# The evolution of prestige-biased transmission

Karthik Panchanathan
Center for Behavior, Evolution & Culture
Department of Anthropology, UCLA

May 27, 2010

## Introduction

Humans are unique in their capacity to acquire and transmit adaptive information through both genetic and cultural channels (Boyd and Richerson, 2005). To understand the dynamics of cultural evolution, we need to understand the nature of our social psychology, cataloguing the various learning biases that underlie cultural transmission. Henrich and Gil-White (2001) argue that a prestige-bias (i.e., a preference for imitating prestigious individuals) would be a particularly effective learning bias, as it increases the likelihood that naive individuals acquire adaptive information. There are two parts to their argument. First, selection would have co-opted pre-existing psychological adaptations designed for ranking conspecifics, enabling imitators to rank models in terms of skill using some measure of success as a proxy. Second, selection would have favored imitators to confer deference on successful individuals in order to gain proximity, enabling imitators to successfully imitate experts.

While there has been theoretical research on the effects of prestige on cultural evolution (e.g., Boyd and Richerson, 1985) and even on the evolution of prestige-seeking (Ihara, 2008), there have been no attempts to formalize and thereby verify Henrich and Gil-White's (2001) conjecture regarding the evolution of a prestige-bias. In this paper, I present a formal model of the evolution of prestige-biased transmission.

After presenting the modeling framework, I demonstrate the effectiveness of a prestige-bias. When individuals imitate randomly-selected models, adaptation is constrained by migration, the rate at which individuals can learn the adaptive behavior on their own, and the difficulty in social learning the adaptive behavior from others. When individuals imitate the prestigious, migration is the only force opposing adaptation, which can result in a much higher frequency of the adaptive behavior.

Next, following Henrich and Gil-White (2001), I study the evolution of the prestige-biased learning bias in two steps. First, assuming that there is no deference paid to the prestigious, when does natural selection favor selective imitation (i.e., the degree to which individuals imitate experts as opposed randomly-selected models)? Second, assuming individuals imitate experts, when does natural selection favor deferential imitation (i.e., the degree to which imitators pay a deference cost to experts in order to gain proximity and thereby reduce social learning errors)?

## Life History

When environments are stationary, adaptation can result from selection on genes. When environments vary temporally or spatially, natural selection can favor social learning, a form of phenotypic plasticity. Here, I use Wright's 'island model' to incorporate spatial variation, assuming that individuals live in an infinite population, structured into infinitely-sized sub-populations. There is a unique adaptive behavior for each sub-population. If an adult possesses the adaptive behavior, he receives a marginal fitness benefit $b$. Adolescents first imitate adults in their sub-population. If they don't acquire the adaptive behavior—and I assume that individuals know when they possess the adaptive behavior—they attempt to learn the adaptive behavior on their own, succeeding with probability $I$, representing the rate of innovation.

Two parameters determine the nature of social learning: the degree to which individuals selectively imitate experts ($s$) and the degree to which they defer to models in order to gain proximity ($d$). The selectivity parameter, $s$, affects the number of models whose behavior is observed. When $s = 0$, imitators pick a model at random from their natal group and attempt to imitate his behavior. If $q$ denotes the fraction of the population in possession of the locally-adaptive behavior, then these random imitators will imitate an expert with probability $q$. When $s = 1$, imitators evaluate a set of $n$ models. This set will contain at least one expert with probability $1 - (1 - q)^n$. With probability $(1 - q)^n$, the set will not include an expert, and so the selective imitator will attempt to innovate the adaptive behavior. Selective imitation comes with a cost, denoted by $c_s$, representing the time and energy necessary to evaluate a set of models. The total cost paid is the product of the cost, $c_s$, and the reliance on selective imitation, $s$. Following Henrich and Gil-White (2001), I assume that the psychological machinery necessary to rank individuals by some measure of prestige already exists, so $c_s$ should only be interpreted as search and evaluation costs. As such, $c_s$ will likely be small relative to $b$.

