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

Joseph Henrich and **Robert Boyd**

Department of Psychology and Department of Economics, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada (joseph.henrich@gmail.com)/Department of Anthropology, University of California, Los Angeles, CA 90254, U.S.A. (robert.t.boyd@gmail.com). 21 II 08

CA+ Online-Only Material: Supplement A

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 process driving cultural evolution increases individuals' economic gains. The key assumptions are that human populations are structured into groups and that cultural learning is more likely to occur within 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.

Explaining social stratification has been an important focus of social thought at least since the Enlightenment. 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 nineteenth century. 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, and 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 (Fried 1967). Our puzzle, however, is not this ubiquitous inequality among individuals but "social stratification," persistent inequality among social groups such as 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 ability among individuals (Richerson and Boyd 2005). Instead, these differences must result from something that individuals acquire as a consequence of group membership. This leads to the obvious question why people on the wrong side of such inequalities do not adopt the skills, practices, behaviors, or strategies of the people 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. Many deny the paradox, arguing either that people are systematically deceived about their interests (e.g., as a result of elite propaganda) or that they are coerced into submission (Cronk 1994; DeMarrais, Castillo, and Earle 1996; Kerbo 2006). Others 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 among groups. For example, either correlated asymmetries such as access to high-quality resources (such as land) or uncorrelated ones such as 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 in investment and income between groups (Lundberg and Startz 1998). 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 this 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 is 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 more 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 population is subdivided into social groups and that people tend to learn more often from members of their own group than from members of 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 subgroups and despite the fact that the only force shaping cultural variation is an adaptive learning mechanism that myopically maximizes payoffs. We also show that this can give

 0011-3204/2008/4904-0008$10.00. DOI: 10.1086/587889

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 (Richerson and Boyd 2005). For example, recent laboratory studies in economics using performance-based monetary incentives indicate that people rely on imitating beliefs, economic strategies, and behaviors of particularly *successful individuals* in social interactions (see Henrich and Henrich 2007, chap. 2). This work is consistent with experimental findings 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 (Henrich and Gil-White 2001).

Cultural transmission is complex, and a number of processes not included in this model are undoubtedly important (Henrich and McElreath 2003; Richerson and Boyd 2005). 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 leads to the spread of individually more successful beliefs and practices? This sets the bar higher than it might otherwise be.

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, people tend to live in and preferentially associate with local aggregations, whether they be villages, neighborhoods, ethnic enclaves, bands, or clans. 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 present an evolutionary game-theoretical model that formalizes these assumptions. 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 (table 1). If both interacting individuals use the same strategy, each receives the baseline payoff, ω. However, if they use different strategies, the individual using H receives a payoff of $\omega + \gamma G$ while the individual using 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 complementary element in the interaction. The parameter γ gives the proportion of the surplus that goes to the individual playing H, and $1 - \gamma$ gives the proportion of the surplus that goes to the individual playing L. We assume that γ ranges from 0.5 to 1.0 without loss of generality. As a real-world referent, H could represent skills and know-how about management, politics, commerce, defense, construction, or metallurgy, while L could represent agricultural production, herding, foraging, or physical labor.

Table 1. Payoff Matrix for Social Interaction

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

The population is divided into two subpopulations, 1 and 2. This structure could result from anything that patterns social interactions, including distance, geographical barriers, or social institutions such as villages, clans, or ethnic groups. The frequency of individuals in subpopulation 1 using 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 with a randomly selected individual from the other population (individuals from subpopulation 1 meet those from subpopulation 2 and vice versa) and that with probability $1 - \delta$ the individual meets someone randomly selected from her 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 mixing between subpopulations occurs. Cultural mixing can happen in two 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 (for details, see CA+ online supplement A) and then mix by 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, it could be 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. In this case, mixing occurs as ideas flow between subpopulations. In supplement A, we show that this is equivalent to assuming that learners observe and learn from a model in the other subpopulation with probability $2m$ and 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 evolutionary game theory (McElreath and Boyd 2007), we can express the change in the frequency of individuals playing H in subpopulation 1 in one time step, Δp_1, as

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

Recursion (1) contains two parts: (*a*) the effects of success-biased transmission and (*b*) the movement of cultural variants between the subpopulations. The symbol π_{H1} gives the expected payoff received by individuals playing H from subpopulation 1, while π_{L1} gives the expected payoff to individuals from subpopulation 1 playing L. The parameter β is a constant that scales differences in payoffs into changes in the frequency of cultural variants (more on this below). The derivation of (1) in supplement A 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 qualitatively change our findings.

