}$$

$$\hat{\bar{\pi}}_2 = \hat{p}_2\hat{\pi}_{H2} + (1 - \hat{p}_2)\hat{\pi}_{L2} \tag{12}$$

Substituting in expressions (2), (3), (5) and (6), evaluated at location $\hat{p}_1$ and $\hat{p}_2$ as expressed by equations (6) and (7), into (11) and (12) gives the average payoff in each subpopulation. We'll use a ratio of the average payoffs (13) to summarize the inequality present in any particular equilibrium condition.

$$\Gamma = \frac{\hat{\bar{\pi}}_2}{\hat{\bar{\pi}}_1} \tag{13}$$

How the Parameters Affect Stratification

To explore the implications of this model we will examine how each of the parameters—G, m, β and γ—influences: (1) the emergence of a stable stratified equilibrium vs. an egalitarian equilibrium, and (2) the degree of inequality in average payoffs between the two subpopulations—stratification may exist, but with greater or lesser degrees of inequality between the subpopulations.

Mixing rate***, m,*** measures the flow of ideas or people between the two subpopulations. Several factors might influence *m*. If populations are spatially separated, *m* would likely be smaller, while if they are interspersed, *m* would, *ceteris paribus*, tend to be bigger. However, we know from psychology that social learning is influenced by both symbolic (accent, language, dress, etc.) and other kinds of phenotypic markings, often related to such phenomena as ethnicity, race, caste and class. If subpopulations are ethnically/symbolically marked, *m* can be small even in an interspersed population.

The larger the value of *m*, the more difficult it is to produce stable social stratification. Moreover, if the stratified equilibrium does exist, increasing *m* reduces the degree of inequality. The effect on the *existence* of a stratified equilibrium is completely summarized in expression (10), and can be visually appreciated in Figure 1. Figure 2 shows the effect of *m* on the degree of inequality ($\Gamma = \hat{\bar{\pi}}_2 / \hat{\bar{\pi}}_1$). Increasing *m decreases* the degree of inequality between the subpopulations, until a stable egalitarian society emerges. Conversely, a decline in *m* eventually, *ceteris paribus*, brings stratification and inequality into existence. This effect is especially dramatic when *G* is small. For values of *m* larger than the stratification threshold, there is no inequality; all individuals have the same expected payoff. Below, we discuss how and why *m* might vary among societies, as well as how it may be influenced by larger-scale sociocultural evolutionary processes.

[Figure 2 about here]

As noted earlier, *m*, the mixing rate, could represent either the physical movement of individuals or the flow of ideas. When the rates of change are small enough that the order of transmission and mixing don't matter, either assumption yields recursions (1) and (4). However, when rates of change are higher, the two models yield different sets of recursions. As shown in the appendix, computer simulations indicate that the different models have very similar qualitative properties.

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 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 solution discussed in the appendix indicate that increasing the parameter *a* **increases—**not decreases**—**the amount of stratification. 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.

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 (Boyd and Richerson 1987, 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 (see our case examples below and 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.

Surplus production from the division of labor, *G*, is the production created by the economic exchange of specialized skills, resources, knowledge or talents. *G* depends on technology, know-how, environment and ecological resources/constraints, norms of interaction and transaction costs.

Greater surplus production (G) expands the conditions favoring stable stratification and increases the degree of inequality if stratification is already stable. This means that, *ceteris paribus*, technologies, practices and forms of organization that favor greater production (through exchange and specialization) favor increasing degrees of stratified inequality. If G increases, the stratified equilibrium is more likely to be stable—and the effect is linear. Both the effect of G on the existence of a stratified equilibrium and on the degree of inequality can be seen in Figure 2. The dashed vertical line on Figure 2a shows that, for the same value of m (and γ, β), the payoff inequality is greatest for $G = 80$. The fact that the dashed line does not cross the $G = 10$ curve implies that no stratified equilibrium exists, so $\Gamma = 1$ and members of different subpopulations do not differ in average payoffs.

Inequality of the division of the proceeds of interaction, *γ*, specifies the proportion of G that is allocated to the individual playing the *H*-strategy. The parameter γ might be influenced by resource availability, skill investment (high skill vs. low), supply and demand, and/or local customs. Changing γ has complicated effects on stratification and the degree of inequality. First, recall that the stratification threshold (10) depends on $\gamma(1-\gamma)$. The greater this product, the more likely stratification is to emerge. This means that the more *equitable* individual-level divisions are, the more likely stable *stratified* equilibria are to emerge. This also indicates, perhaps non-intuitively, that a sudden increase in γ in a stratified society may cause a shift *from* a stratified equilibrium to an egalitarian situation—sociocultural evolution will drive the system to the egalitarian situation. This can be seen in Figure 2 by comparing 2a and 2b. Looking at the stratification threshold for each value of *G* (where the curves intersect the x-axis), we see that when $\gamma = 0.7$ (Figure 2a) stratification persists up to higher values of *m* than when $\gamma = 0.9$. This means that more equitable divisions of surplus (γ closer to 0.50) at the individual level favor stable *inequality* at societal level. Remember, this is not the same as saying the *degree* of inequality is higher.

There is also a potent interaction between *G* and γ on the degree of inequality, *when* the stratification equilibrium exists. Compare Figures 2a and b. When $\gamma = 0.9$, an increase in *G* from 50 to 80 creates a substantial increase in Γ compared to the effect on Γ of the same increase in *G* when $\gamma = 0.7$. This effect can be further observed on Figure 3, which plots Γ against γ for several *G* values. This plot emphasizes and informs an effect we mentioned earlier: greater *G* permits stratification to be maintained at greater γ values. Together, higher *G* and higher γ drastically increase the degree of inequality observed (Γ).