The deference parameter, $d$, affects the accuracy of social learning. When $d = 0$, imitators learn from a distance, successfully imitating the observed behavior with probability $1 - e$. The parameter $e$ represents the difficulty of imitation from a distance. When $d = 1$, imitators pay a sufficient deference cost in order to gain privileged access to models, thereby guaranteeing successful imitation.

Henrich and Gil-White (2001) argued that the dynamics of prestige and deference would set up a market-like situation. Prestigious individuals will attempt to extract as much deference as they can, while imitators will seek out the lowest deference cost. To capture this market-like mechanism, I assume that the cost of deference is given by $D(1 - q)$. $D$, which will be formally defined in a subsequent section, represents the marginal benefit an imitator gains by switching from learning at a distance to deferential imitation. This is the surplus to be divvied up between the expert and the imitator. When the adaptive behavior is rare ($q \approx 0$), experts can extract nearly all of the surplus. When the adaptive behavior is common ($q \approx 1$), imitators won't pay much to defer to experts. The total deference cost paid is the product of the deference cost, $D(1 - q)$, and the reliance on deferential imitation, $d$.

After social and individual learning, a fraction $m$ of the individuals migrate to a new group. Migration is assumed to be unrelated to the behavior being modeled, driven by processes like marriage. Following Wright's island model, the distribution of immigrants are assumed to reflect the population-wide distribution. Individuals who emigrate will not possess the adaptive behavior as adults, as each group has a uniquely adaptive behavior. Individuals then reproduce proportionate to fitness.

## Cultural Dynamics

When individual learning is coupled to a cultural inheritance system, a population can adapt to its environment in a process analogous to natural selection, a process Boyd and Richerson (1985) call 'guided variation'. Even if the rate of innovation is low, a population can be well-adapted to the environment because adaptations are transmitted culturally. Biased learning rules can amplify the effects of guided variation, increasing the rate of adaptation. If individuals are biased towards imitating adaptive variants, then there are two forces which increase adaptation, innovation and social learning.

To see how this works, let's look at the cultural dynamics of different social learning rules. Let $q$ denote the frequency of the adaptive behavior among adults and $q'$ the frequency among juveniles after social and individual learning in a focal island population. Then, we have the following recursion:

$$q' = (1-m)\big[Q + I(1-Q)\big] \tag{1}$$

where $Q$ is the probability of successfully acquiring the adaptive behavior through social learning based on the distributions of social learning biases in the population, and is a function of $s$ and $d$. Following the assumed life history, juveniles first attempt to socially learn the adaptive behavior. If unsuccessful, they attempt to innovate. Only those individuals who remain in their natal group (a fraction $1-m$) can possess the adaptive behavior as adults. $Q$ is given by:

$$Q(s,d) = \begin{array}{l} (1-s)(1-d)q(1-e) + (1-s)dq + \\ s(1-d)(1-(1-q)^n)(1-e) + sd(1-(1-q)^n) \end{array} \tag{2}$$

In this model, migration opposes adaptation. When the migration rate is very high, then most adults will be maladapted to their environments. However, when migration rates are high, it's unlikely that selection will favor cultural transmission. Acquiring information through a cultural channel is only useful when there is some correlation in the environments experienced across generations. When individuals are unlikely to face the same selective pressures as their parents, there is little reason to imitate parents. In this model, I assume that migration rates will be low ($m \ll 1$). With this assumption, some meaningful fraction of the population will be in possession of the adaptive behavior ($q \gg 0$).

Equation (2) can now be usefully simplified with one further assumption. If selective imitators cast a wide net in selecting a role model ($n \gg 1$), then $(1-q)^n \approx 0$, implying they will almost certainly find an expert so long as one exists in their group. With these two assumptions, equation (2) can be approximated by:

$$Q(s,d) \approx \; (1-s)(1-d)q(1-e) + (1-s)dq + s(1-d)(1-e) + sd \tag{3}$$

Equation (3) lists the outcomes for different types of social learning. When there is no reliance on selective or deferential imitation ($s = 0$, $d = 0$), an imitator acquires the adaptive behavior with probability $q(1-e)$. With no reliance on selective imitation but full reliance on deferential imitation ($s = 0$, $d = 1$), an imitator acquires the adaptive behavior with probability $q$. With complete selective imitation but no deferential imitation ($s = 1$, $d = 0$), an imitator acquires the adaptive behavior with probability $1-e$. Finally, with full reliance on selective and deferential imitation ($s = 1$, $d = 1$), an imitator will acquire the adaptive behavior with certainty.