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 the probability of interacting with an individual from the other subpopulation (δ):

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

$$\pi_{L1} = \delta\{p_2[\omega+G(1-\gamma)]+(1-p_2)\omega\} + (1-\delta)\{p_1[\omega+G(1-\gamma)]+(1-p_1)\omega\}. \tag{3}$$

The first terms on the right-hand side of equations (2) and (3) give the expected payoff to H and L individuals as a result of interacting with an individual from the other subpopulation, and the second terms give the expected payoffs when interacting with a member of the player's own subpopulation given the chance of meeting either an H or an L (p_1 and $1-p_1$, respectively) individual.

Similarly, the change in the frequency of H strategies in subpopulation 2, Δp_2, can be expressed as

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

where, as above,

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

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

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 longer run. Of particular interest are the stable equilibria. The equilibria are combinations of p_1 and p_2 that, according to equations (1) and (4), lead to no further change in behavior. An equilibrium is locally stable when the population will return to that equilibrium if perturbed. It is unstable if small shocks cause the population to evolve 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 H and L. This situation could be interpreted as each individual using a mixed strategy of H and L (individuals lack task-specialized skills). At such egalitarian equilibria, the *average* payoff of all individuals is the same no matter which subpopulation they are from. There are also *stratified equilibria*, in which most individuals in one subpopulation play H while most individuals in the other subpopulation use L. In this case, the average payoff to the subpopulation that consists mostly of H individuals is higher than the average payoff of individuals in the subpopulation consisting of mostly L individuals. The analysis presented below indicates that, depending on the parameters, either one or the other type of stable equilibrium exists 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 simulations and analytical methods 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 of δ by the majority of the subpopulation any time 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 $\delta < 1$ because real-world obstacles may prevent δ from reaching 1.

To find the equilibrium values we set Δp_1 and Δp_2 equal to 0 and solved for the equilibrium frequencies $\hat{p}_1$ and $\hat{p}_2$. This yields two interesting solutions (see supplement A), an egalitarian and a stratified equilibrium. At the egalitarian equilibrium

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

This tells us that the frequency of H in each subpopulation at equilibrium will be equal to the fraction of the surplus received by H during an interaction (this also holds for

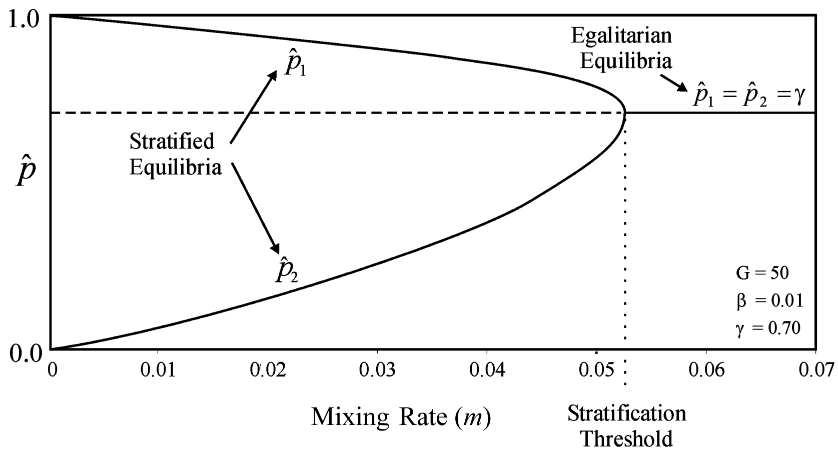

Figure 1. Plot of the location of the two equilibrium solutions. The dashed line indicates the location of the unstable egalitarian equilibrium. When the dashed line changes to solid, at $m = 0.0525$ (in this case), the egalitarian equilibrium becomes stable, and the stratified equilibrium disappears. This is the *stratification threshold.*

$0.5 \leq \delta \leq 1$). The stratified equilibrium is locally stable and the egalitarian equilibrium is unstable if

$$G\beta\gamma(1-\gamma) > 2m. \tag{8}$$

If (8) is not satisfied, *only* the egalitarian equilibrium exists, and it is stable.