[Figure 3]

Social learning scale parameter, *β*, converts differences in payoffs between the strategies observed by social learners into probabilistic changes in individual behavior. For this reason, β must be less than $1/(G\gamma)$, and greater than zero. Psychologically, β can be thought of as the degree to which individuals' behavior is influenced by a particular learning event. If $\beta \ll 1/G\gamma$ then the effects of one incident of social learning are small and cultural evolution will proceed slowly. Since β scales payoffs into learning, β is in the reciprocal units of G. We will have more to say about the kinds of real world, cross-society differences that might influence β in the discussion below. From the perspective of what we have derived so far, β consistently occupies a position alongside *G,* and thus change have the same effects as *G*.[4]

Mixed interaction reduces stratification

When the restriction that individuals only interact with members of the other subpopulation (captured in the assumption that $\delta = 1$) is relaxed, and we allow δ to vary such that $0.5 \leq \delta \leq 1$, the location of the egalitarian equilibrium is $\hat{p}_1 = \hat{p}_2 = \gamma$, the same as above when $\delta = 1$. The stability of that equilibrium is captured by this inequality:

$$\beta G\gamma(1-\gamma)(2\delta-1) < 2m \tag{14}$$

When inequality (14) is satisfied, the egalitarian equilibrium is stable. Note that the only change in (14), from (10), is the term $(2\delta - 1)$. This tells us that δ values less than 1 increase the stability of the egalitarian equilibrium. While we have not been able to analytically solve our system of equations for the location and stability of the stratified equilibrium, extensive simulations indicate that, as above, either the egalitarian equilibrium exists and is stable, OR the stratified equilibrium exists and is stable (and the egalitarian equilibrium exists but is unstable). This means that our condition (14) likely provides the conditions for the existence of a stable stratified state, and all of our above analysis regards G, γ, m and β apply to this case. A user-friendly version of the program we used to establish this is available on the web, see endnote 3.

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: 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. The means that individuals, using a range of adaptive learning mechanisms, 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. The answer is that constraints on economic interaction inhibit the emergence of stratification.

Stratification Increases Total Group Payoffs

Our results so far indicate that restricted migration gives rise to stratified societies with more economic specialization. This means that, all other things being equal, more stratification will lead to higher average production than in egalitarian societies. The total average payoff of societies at a stratified equilibrium is (assuming $\delta = 1$):

$$\hat{\pi} = 2\omega + G\left(\hat{p}_1(1-\hat{p}_2) + \hat{p}_2(1-\hat{p}_1)\right) \tag{15}$$

The first term is the baseline payoff achieved by all individuals, and the second is the surplus created by exchange weighted by a term that measures the amount of coordination. For example, if $\hat{p}_1 = 1$, $\hat{p}_2 = 0$ and $\delta = 1$, there is no mis-coordination and the population gets all surplus. More mis-coordination causes some of the surplus to be lost and the average payoffs to decrease.[5]

Substituting in the equilibrium values of p_1 and p_2 from equations (8) and (9) and setting $\omega = 0$, yields:

$$\hat{\pi} = \frac{(2m - G\beta(1-\gamma))(G\beta\gamma - 2m)}{\beta(m - G\beta\gamma(1-\gamma))} \tag{16}$$

This result suggests that cultural group selection might affect social stratification. It is plausible that factors like migration and surplus that are represented by model parameters are influenced by culturally transmitted beliefs and values (e.g., norms, institutional forms, ethnicity, etc.). If so, then different societies may arrive at different stable equilibria, including stratified equilibria that differ in total payoffs. The existence of different societies at stable equilibria with different total payoffs creates the conditions for cultural group selection (CGS: Boyd and Richerson 2002, Boyd and Richerson 1990, Henrich 2004a).[6] For cultural group selection to operate, group payoffs ($\hat{\hat{\pi}}$) must influence the outcomes of competition among societies. Warfare, for example, might cause societies with higher payoffs ($\hat{\hat{\pi}}$) to proliferate because economic production can supply more weapons, supplies, allies, feet-on-the-ground, skilled warriors, etc. Alternatively, extra production could lead to faster relative population growth. Or, more indirectly but perhaps more importantly, people tend to imitate people with higher payoffs. This means that when people from a poorer society meet people from a richer society, there will be a tendency (*ceteris paribus*) for cultural traits and institutions to flow from rich societies to poorer ones, as an unintended result of contact between societies and the nature of human learning psychology (Boyd and Richerson 2002). This last process has been widely observed among ancient polities (Renfrew and Cherry 1986), contemporary business organizations (DiMaggio 1994), and in ethnographically known small scale societies (Atran, Ross, and Coley 1999, Boyd 2001, Wiessner 2002). Naroll and co-workers present data which indicates that imitation-driven processes between societies account for more cultural change than does warfare (Naroll and Divale 1976, Naroll and Wirsing 1976).

Assuming cultural group selection leads to the spread of societies with higher average payoffs we can use comparative statics to predict the directional effect of this process on our five parameters and on the frequency of stratified societies (with economic specialization).

1) Cultural group selection favors economic specialization and stratification equilibria over egalitarian equilibria because stratification yields the surplus benefits of economic specialization. The highest total payoff a non-stratified society can achieve is $2\omega + G/2$ while a stratified society can achieve $2\omega + G$.

2) Cultural group selection will tend, to the degree possible, to drive δ to 1 and m to zero. Higher values of δ and lower values of m maximize the coordination of strategies, and the economic benefits of specialization.

3) Cultural group selection favors greater values of G—this would include the technologies,

skills and know-how that increase economic production and increase between-group competitiveness. Interestingly, modeling work on the evolution of technological complexity (Henrich 2004b, Shennan 2001) indicates that larger, denser social groups will be able to maintain greater levels of technological complexity, knowledge, and skill. This implies that population characteristics may be indirectly linked to stratification and economic specialization: bigger, more culturally interconnected populations generate more productive technologies and skills, which lead to greater values of *G*. Higher values of *G* lead to stratification and higher group payoffs, $\hat{\pi}$.