Assuming that the evolutionary dynamics of the two psychological biases (represented by $s$ and $d$) are slow compared to the cultural dynamics of adaptation, we can set $q' = q$ to find the equilibrium frequency of the adaptive behavior ($\hat{q}$), resulting in:

$$\hat{q} \approx (1-m)\frac{I + s(1-I)\big[1-e(1-d)\big]}{1-(1-m)(1-I)(1-s)\big[1-e(1-d)\big]} \tag{4}$$

**INSERT FIGURE 1 HERE**

## Random Imitation

Let's begin with the base case: neither selective nor deferential imitation. Setting $s = 0$ and $d = 0$, equation (4) reduces to:

$$\hat{q} = (1-m)\frac{I}{1-(1-m)(1-I)(1-e)} \tag{5}$$

Equation (5) is plotted in figure 1. To gain further insight, we can assume that the rates of migration, innovation, and the social learning error are low ($m \ll 1$, $I \ll 1$, and $e \ll 1$), resulting in:

$$\hat{q} \approx \frac{I}{m+e+I} \tag{6}$$

With random imitation ($s = 0$ and $d = 0$), cultural adaptation is constrained by ratio of innovation ($I$) to migration ($m$) and social learning error ($e$) rates. When the rate of innovation is much larger than the sum of migration and social learning error rates ($I \gg m + e$), the population will be well adapted to the environment ($q \approx 1$). When innovation rates are much lower than the sum of migration and social learning error rates ($I \ll m + e$), the population will be poorly adapted to the environment ($q \approx 0$). With purely guided variation, the rate of adaptation is constrained by two factors, the migration rate and the difficulty of social learning.

**INSERT FIGURE 2 HERE**

## Selective Imitation

Next, let's look at the effect of selective imitation without any deferential imitation. Setting $s = 1$ and $d = 0$, equation (4) reduces to:

$$\hat{q} \approx (1-m)\big[I + (1-I)(1-e)\big] \tag{7}$$

Equation (7) is plotted in figure 2. Again, making the assumption that the rates of migration, innovation, and the social learning error are low ($m \ll 1$, $I \ll 1$, and $e \ll 1$), we can approximate equation (7) with:

$$\hat{q} \approx 1 - m - e \tag{8}$$

Adding selective imitation has lifted the constraint of innovation. With selective imitation, the adaptation of the population is limited only by the rates of migration and social learning error. Selective imitators expend sufficient time and energy to evaluate a set of role models to increase the probability of imitating from the successful. Thus, an innovation somewhere, however rare, becomes an innovation everywhere, for all to share.

## Deferential Imitation

We now turn to deferential imitation without any selective imitation. Setting $s = 0$ and $d = 1$, equation (4) reduces to:

$$\hat{q} \approx (1-m)\frac{I}{1-(1-m)(1-I)} \tag{9}$$

Equation (9) is plotted in figure 2. Making the assumption that the rates of migration, innovation, and the social learning error are low ($m \ll 1$, $I \ll 1$, and $e \ll 1$), we can approximate equation (9) with:

$$\hat{q} \approx \frac{I}{m+I} \tag{10}$$

With deferential imitation, the constraint of social learning error has been eliminated. Now, the adaptation of the population is driven by the relative magnitudes of migration and innovation. When innovation rates are much higher than migration rates ($I \gg m$), the population is well adapted to its environment ($q \approx 1$). When migration rates are much larger than innovation rates, the frequency of adaptation is low.