Thus, the system has two qualitatively different and mutually exclusive 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.* Figure 1 shows the frequencies of H in the two subpopulations at the two equilibria 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. The curves labeled $\hat{p}_1$ and $\hat{p}_2$ give the frequencies of H in subpopulations 1 and 2 at the stratified equilibrium. When m is low, the stratified equilibria exist and are stable and the egalitarian equilibrium exists but is not stable. As the flow of strategies or people between subpopulations (m) increases, the frequencies of H in the two subpopulations converge toward each other. Stratification disappears at exactly the point at which the frequencies of individuals adopting H become equal in the two populations. At higher values of m only the egalitarian equilibrium exists and is stable.

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 individuals 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 (see supplement A).

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

$$\hat{\bar{\pi}}_1 = \hat{p}_1\hat{\pi}_{H1} + (1-\hat{p}_1)\hat{\pi}_{L1}, \tag{9}$$

$$\hat{\bar{\pi}}_2 = \hat{p}_2\hat{\pi}_{H2} + (1-\hat{p}_2)\hat{\pi}_{L2.} \tag{10}$$

Substituting expressions (2), (3), (5), and (6), evaluated at location $\hat{p}_1$ and $\hat{p}_2$ as expressed by equations (A5) and (A6) (see supplement A), into (9) and (10) gives the average payoff in each subpopulation. We will use a ratio of the average payoffs to summarize inequality:

$$\Gamma = \frac{\hat{\bar{\pi}}_2}{\hat{\bar{\pi}}_1}. \tag{11}$$

How the Parameters Affect Stratification

Next, we examine how varying each of the parameters—G, m, β, and γ—influences (1) the emergence of a stable stratified equilibrium versus 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.

The mixing rate, m, measures the flow of ideas or people between the two subpopulations. Several factors might influ-