4) For β, bigger is better. Institutions, beliefs and values that lead individuals, in their cultural learning, to weigh economically successful members of their subpopulation more heavily will be favored by cultural group selection.

5) For γ, the situation is particularly interesting. Figure 4 plots the total payoff for the stratified equilibrium against γ for three values of G ($\beta = 0.01$, $m = 0.01$, $\delta = 1$). The strength of cultural group selection is equal to the slope of the line ($\partial\hat{\pi}/\partial\gamma$). This means that for the extreme values of γ (highest inequality), cultural group selection strongly favors societies with less inequality (lower γ), while for moderate and low levels of individual-level inequality in exchange, cultural group selection only weakly favors greater equality. In the long-run, cultural group selection favors $\gamma = 0.5$ (equality) *with* economic specialization and stratification.

These predictions can be verified by differentiating (16) with respect to each of the parameters. This means that we assume that only one parameter varies at a time. It's probable that some of these parameters are causally interconnected, probably in different ways in different specific circumstance, so a more complete analysis of a particular historical situation will require consideration and modeling of the interconnections among parameters.

DISCUSSION

Human populations are frequently divided into subpopulations that are *partially* culturally isolated. In this paper we describe a very simple mathematical model that shows how such partial isolation can give rise to stable social stratification even when adaptive individual learning is guided only to improve payoffs. While the use of such intentionally simplified models is commonplace in other disciplines like economics, ecology, and evolutionary biology that focus on complex historical subject matter, it is unusual in anthropology and archaeology (exemplary exceptions include Eerkens and Lipo 2005, Neiman 1995, Shennan 2001, Ugan, Bright, and Rogers 2003), and we know that many will be uncomfortable with the approach. How can such a simplified abstract model capture all the richness of human cultural change? The answer, of course, is that it is not intended to capture the richness of any particular evolutionary sequence. Rather, we seek to understand the effect of a particular set of cultural evolutionary processes. We begin with several empirically grounded assumptions: Human populations are subdivided, people tend to learn from successful others in their local group, and division of labor can create economic surplus. By analyzing a simple model incorporating these assumptions we derive qualitative conclusions about the resulting evolutionary process. Some are not surprising: increased mixing between subpopulations decreases stratification; but, others are not obvious: increasing the surplus available tends to increase the degree of stratification. The empirical payoff from such simplified models is a clearer qualitative understanding of a set of generic processes that then can be applied, along with our understanding of other unmodeled processes, or processes modeled elsewhere, such as intergroup competition (cultural group selection).

Below, we will first apply our findings to an ethnographic case example, then briefly discuss how it relates to, or informs, some of the existing work on the evolution of social stratification, and finally, close by clarifying why this model of cultural evolution gives rise to heritable stratification while models of genetic evolution usually do not.

Interpreting an Ethnographic Example

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.[7] 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 economicallyadjacent 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).[8] 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.[9] 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.

Relationships to Existing Theoretical Approaches

This model does not capture the only route to stratification, and it certainly only contributes to understanding one aspect of the more general problem of the evolution of societal complexity. What it does do is to show how economic specialization and cultural differentiation can sometimes produce stratified inequality. We believe that this model can inform a number of existing theoretical approaches, especially when the model is seen as setting a foundation for the evolution of a political elite by supplying a "surplus" (γG—the increased production created by specialization) that could be used for such things as building monuments, employing armies and labor, constructing boats and fortresses, purchasing capital equipment, etc. In this section we briefly highlight some relationships between our models and existing anthropological work on stratified inequality.

Economic Specialization, Exchange and Surplus. Our model focuses on how economic specialization can lead to both additional production and stratified inequality when people occupying different economic roles are culturally isolated. As one might call this extra production (G) a 'surplus' that derives from an exchange of some kind, our model illuminates how and when economic specialization and exchange can lead to stratified inequality. Unlike previous verbal depictions, we show that simply generating some surplus (G) is insufficient. More surplus makes it more likely that the basic conditions for stratification will be satisfied, but in the end everything depends on the confluence of several factors. If populations are homogeneous and well-mixed, for example, stratification will never emerge no matter how large G is.

Population Pressure, Intensification, and Social Stratification. Theories of societal evolution often emphasize population pressure as the prime mover (Johnson and Earle 2000, Netting 1990). In these cases economic specialization and social stratification are seen as responses to the economic need for intensification (declining productivity yields per capita). While there are good reasons to doubt that population pressure causes stratification (Richerson, Boyd, and Bettinger 2001), it is empirically true that denser, populations tend to co-occur with stratification and inequality (Naroll 1956). One possible explanation for this correlation is that larger, more densely connected populations are likely to produce faster cultural evolutionary rates for sophisticated technology, complex skills and knowledge (Henrich 2004b, Shennan 2001), which in turn will generate more surplus (i.e. higher values of *G*) (Carneiro and Tobias 1963), which favors stratification and permits greater degrees of inequality. Higher levels of productivity often support an even larger populations, which in turn, will support a higher equilibrium level of technological/skill sophistication. Such a feedback loop could explain the observed relationship between stratification and population variables.

Conflict and Circumscription Theories: Warfare between social groups does not cause stratification in our model. However, cultural group selection will spread *certain kinds* of stratification through inter-group competition. In this vein, stratified specializations allow for both warrior and weapon-maker castes, which have benefits in violent, competitive socio-ecologies. Further, we expect the competitive interaction among circumscribed societies to favor those combinations of parameters that maximize overall group production, thereby freeing more of the population for military participation and providing more resources (e.g. food, weapons, etc.). Over time, these higher-level processes should favor greater joint production (*G* increasing), more economic specialization and stratification (to the degree that it improves production), less flow of strategies (m decreasing), more well-defined patterns of interaction between subpopulations (δ increasing) and, over longer time scales, greater equality between subpopulations ($\gamma \rightarrow 0.5$).