## Selective Deferential Imitation

Finally, assuming completely selective, deferential imitation ($s = 1$ and $d = 1$), equation (4) reduces to:

$$\hat{q} \approx 1 - m \tag{11}$$

Equation (11) is plotted in figure 2. When imitators are highly selective in choosing role models and then pay a deference cost to gain access, thereby eliminating errors associated with social learning, the only force mitigating adaptation is migration. This is the core of Henrich and Gil-White's (2001) argument. A prestige-bias can vastly improve social learning.

# Evolutionary Dynamics

That a prestige-bias amplifies guided variation, increasing adaptation at the population level, is unsurprising. It is not, however, obvious whether natural selection will favor such a bias. A prestige-bias doesn't come for free. In this model, imitators can pay two types of cost in order to increase the probability they acquire adaptive information. They can pay a cost to increase the probability that they imitate from a successful model, which I have called selective imitation, and they can pay a deference cost to gain proximity to their chosen model thereby decreasing social learning errors, which I have called deferential imitation. These costs are borne by individuals, not populations. We can now ask under what conditions natural selection favors the payment of these costs in order to increase the success rate of social learning.

Letting $w$ denote fitness, $w_0$ the baseline fitness, $Q$ the probability of socially learning the adaptive behavior averaged across the different social learning biases, and $D$ the maximum deference cost, the fitness is given by:

$$w = w_0 + b(1-m)\big[Q + I(1-Q)\big] - sc_s - dc_d - Dd(1-\hat{q}) \qquad (12)$$

In this model, adolescents first attempt to socially learn the adaptive behavior, succeeding with probability $Q$, given by equation (3), substituting in equation (4) for $q$. This assumes that the dynamics of culture are fast relative to the dynamics of selection on the psychology underpinning social learning biases.

If an individual is unsuccessful in social learning, which happens with probability $1-Q$, he attempts to innovate, succeeding with probability $I$. If the possessor of adaptive information doesn't emigrate upon reaching adulthood, which happens with probability $1-m$, he receives a marginal fitness benefit $b$.

A social learner can pay two kinds of costs based on his reliance on the two types of social learning biases. A social learner pays a fraction of the cost of selective imitation ($c_s$) based on his reliance on selective imitation ($s$). This cost includes the time and energy necessary to find $n$ models and rank them in terms of success. The deferential imitator also pays a deference cost relative to his reliance on deferential imitation ($d$), which has two components. There is a fixed cost of engaging in deferential imitation ($c_d$) which includes the time and energy necessary to successfully imitate a model.

There is also a variable cost component to deference, which is determined by a market-like process. When the adaptive behavior is common ($q \approx 1$), the deference cost will be negligible. When the adaptive behavior is rare ($q \approx 0$), an expert can demand nearly all of the imitator's marginal gain from switching from distance learning to deferential imitation. Let $D$ represent the marginal fitness difference between deferential and distance learning:

$$D = be(1-m)(I-I)\big[(1-s)\hat{q} + s\big] - c_d \tag{13}$$

Imitators face a choice: learn from a distance and succeed with probability $1-e$, or pay the cost of deference, gain proximity to the expert, and succeed with certainty. $D$ thus represents the fitness gain (or loss) of switching from no deference to full deference, assuming the population engages in selective and deferential imitation with probabilities $s$ and $d$. This fitness difference increases with the benefit of the adaptive behavior ($b$), the rate of social learning error ($e$), the degree to which the population relies on selective imitation ($s$), and the frequency of the adaptive behavior in the population ($\hat{q}$). $D$ decreases with the migration rate ($m$), the innovation rate ($I$), and the fixed-cost associated with deference ($c_d$).