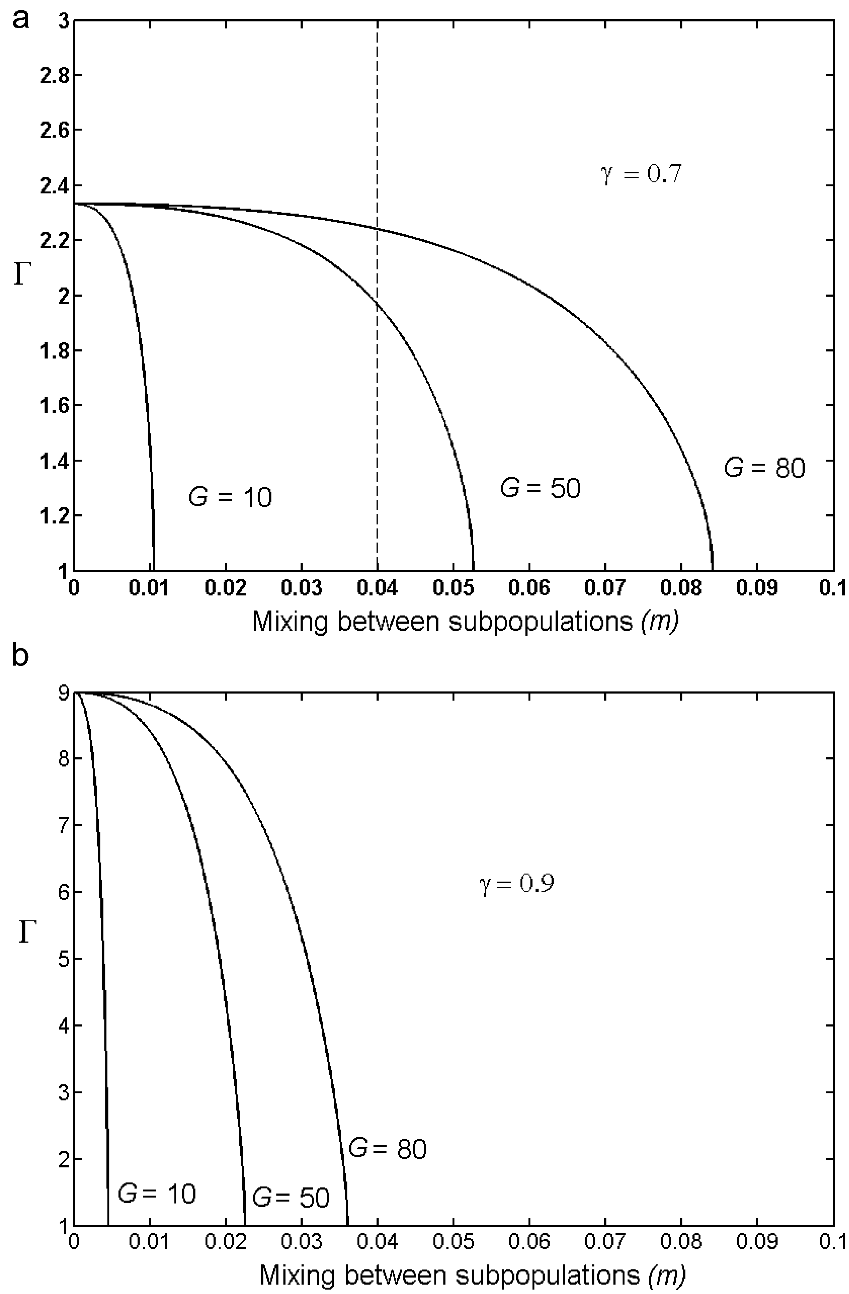

Figure 2. Plots of the ratio of the payoff to the H subpopulation to the L subpopulation, Γ, against m. The plots represent the same parameters, except that $\gamma = 0.7$ in *a* and $\gamma = 0.9$ in *b*. $\beta = 0.01$.

ence 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, because both theoretical and empirical considerations suggest that social learning is likely influenced by symbolic (accent, language, dress, etc.) and other kinds of phenotypic markings, often related to such phenomena as ethnicity, race, caste, and class, m can be small even in an interspersed population.

Decreasing m makes it easier to produce stable social stratification. When m is greater than the stratification threshold, there is no inequality; all individuals have the same expected payoff. When m is below the threshold, only the stratified equilibrium is stable, and decreasing m increases the degree of inequality (fig. 2).

The results presented so far are based on the assumption that m is low enough that the order of transmission and mixing do not matter, and as a result, the same model can be used to represent movement of individuals or the flow of

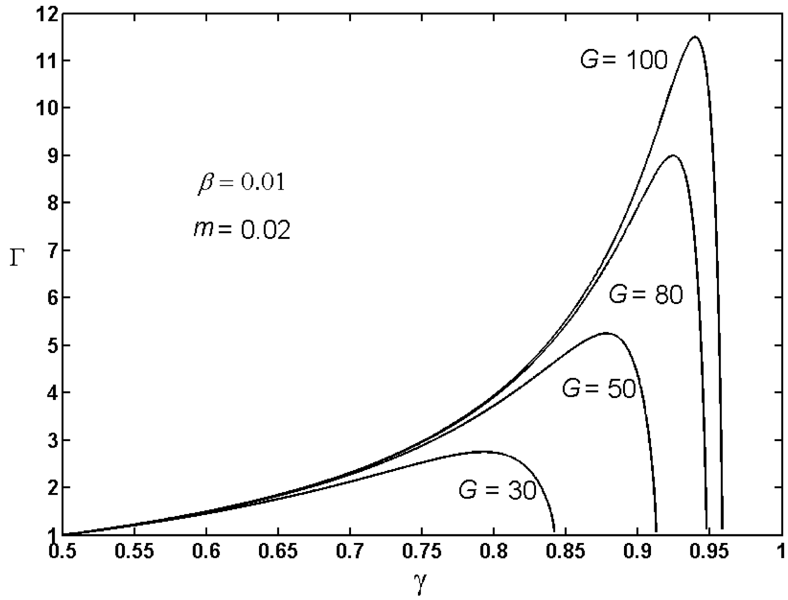

Figure 3. Plots of γ, the individual-level division of surplus, against Γ for four values of G.

ideas. When rates of change are higher, the two models yield different sets of recursions. Simulations, however, discussed in supplement A, indicate that these different models have very similar qualitative properties.