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 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 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 fitness. 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.[10] 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 ones), religious functions (14 different castes), landownership, and foraging (just to name a few). Among the foraging castes alone, some castes 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.

CONCLUSION

We know that many anthropologists will object to this model as abstract and oversimplified. However, we believe its simplicity is an advantage. By focusing on the cultural dynamics that result from simple assumptions about social learning, economic interaction, and population structure, we are able to draw clear, qualitative lessons from the model. These lessons can then be applied to a wide range of specific cases including social phenomena such as social classes, ethnic occupations, castes, guilds, and occupation-based clan divisions by asking how specific historical developments like agriculture, irrigation, warfare, steel plows, craft specialization influence the parameters of the model, and thereby led to stratification or greater inequality in an already stratified society. Such simple models have proven useful in a number of disciplines that study complex, diverse phenomena such as ecology, evolutionary biology, and economics. We believe that they are beginning to prove useful in anthropology and archaeology as well.

APPENDIX

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

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

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).

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.

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

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

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

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 ½ 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:

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

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

We implemented and explored success biased physical migration as described in the text. This migration was combined with the within-group learning model described above. 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. As noted in the main text, 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]

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.

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:

$$\hat{\pi} = p_1(\pi_{H1}) + (1 - p_1)\pi_{L1} + p_2\pi_{H2} + (1 - p_2)\pi_{L2} \qquad \text{(A5)}$$

We then substitute in equations (2), (3), (5) and (6) in (A5) to yield (15) in the main text.

Table 1		*Individual 2*	
	Strategy	**H**	**L**
Individual 1	**H**	ω	$\omega + \gamma G$
	L	$\omega + (1 - \gamma)G$	ω

References Cited

Alpesteguia, J., S. Huck, and J. Oeschssler. 2003. "Imitation: Theory and Experimental Evidence," in *CESifo Working Papers*.

Arnold, J. E. 1993. Labor and the Rise of Complex Hunter-Gatherers. *Journal of Anthropological Archaeology* 12:75-119.

Atran, S. M., Douglas, N. Ross, Lynch, Elizabeth, and J. Coley, Ek, Edilberto Ucan and Vapnarsky, Valentina. 1999. Folkecology and commons management in the Maya Lowlands. *Proc. Natl. Acad. Sci.* 96:7598-7603.

Axtell, R. L., J. M. Epstein, and H. P. Young. 2001. "The Emergence of Classes in a Multi Agent Bargaining Model," in *Social Dynamics*. Edited by S. Durlauf and H. P. Young, pp. 191-211. Cambridge: MIT Press.

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.

Boone, J. L. 1992. "Competition, Conflict, and the Development of Hierarchies," in *Evolutionary Ecology and Human Behavior*. Edited by E. A. Smith and B. Winterhalder, pp. 301-337. Hawthorne, N.Y.: Aldine de Gruyter.

Boyd, D. 2001. Life Without Pigs: Recent Subsistence Changes Among the Irakia Awa, Papua New Guinea. *Human Ecology* 29:259-281.

Boyd, R., and P. Richerson. 2002. Group Beneficial Norms Can Spread Rapidly in a Structured Population. *Journal of Theoretical Biology* 215:287–296.

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

—. 1987. The Evolution of Ethnic Markers. *Cultural Anthropology* 2:27-38.

—. 1990. Group Selection Among Alternative Evolutionarily Stable Strategies. *Journal of Theoretical Biology* 145:331-342.

Carneiro, R. L., and S. F. Tobias. 1963. The Application of Scale Analysis to the Study of Cultural Evolution. *Transaction of the New York Academy of Sciences* 26:196-297.

Clark, J. E., and M. Blake. 1994. "The Power of Prestige: Competitive Generosity and the Emergence of Rank Societies in Lowland Mesoamerica," in *Factional Competition and Political Development in the New World*. Edited by E. M. B. a. J. W. Fox, pp. 17-30. Cambridge: Cambridge University Press.

Cronk, L. 1994. Evolutionary theories of morality and the manipulative use of signals. *Zygon* 29:81-101.

DeMarrais, E., L. J. Castillo, and T. Earle. 1996. Ideology, Materialization, and Power Strategies. *Current Anthropology* 37:15-31.

DiMaggio, P. 1994. "Culture and Economy," in *The Handbook of Economic Sociology*. Edited by N. J. Smelser and R. Swedberg. Princeton: Princeton University Press.

Durkheim, E. 1933. *The division of labor in society*. New York: Free Press.

Eerkens, J. W., and C. P. Lipo. 2005. Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology* 24:316-334.

Fiske, A. P. 1998. "Learning A Culture The Way Informants Do: Observing, Imitating, and Participating."

Fried, M. 1967. *The Evolution of Political Society: An Essay in Political Anthropology*. New York: Random House.

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.

Gilman, A. 1981. The Development of Social Stratification in Bronze Age Europe. *Current Anthropology* 22:1-23.

Gintis, H. 2000. *Game Theory Evolving*. Princeton: Princeton University Press.

Henrich, J. 2004a. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization* 53:3-35.

—. 2004b. 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, J., and F. Gil-White. 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.

Henrich, J., and R. McElreath. 2003. The Evolution of Cultural Evolution. *Evolutionary Anthropology* 12:123-135.

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.

Johnson, A., and T. Earle. 2000. *The Evolution of Human Societies*, Second edition. Stanford: Stanford University Press.

Kerbo, H. R. 2006. *Social stratification and inequality : class conflict in historical, comparative, and global perspective*, 6th edition. New York: McGraw-Hill.