## The evolution of selective imitation

Following Henrich and Gil-White (2001), we study the evolution of prestige-biased transmission in a sequential fashion, first investigating the evolution of selective imitation, then the evolution of deferential imitation. In this section, we assume that there is no deference and find the optimal level of selectivity in finding an expert from whom to imitate. Setting $d = 0$, the fitness function (equation (12)) reduces to:

$$w = w_0 + b(1-m)\big[Q + I(1-Q)\big] - sc_s \tag{14}$$

And the equilibrium frequency of the adaptive behavior (equation (4)) reduces to:

$$\hat{q} \approx (1-m)\frac{I + s(1-I)(1-e)}{1-(1-m)(1-I)(1-s)(1-e)} \tag{15}$$

To find the optimal amount of selectivity, we take the partial derivative of fitness with respect to selectivity and set it equal to zero ($\partial w/\partial s = 0$), which results in the following quadratic equation:

$$0 = \begin{array}{l} b(1-m)(1-e)(1-I)\big[1-(1-m)[I+(1-I)(1-e)]\big] \\ -c_s\big[1-(1-m)(1-I)(1-e)\big]^2 \\ -2sc_s(1-m)(1-I)(1-e)\big[1-(1-m)(1-I)(1-e)\big] \\ -s^2c_s(1-m)^2(1-e)^2(1-I)^2 \end{array} \tag{16}$$

**INSERT FIGURE 3 HERE**

Figure 3 plots the optimal level of selectivity for a range of parameter values. To gain some insight regarding the factors which favor selective imitation, we can we assume that migration is weak ($m \ll 1$) and the rates of social learning error and innovation are low ($e \ll 1$ and $I \ll 1$). This simplifies equation (16) to:

$$0 \approx b(m+e) - 2sc_s(m+I+e) - s^2c_s(1-2m-2I-2e) \tag{17}$$

When there is no selective imitation ($s = 0$), this equation reaches it's maximum value of $b(m+e)$. So long as there is any migration ($m > 0$), there is some probability of committing an error in social learning ($e > 0$), and there is some marginal benefit of the adaptive behavior ($b > 0$), then selection will favor at least some selective imitation ($\partial w/\partial s|_{s=0} > 0$). That is, non-selective imitation is not an evolutionary stable equilibrium.

We can find when selection favors a complete reliance on selective imitation, by setting $s = 1$ and finding when equation (17) is positive. Because equation (17) is quadratic and $\partial^2 w/\partial s^2 < 0$ when $0 \leq s$, there is only one value of $s$ when $\partial w/\partial s = 0$. If $\partial w/\partial s|_{s=1} \geq 0$, then it means that this value of $s$ lies at 1 or more, implying that the evolutionary stable outcome is a complete reliance on selective imitation.

This threshold occurs when $b/c_s \geq 1/(m + e)$. Obviously, selection will favor more selective imitation as the ratio of the benefit of the adaptive behavior relative to the cost of being a selective imitator increases. This condition tells us that this benefit-cost ratio is related to the inverse of the the sum of migration and social learning error. When migration and social learning error rates are very low, then there is little reason to be a selective imitator. The frequency of the adaptive behavior will be high, and so imitating at random is a good strategy. As migration or social learning error rates increase, then the frequency of the adaptive behavior will decrease. When this frequency is low, then it pays to expend the time and energy to find an expert from whom to model behavior.

To summarize, unless there is no benefit to the adaptive behavior ($b = 0$), there is no migration ($m = 0$), and there is no possibility of committing an error in social learning ($e = 0$), selection will favor at least some reliance on selective imitation. This result is fairly obvious. If there is no benefit to the adaptive behavior, then there is no purpose being choosy in picking a role model. Similarly, with no migration and no social learning error, the population will be perfectly adapted to the environment, and so there is no need to be choosy in picking a role model. For the parameter values that interest us, we can conclude that selective imitation will occur to some degree.

More interesting, we see that selection favors a complete reliance on selection imitation when the ratio of the benefit of the adaptive behavior relative to the cost of being selective is larger than the inverse of the sum of migration and social learning error rates. When the cost of being selective is small compared to the benefit of the adaptive behavior ($b \gg c_s$), then even small amounts of migration and social learning error will result in complete reliance on selective imitation. When this condition is not satisfied, then the population will engage in some intermediate amount of selective imitation.