Another concern is the assumption that the amount of mixing is unaffected by the difference in payoffs between subpopulations. It seems plausible that such payoff differences might increase the flow of people from the lower- to the higher-payoff subpopulation and reduce flow in the opposite direction and that this would undermine the results presented above. This is not the case, however. In supplement A we show that adding success-biased physical migration to the model actually *increases* the range of conditions conducive to social stratification.

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

Greater values of G expand the conditions favoring stable stratification and increase the degree of inequality if stratification is already stable. This means that, ceteris paribus, technologies, practices, or forms of organization that favor greater production (through exchange and specialization) favor increasing the degree of stratified inequality. Both of these effects can be seen in figure 2. The dashed vertical line in figure 2, *a*, shows that for the same value of m (and γ and 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 individuals playing H. The parameter γ might be influenced by resource availability, skill investment (high skill vs. low skill), 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 (8) is proportional to $\gamma(1-\gamma)$. The greater this product is, 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 nonintuitively, 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 by comparing *a* and *b* in figure 2. Looking at the stratification threshold for each value of G (where the curves intersect the X-axis), we see that when $\gamma = 0.7$, 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 the societal level. (This is not the same as saying that the *degree* of inequality is higher.)

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

The social learning scale parameter, β, converts differences in payoffs between the strategies observed by social learners into probabilistic changes in behavior, and therefore β has units of $1/G$. For this reason, β must be $< 1/(G\gamma)$ and > 0. 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 effect of one incident of social learning is small and cultural evolution will proceed slowly. Below we have more to say about the kinds of real-world, cross-society differences that might influence β. From the perspective of what we have derived so far, β occupies a position alongside G, and therefore changes to β have the same effects as changes to G.

Mixed Interaction and Stratification

When the restriction that individuals interact only with members of the other subpopulation ($\delta = 1$) is relaxed, the location of the egalitarian equilibrium is $\hat{p}_1 = \hat{p}_2 = \gamma$, the same as above when $\delta = 1$. When

$$\beta G\gamma(1-\gamma)(2\delta - 1) < 2m, \tag{12}$$

only the egalitarian equilibrium is stable. The only difference between (12) and (8) is the term $2\delta - 1$. This tells us that δ values < 1 increase the range of conditions under which the egalitarian equilibrium is stable. While we have not been able to analytically solve the 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 condition (12) likely provides the conditions for the existence of a stable stratified state and all of our above analysis regarding G, γ, m, and β applies to this case as well.

Stratification and Total Group Payoffs

Our results so far indicate that restricted mixing 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 $\delta = 1$)

$$\hat{\pi} = 2\omega + G[\hat{p}_1(1-\hat{p}_2) + \hat{p}_2(1-\hat{p}_1)]. \tag{13}$$

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 miscoordination, and the population gets all of the surplus. More miscoordination causes some of the surplus to be lost and the average payoffs to decrease.

Substituting the values of p_1 and p_2 at the stratified equilibrium, equations (A5) and (A6) and setting $\omega = 0$ yields

$$\hat{\pi} = \frac{[2m - G\beta(1-\gamma)](G\beta\gamma - 2m)}{\beta[m - G\beta\gamma(1-\gamma)]}. \tag{14}$$

Cultural Group Selection and Social Stratification

Equation (14) suggests that cultural group selection might affect social stratification. It is plausible that the amount of mixing, the magnitude of the surplus, and the equality of the division are influenced by culturally transmitted beliefs and values. 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 (Henrich 2004*a*). For cultural group selection to operate, group payoffs ($\hat{\pi}$) must influence the outcomes of competition among societies. Warfare, for example, might cause societies with higher $\hat{\pi}$ to proliferate because economic production can generate more weapons, supplies, allies, feet on the ground, skilled warriors, etc. Alternatively, extra production could lead to faster relative population growth. Or, perhaps more important, because individuals tend to imitate those with higher payoffs, when people from a poorer society meet people from a richer society, there will be a tendency for cultural traits and institutions to flow from richer to poorer societies (Boyd and Richerson 2002).

Assuming that 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 specialization and stratified equilibria over egalitarian equilibria because stratification yields the surplus benefits of specialization. The highest total payoff a nonstratified 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 0. 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 production and increase between-group competitiveness. Modeling work on the evolution of technological complexity (Henrich 2004*b*; Shennan 2001) indicates that larger, denser social groups should 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 $\hat{\pi}$.