Kroll, Y., and H. Levy. 1992. Further Tests of the Separation Theorem and the Capital Asset Pricing Model. *American Economic Review* 82:664-670.

Lancy, D. F. 1996. *Playing on Mother Ground: Cultural Routines for Children's Development*. *Culture and Human Development*. London: The Guilford Press.

Lenski, G. E., J. Lenski, and P. Nolan. 1991. *Human societies : an introduction to macrosociology*, 6th edition. New York: McGraw-Hill.

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.

Naroll, R. 1956. A Preliminary Index of Social Development. *American Anthropologist* 58:687-715.

Naroll, R., and W. T. Divale. 1976. Natural Selection in Cultural Evolution: Warfare. *American Ethnologist* 3:97-129.

Naroll, R., and R. Wirsing. 1976. Borrowing Versus Migration as Selective factors in Cultural evolution. *J. Conflict Resolution* 20:187-212.

Neiman, F. D. 1995. Stylistic Variation in Evolutionary Perspective: Inferences from Decorative Diversity and Interassemblages Distance in Illinois Woodland Ceramic Assemblages. *American Antiquity* 60:7-36.

Netting, R. M. 1990. "Population, permanent agriculture, and polities: unpacking the evolutionary portmanteau," in *The Evolution of Political Systems*. Edited by S. Upham. New York: Cambridge University Press.

Offerman, T., and J. Sonnemans. 1998. Learning by experience and learning by imitating others. *Journal of Economic Behavior and Organization* 34:559-575.

Oster, G., and E. O. Wilson. 1979. *Caste and Ecology in Social Insects* Princeton: Princeton University Press.

Pingle, M. 1995. Imitation vs. rationality: An experimental perspective on decision-making. *Journal of Socio-Economics* 24:281-315.

Pingle, M., and R. H. Day. 1996. Modes of economizing behavior: Experimental evidence. *Journal of Economic Behavior & Organization* 29:191-209.

Renfrew, C., and J. F. Cherry. Editors. 1986. *Peer Polity Interaction and Sociopolitical Change*. Cambridge: Cambridge University Press.

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

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

Richerson, P. J., R. Boyd, and R. L. Bettinger. 2001. Was Agriculture Impossible During the Pleistocene But Mandatory During the Holocene? A Climate Change Hypothesis. *American Antiquity* 66:387-411.

Rousseau, J. J. 1755. "Discourse on the Origin and the Foundations of Inequality Among Men," The Online Library Of Liberty.

Ruyle, E. E. 1973. Slavery, Surplus, and Stratification on the Northwest Coast: the Ethnoenergetics of an Incipient Stratification System. *Current Anthropology* 14:603-631.

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.

Selten, R., and J. Apestegula. 2002. "Experimentally Observed Imitation and Cooperation in Price

Competition on the Circle," in *Bonn Econ Discussion Papers*.

Service, E. R. 1975. *Origins of the State and Civilization: The Process of Cultural Evolution*. New York: W.W. Norton & Company.

Shennan, S. 2001. Demography and Cultural Innovation: A Model and Its Implications for the Emergence of Modern Human Culture. *Cambridge Archaeology Journal* 11:5-16.

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.

Smith, E., and J.-K. Choi. 2007. "The Emergence of Inequality in Small-Scale Societies: Simple Scenarios and Agent-Based Simulations," in *The Model-based Archaology of Socionatural Systems*. Edited by T. Kohler and S. van der Leeuw. Santa Fe: SAR Press.

Ugan, A., J. Bright, and A. Rogers. 2003. When is technology worth the trouble? *Journal of Archaeological Science* 30:1315-1329.

Weibull, J. W. 1995. *Evolutionary Game Theory*. Cambridge, Mass: MIT Press.

Wiessner, P. 2002. The Vines of Complexity. *Current Anthropology* 43:233-269.

Young, H. P. 1998. *Individual strategy and social structure : an evolutionary theory of institutions*. Princeton, N.J: Princeton University Press.

NOTES

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

[2] There are actually five solutions. 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), so we've focus only on one of them in the text.

[3] A Visual Basic program that visually simulates this evolutionary process our system is available at http://www.sscnet.ucla.edu/anthro/faculty/boyd/MESB/ClassesSimulationFiles.zip.

[4] β need not be a purely theoretical parameter. It can be estimated in laboratory experiments or by fitting real world data (such as the diffusion of innovations literature) to formal learning models.

[5] 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.

[6] Selection among stable equilibrium does not involve the evolutionary challenges often associated with altruism and genetic group selection. Unlike cooperative situations with a free-rider problem, a stable equilibrium means that the within-group selection forces have nearly been exhausted (zeroed out, if you will).

[7] 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.

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

[9] 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.

[10] 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.