## The evolution of deference

Now, let's assume that full selectivity is favored ($s = 1$), and find the optimal level of deference. The fitness function (equation (12)) reduces to:

$$w = w_0 + b(1 - m)\big[Q + I(1 - Q)\big] - c_s - dc_d - Dd(1 - \hat{q}) \tag{18}$$

The equilibrium frequency of the adaptive behavior (equation (4)) reduces to:

$$\hat{q} \approx (1 - m)\big[I + (1 - I)\big(1 - e(1 - d)\big)\big] \tag{19}$$

And, the maximum variable component of deference (equation (13)) reduces to:

$$D = be(1 - m)(I - I) - c_d \tag{20}$$

To find the optimal value of deference, we take the partial derivative of fitness with respect to the reliance on deference, resulting in:

$$\partial w/\partial d = D(1 - m)\big[1 + e(1 - I)(2d - 1)\big] \tag{21}$$

From equation (21), we see that fitness increases with deference, so long as $D > 0$. If $D$ is positive, selection will favor full reliance on deference ($d = 1$); if $D$ is negative, selection favors no deference ($d = 0$). $D$ will be positive when the following condition is satisfied:

$$\frac{b}{c_d} = \frac{1}{e(1 - m)(1 - I)} \tag{22}$$

This condition is plotted in figure 4. To get a better sense of what this means, we can assume that innovation rates are low ($I \ll 1$), social learning error rates are low ($e \ll 1$), and the rate of migration is low ($m \ll 1$), resulting in:

$$\frac{b}{c_d} \approx \frac{1}{e} \tag{23}$$

When the rate of social learning error is low ($e \ll 1$), then it only pays to be a deferential imitator when the benefit from the adaptive behavior is much larger than the cost associated with spending the time with the expert to become an expert ($b \gg c_d$). This is fairly obvious; when it's easy to learn from a distance, deference only makes sense when it is relatively cheap. As the rate of social learning error increases, then selection favors deference, even when considerable time and effort are needed to master the skill.

**INSERT FIGURE 4 HERE**

# Discussion

In this paper, I developed a model of the evolution of prestige-bias. Overall, the results support Henrich and Gil-White's (2001) argument: a prestige-bias in social learning increases ability of the population to adapt to its environment, and, for a broad range of conditions, selection favors a bias to imitate the prestigious.

I modeled prestige-bias with two components. Imitators can be selective, choosing from an expert from a set of models, rather than random with respect to models. And, imitators can be deferential, paying an expert to gain proximity and thereby increase the success of social learning, rather than error-prone distance learners. Relying on random imitation and learning from a distance, the adaptiveness of a population is constrained by the rates of migration, innovation, and social learning error. When the rate of innovation is large relative to the sum of migration and social learning error, the population will be well adapted. Here, social learning biases don't greatly improve adaptation.

When the sum of migration and social learning error are large relative to the rate of innovation, then the population will be poorly adapted. Now, social learning biases have room to better adapt populations. Selective imitation eliminates the constraint of low rates of innovation. When imitators expend sufficient time and energy to find an expert, then there need only be a few individual experts for the population to be well adapted. Deferential imitation eliminates the constraint of social learning error. When imitators expend sufficient time, energy, and resources to defer to an expert, they are certain to acquire the adaptive behavior.

When these two social learning biases are combined, adaptation is constrained only by migration.

While it's not surprise that social learning biases increase adaptation, it's not obvious whether natural selection will favor such biases. After all, social learning biases generate positive externalities, increasing the frequency of adaptive behavior, thereby making it easier the imitation of others more successful. The costs associated with biased learning, however, are strictly personal. Selective imitation is favored when social learning is error prone and migration is high. If we assume that the benefit of adaptive behavior is large relative to the cost of ranking a set of models, then we should expect a large degree of selectivity.

Like selective imitation, deferential imitation is favored when social learning is relatively error prone. This is fairly intuitive; if social learning from a distance were easy, there would be no reason to defer to an expert. Interestingly, selection doesn't favor intermediate amounts of deference. Either full or no deference is favored. Unlike selective imitation, migration rate opposes deference. This effect is weak until migration rates are substantial. When migration rates are high, the adaptive behavior will be more rare. As such, experts can demand more for access.