4. Cultural group selection favors bigger β values because any institutions, beliefs, or values that lead individuals to weigh economically successful members of their subpopulation more heavily in their cultural learning will create more competitive groups.

5. Figure 4 plots the total payoff for the stratified equilibrium against γ for three values of G ($\beta = 0.01$, $m = 0.01$, and $\delta = 1$). The strength of cultural group selection is equal to the slope of the line ($\partial\hat{\pi}/\partial\gamma$). This means that for the large 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, it only weakly favors greater equality. In the longer run, cultural group selection favors $\gamma = 0.5$ (equality) *with* specialization and stratification. These predictions assume that only one parameter varies at a time. It is probable that some of these parameters are causally interconnected, probably in different ways in different circumstances, and therefore a more complete analysis of particular historical situations will require modeling of these interconnections.

Connections to Existing Approaches

This model does not capture the only route to stratification, and it contributes to an understanding of only one aspect of the more general problem of the evolution of societal complexity. Nevertheless it does show how economic specialization and cultural differentiation can sometimes produce stratified inequality. We believe that this model informs existing theoretical approaches, especially when it 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. Here, we highlight some of the relationships between our model and existing 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 partially culturally isolated. It suggests that increasing economic surplus does not necessarily lead to stratified inequality. A large surplus does make it more likely that the conditions for stratification will be satisfied, but stratification also depends on the amount of mixing and the degree of inequality. Because we would like to emphasize that this model applies to emergence of occupational specialization in general, we deploy our model in supplement A to interpret an ethnographic case involving the stable coexistence of highly interdependent specialized occupational castes in the Swat Valley, Pakistan (Barth 1965). This example illustrates how self-interested dyadic exchanges among partially culturally isolated (endogamously marrying) but geographically and economically intermixing groups can yield a stable, stratified situation with different mean payoffs among occupations.

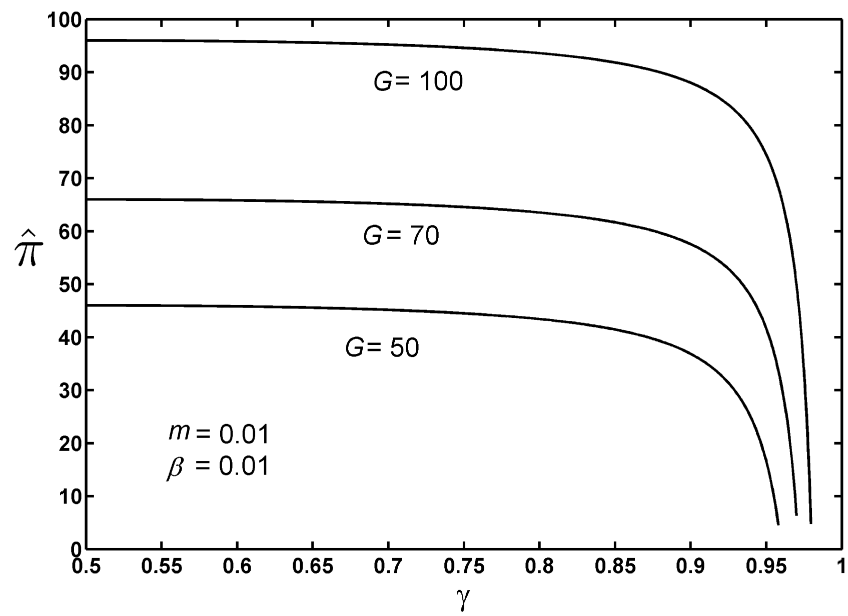

Figure 4. Plots of the overall population payoff ($\hat{\pi}$) against γ for three values of G ($m = 0.01$, $\delta = 1$, $\beta = 0.01$).

Population Pressure, Intensification, and Social Stratification

Theories of societal evolution often emphasize population pressure as the prime mover (Johnson and Earle 2000; Netting 1990). Declining yields per capita create a need for intensification and this, in turn, gives rise to stratification. While there are good reasons to doubt that population pressure causes stratification (Richerson, Boyd, and Bettinger 2001), denser populations do 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 2004*b*; Shennan 2001). This, in turn, generates more surplus (i.e., higher values of G) (Carneiro and Tobias 1963) and favors stratification, permitting greater degrees of societal inequality to emerge. Higher levels of productivity often support even larger populations, which will support a higher equilibrium level of technological/skill sophistication. Such a feedback loop could create the observed relationship between stratification and population variables.