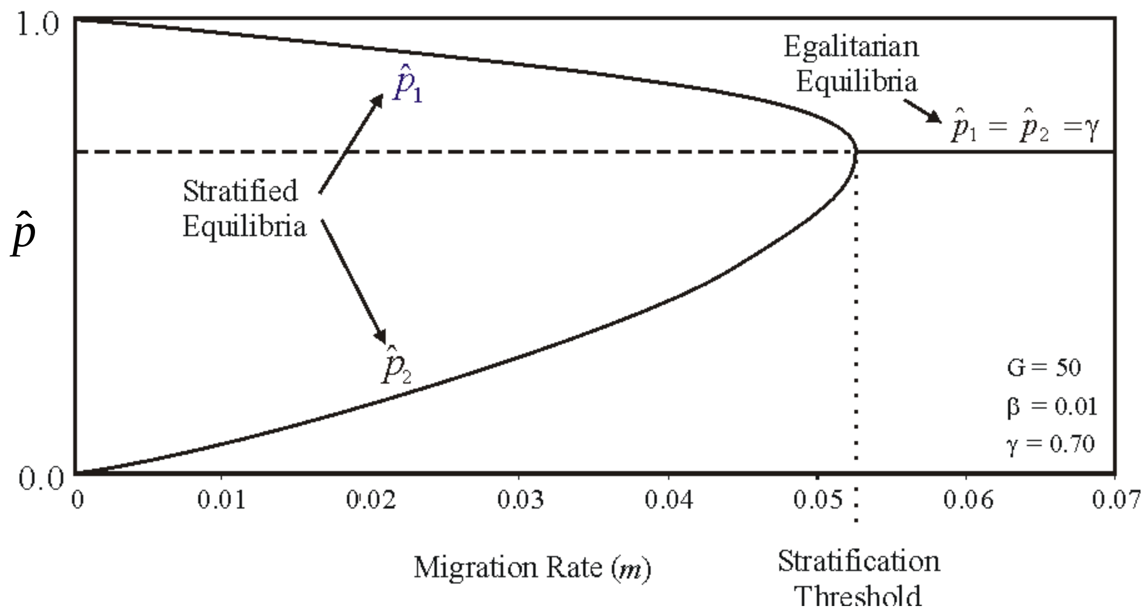

Figure 1. This plots the location of our two equilibrium solutions. For low migration rates, both equilibria exist, but only the stratified equilibrium is stable. The dashed line indicates the location of the unstable egalitarian equilibrium. Under such condition the system evolves to the stratified equilibrium. When dashed line changes to solid, at m = 0.0525 (in this case), the egalitarian equilibrium becomes stable, and stratified equilibria goes out of existence. This is the *stratification threshold*.

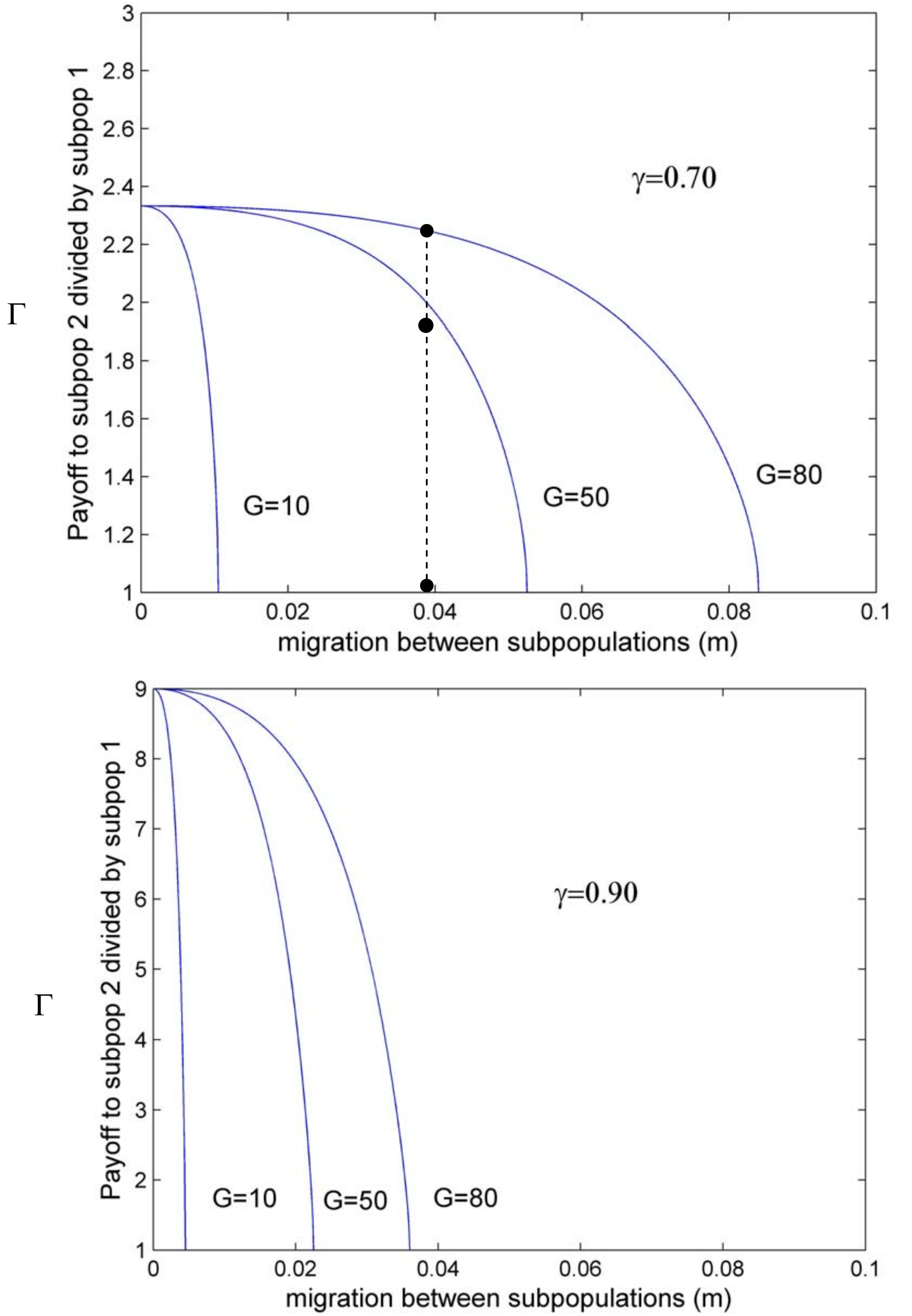

Figure 2. These figure plots the ratio of the payoff to the 'high' *subpop* to the 'low' *subpop*, Γ, against *m*. The plots represent the same parameters, except that γ = 0.70 in the upper plot (Figure 2a) and γ = 0.90 in the lower plot (Figure 2b).