In this paper, I have assumed that social learning biases are continuous characters. However, biases can also be interpreted as dichotomous traits. That is, individuals either are or are not selective imitators. $s$ and $d$ then become frequencies within the population, rather than individuals' degrees of reliance. Looking at figure 3, there is a range of parameter values in which selection favors an intermediate amount of selectivity ($0 < s < 1$). If $s$ is interpreted as the frequency of selective imitators in the population, rather than the degree of selectivity, then there will be two types of imitators, selective and un-selective ones.

Social learning can be a form of scrounging, benefiting the imitator, but not necessarily the population (Rogers, 1988; Boyd and Richerson, 1995). In this model, selective imitators pay a cost to increase their success rate. Selective imitation also provides a benefit to others. By choosing to imitate from an expert, selective imitators act as filters, increasing the frequency of adaptive behavior in the population. Because of this, there is scope for second-order scrounging, here represented by un-selective imitators. The same doesn't seem to apply to deference. Selection either favors all deferential imitators or none.

Table 1: Model Parameters

| Parameter | Definition |
|---|---|
| $s$ | Degree of selective imitation |
| $d$ | Degree of deferential imitation |
| $q$ | Frequency of the adaptive behavior |
| $I$ | Innovation rate |
| $e$ | Social learning error rate |
| $m$ | Migration rate |
| $n$ | Size of the model set for selective imitators |
| $b$ | Benefit of possessing the adaptive behavior |
| $c_s$ | Cost of selective imitation |
| $c_d$ | Fixed-cost associated with deferential imitation |
| $D$ | Maximum variable cost associated with deference |
| $Q$ | Probability of successfully imitation |

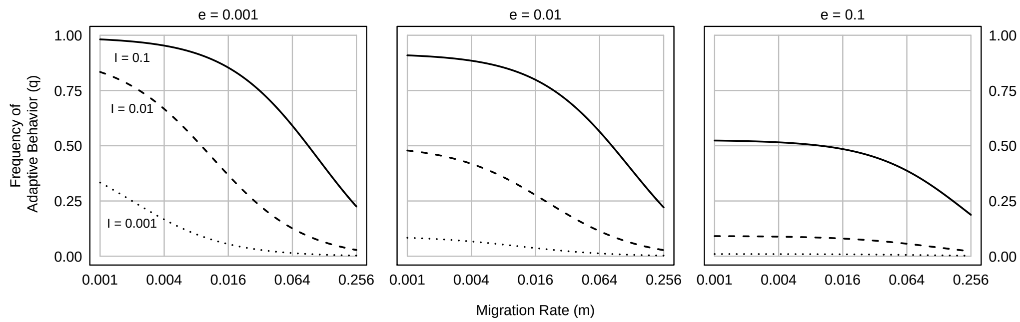


Figure 1: Frequency of the adaptive behavior when imitation is random with respect to model ($s = 0$) and imitators learn from a distance ($d = 0$). The horizontal axis shows the migration rate on a log scale. Each panel depicts a different rate of social learning error, ranging from 0.001 to 0.1. Within each panel, three innovation rates are depicted, ranging from 0.001 to 0.1.