Conflict and Circumscription

Warfare between social groups does not cause stratification in our model. However, cultural group selection will spread *certain kinds* of stratification through intergroup competition. In this vein, stratified specializations allow for both warrior and weapon-maker castes, which have benefits in violent, competitive socioecologies. We would 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 timescales, greater equality between subpopulations ($\gamma \rightarrow 0.5$).

Division of Labor in Other Species

While the profitable division of labor is relatively common *between* different species (interspecies mutualisms), the heritable division of labor *within* species, involving different types or morphs, is relatively rare except among eusocial insects. This rarity within species is explained by the Bishop-Cannings theorem, which shows that any division of labor that leads to mean fitness differences among occupational types will be unstable because the type with higher fitness should outcompete the type with lower fitness (Bishop and Cannings 1978). Different types may persist at equilibrium, but they must have the same mean fitness (payoff). Between species, however, mutually beneficial divisions of labor can arise and remain stable because different species do not compete directly in the same gene pool. The model presented here shows that cultural evolution may provide an intermediate case between the genetic evolutionary circumstances within species and the ecological mutualism found in interspecies interaction. This illustrates the potential pitfalls of any direct mapping of theoretical findings from genetic to cultural evolution. We discuss this and an ethnographic example of culturally evolved niche partitioning in supplement A.

Conclusion

The model described above leads to a number of qualitative conclusions about the evolution of social stratification. Some of these conclusions are not surprising (e.g., that increased mixing between subpopulations decreases stratification), but others are less obvious (e.g., that increasing the surplus available tends to increase the degree of stratification). The payoff from such simplified models is a clearer qualitative understanding of a set of generic processes that, along with our understanding of other unmodeled processes or processes modeled elsewhere (such as intergroup competition), can be applied to a wide range of specific cases, including such phenomena as social classes, ethnic occupations, castes, guilds, and occupation-based clan divisions, by asking how specific historical developments such as agriculture, irrigation, steel plows, and craft specialization influenced the parameters of the model and thereby led to stratification or greater inequality in an already stratified society.

References Cited

Axtell, R. L., J. M. Epstein, and H. P. Young. 2001. The emergence of classes in a multi-agent bargaining model. In *Social dynamics*, ed. S. Durlauf and H. P. Young, 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* 70:85–124.

Boyd, R., and P. Richerson. 2002. Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology* 215:287–96.

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.

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.

Fiske, A. P. 1998. Learning a culture the way informants do: Observing, imitating, and participating. http://www.sscnet.ucla.edu/anthro/bec/papers/learning_culture.htm.

Fried, M. 1967. *The evolution of political society: An essay in political anthropology.* New York: Random House.

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

———. 2004*b*. Demography and cultural evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity* 69:197–214.

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* 22:165–96.

Henrich, J., and R. McElreath. 2003. The evolution of cultural evolution. *Evolutionary Anthropology* 12:123–35.

Henrich, N. S., and J. Henrich. 2007. *Why humans cooperate: A cultural and evolutionary explanation.* Oxford: Oxford University Press.

Johnson, A., and T. Earle. 2000. *The evolution of human societies.* 2d 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.

Lancy, D. F. 1996. *Playing on mother ground: Cultural routines for children's development.* London: Guilford Press.

Lundberg, Shelly, and R. Startz. 1998. On the persistence of racial inequality. *Journal of Labor Economics* 16:292–323.

McElreath, R., and R. Boyd. 2007. *Modeling the evolution of social behavior.* Princeton: Princeton University Press.

Naroll, R. 1956. A preliminary index of social development. *American Anthropologist* 58:687–715.

Netting, R. M. 1990. Population, permanent agriculture, and polities: Unpacking the evolutionary portmanteau. In *The evolution of political systems*, ed. S. Upham, 21–61. New York: Cambridge University Press.

Oster, G., and E. O. Wilson. 1979. *Caste and ecology in social insects.* Princeton: Princeton University Press.

Richerson, P., and R. Boyd. 2005. *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.

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.

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 archaeology of socionatural systems*, ed. T. Kohler and S. van der Leeuw, 105–20. Santa Fe: SAR Press.