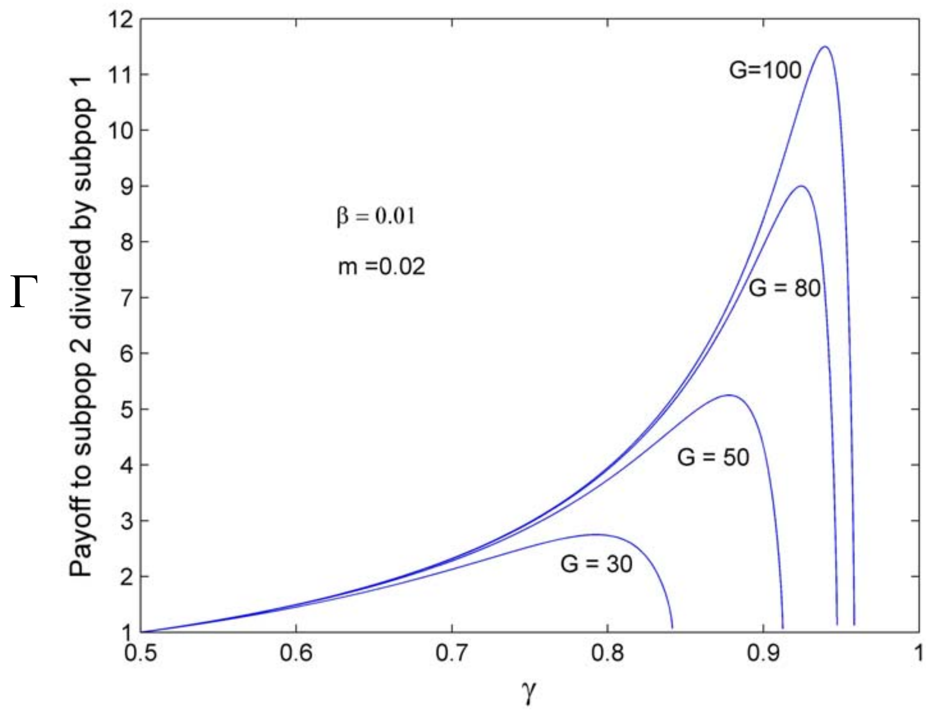

Figure 3. This plot γ, the individual-level division of surplus, against Γ for four values of G. Combinations of larger G values and larger γ values dramatically increases the inequality between subpopulations.

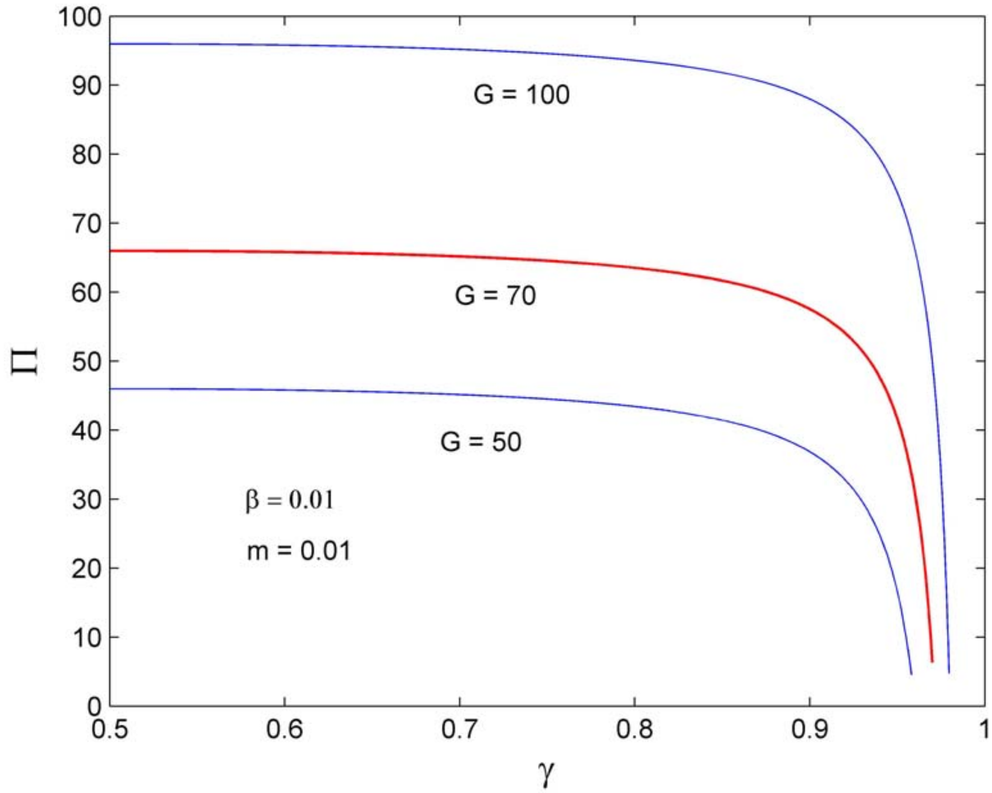

Figure 4. This plots the overall population payoff against γ for three values of G (m = 0.01, δ = 1, β = 0.01). The curve illustrate the non-linear influence on population payoff of decreasing γ. Cultural group selection favors more equitable stratification.