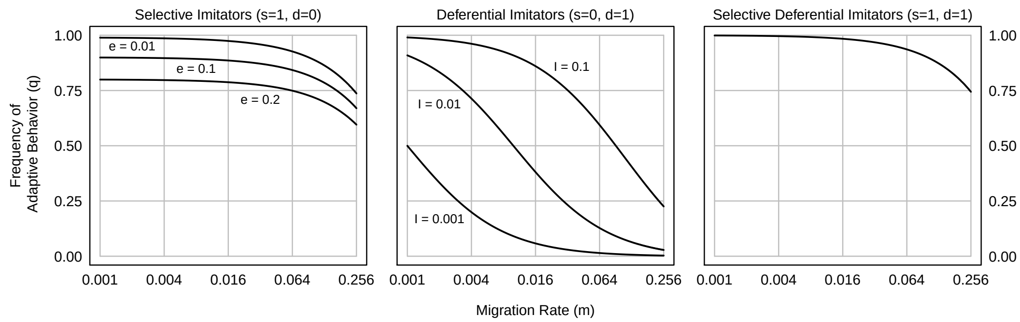


Figure 2: Frequency of the adaptive behavior for different combinations of biased social learning. The horizontal axis shows the migration rate on a log scale. In the first panel, imitators are selective with respect to models ($s = 1$), but learn from a distance ($d = 0$). Three different rates of social learning error are depicted, ranging from 0.001 to 0.1. In the second panel, imitators are not selective with respect to models ($s = 0$), but are deferential ($d = 1$). Three different rates of innovation are depicted, ranging from 0.001 to 0.1. For the third panel, imitators are both selective ($s = 1$) and deferential ($d = 1$). For these types, migration is the only force constraining adaptation. For panels 1 and 3, the size of the model set is assumed to be large ($n \gg 1$). For panel 1, the rate of innovation is fixed ($I = 0.001$). So long as the innovation rate is small ($I \ll 1$), it has little impact on the rate of cultural adaptation.

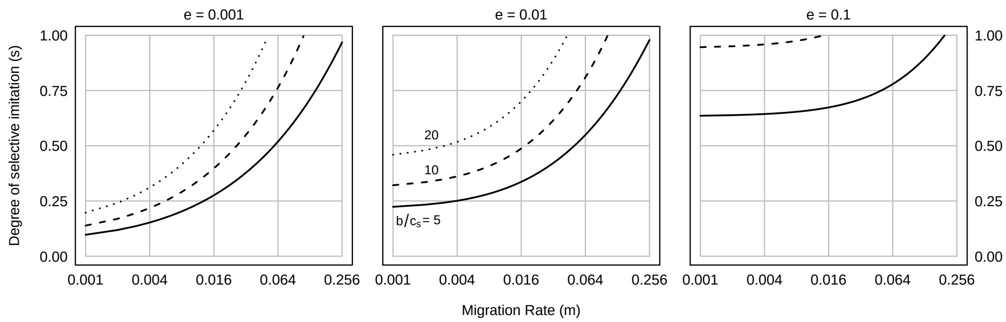


Figure 3: The optimal degree of selective imitation. The horizontal axis shows the migration rate on a log scale. Each panel depicts a different rate of social learning error, ranging from 0.001 to 0.1. Within each panel, different benefit to cost ratios are depicted, ranging from 5 to 20. The size of the model set is assumed to be large ($n \gg 1$). The rate of innovation is fixed ($I = 0.001$). So long as the innovation rate is small ($I \ll 1$), it has little impact on the results.

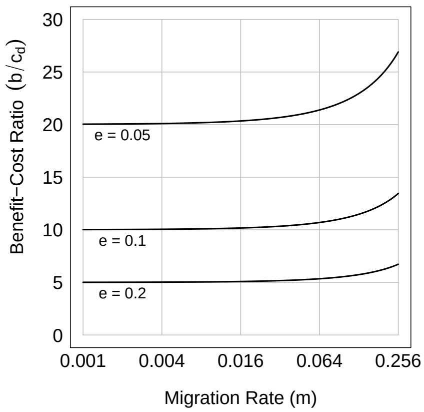


Figure 4: The optimal degree of deferential imitation. The horizontal axis shows the migration rate on a log scale. The vertical axis shows the ratio of benefit of the adaptive behavior to the fixed cost of deference ($b/c_d$). A range of social learning errors, ranging from 0.05 to 0.2, are depicted. When the benefit–cost is above the line, then selection favors full reliance on deferential imitation ($d = 1$). When the benefit–cost ratio is below the line, selection favors no deferential imitation ($d = 0$). The size of the model set is assumed to be large ($n \gg 1$). The rate of innovation is fixed ($I = 0.001$). So long as the innovation rate is small ($I \ll 1$), it has little impact on the results.