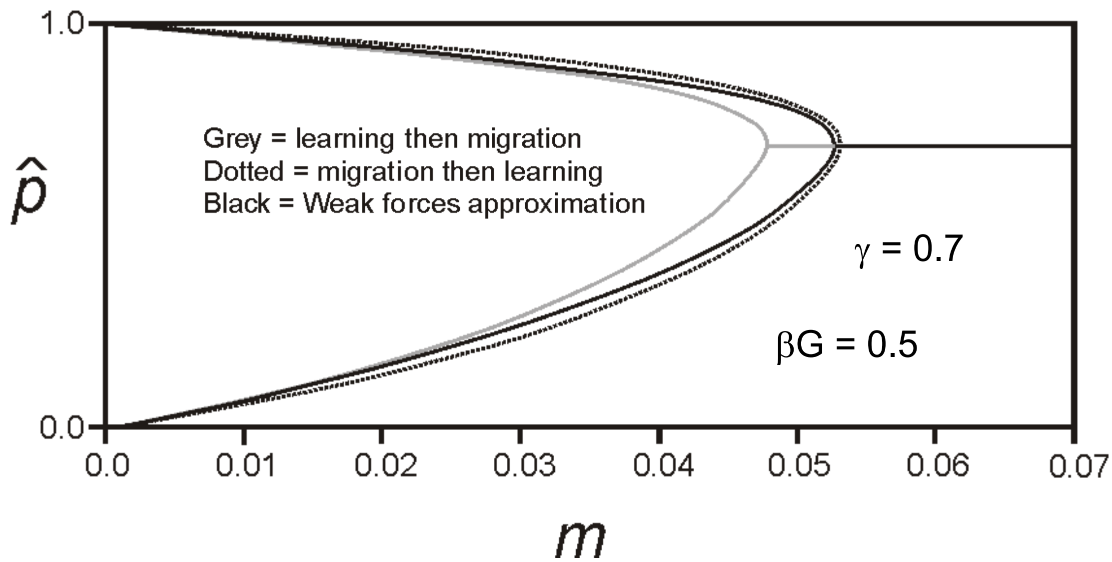

Figure A1. Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, *m*, for three different models. The black line is the analytical solution assuming weak forces. The grey line assumes learning occurs first followed by migration, and the dashed line assumes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

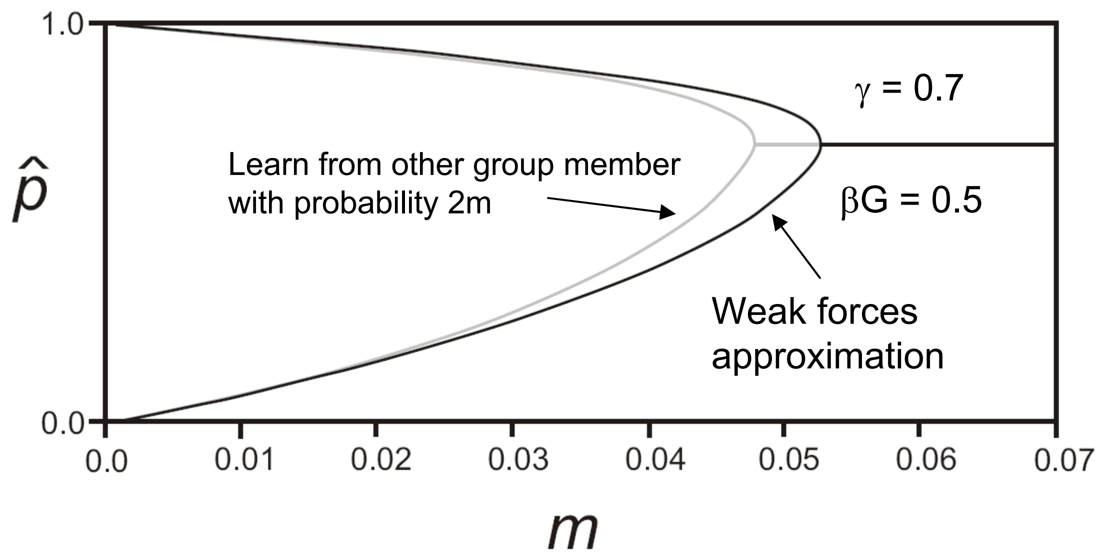

Figure A2. Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, *m,* for two models. The black line is the analytical solution assuming weak forces. The grey line assumes that individual imitate an individual from their own subpopulation with probability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

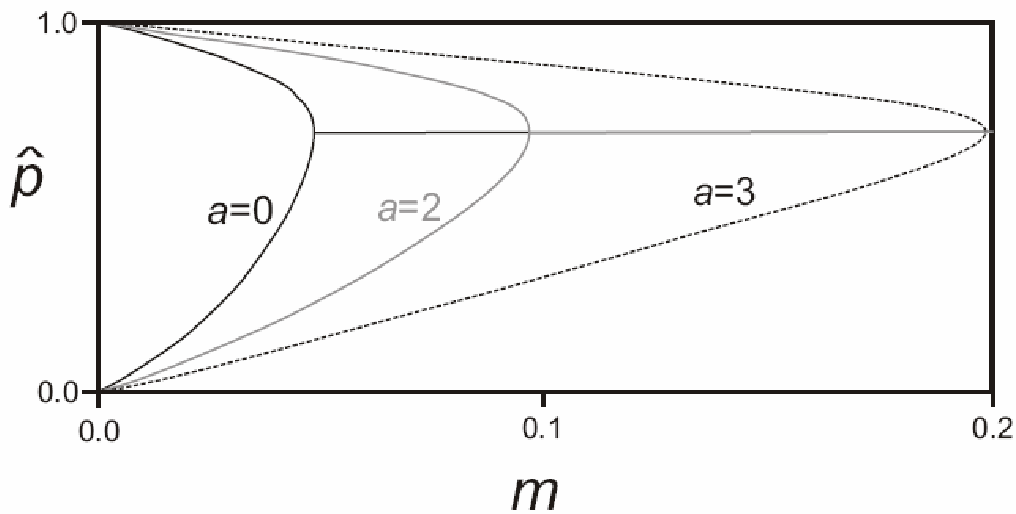

Figure A3: Shows the equilibrium frequencies of H in the two subpopulations as a function of the mixing rate, *m*, for three amounts of success biased migration. In each case imitation occurs first, then migration. The black line is assumes no payoff bias (*a* = 0). The black and dashed lines show greater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