18 Rationality, Imitation, and Tradition

When the quality of information is poor, people often rely on tradition in making economic decisions. What is the best retail markup percentage? When should one refinance one's home? What is the right safety factor in designing a building? Retailers, homeowners, and engineers typically make such decisions using traditionally acquired rules-of-thumb. This tactic has both advantages and disadvantages. It can be useful because solving problems from scratch is difficult and costly. On the other hand, the uncritical adoption of traditional solutions to problems can lead people to acquire outmoded or even completely unfounded beliefs. Peasants sometimes resist beneficial innovations proffered by development agencies and retain traditional agricultural practices; many contemporary Americans maintain the unfounded belief that there are innate differences between the members of different ethnic groups.

The fact that tradition is sometimes reliable and other times misleading creates an interesting problem for economists. Traditions often work; when they do, they are useful because they reduce the costs of acquiring information and lower the possibility of making errors. However, if everyone were to depend exclusively on traditional rules, what would cause traditional rules to be modified in response to changes in the environment, and what would initially cause useful and reliable behaviors to become traditions?

Conventional economic theory is not helpful in answering this question (Conlisk, 1980). Economists have adopted the Bayesian theory of rational choice as the natural extension of the utility-maximizing view of human behavior when there is uncertainty and use it as a positive theory to predict people's behavior in a wide variety of contexts (Hirshleifer and Riley, 1978). Within the context of this theory, a person's beliefs about the world are represented as a subjective probability distribution. Once this distribution is specified, the theory tells us how rational people should behave and how they should modify their beliefs in accord with their experience. The theory does not tell us why people initially come to have the beliefs that they do but simply takes them as given.

The role of traditional knowledge has been discussed by some economists, but the processes that lead to sensible traditions seem to have been largely ignored. Hayek (1978) believes that limited knowledge and cognitive abilities force people to rely on traditional beliefs and values and argues that traditions are sensible because groups with favorable traditions survive longer and attract more members. Proponents of evolutionary models of firms (Alchian, 1950; Nelson and Winter, 1982) assume that beliefs, values, and other determinants of firm behavior are transmitted within firms and that these beliefs are shaped by the natural selection of firms. The only formal theoretical treatment of tradition seems to be the interesting article of Conlisk (1980) in which the individuals who optimize compete with individuals who acquire their behavior by imitation. If optimization is costly, Conlisk shows that imitation can persist in the population.

In this chapter, we introduce tradition into conventional theory by assuming that people acquire their initial subjective probabilities by imitating their parents, relatives, teachers, business associates, and friends, but otherwise behave as classical Bayesian rationalists. Several lines of empirical evidence support the assumption that people acquire their beliefs about the world by imitation and similar processes. Psychologists have shown that children readily acquire behavioral traits from moral beliefs to rules of grammar by imitating adult models (Bandura, 1977; Rosenthal and Zimmerman, 1978). Data collected on familial resemblances show high parent-offspring correlations for a wide variety of cognitive traits (I.Q.; Scarr and Weinberg, 1976), behaviors (child abuse, alcoholism; Smith, 1975), and indicators of beliefs (religious and political-party affiliation; Fuller and Thompson, 1960). A wealth of anthropological data suggests that human groups possess considerable cultural inertia; members of groups with different cultural histories behave quite differently even when living in similar environments (e.g., Edgerton, 1971). There is also evidence that individuals acquire new beliefs by imitation when they enter organizations such as business firms (Van Maanen and Schein, 1979) and that this process causes distinct cultures to develop in different organizations. (This body of evidence is reviewed in more detail in Boyd and Richerson, 1985:38–60.)

The assumption that people acquire their beliefs by imitation leads to models that keep track of the processes that change the frequency of alternative beliefs in a population of decision makers. To understand why a particular person acquires a particular set of beliefs, we must know to what kinds of behavior naive individuals are exposed. This in turn will depend on the distribution of beliefs (and thus behaviors) that exist in the population. A person in a village in which many people have adopted modern farming practices is more likely to acquire the beliefs that underlie such practices than a person exposed only to traditional lifeways. To predict the distribution of beliefs in the population at some future time, we must know the present distribution of beliefs and account for all of the processes that change that distribution through time. Here we present several such models of cultural change. For a more extensive exposition of our views, see Boyd and Richerson (1985), and for related work, see Pulliam and Dunford (1980), Cavalli-Sforza and Feldman (1981), Lumsden and Wilson (1981), and Rogers (1989).

These models are different from Conlisk's in two important ways: (1) Conlisk regards imitation as an alternative to optimization; individuals are either imitators or optimizers. We assume that imitation is a precondition for optimization; everyone must acquire beliefs about the world before they can optimize. (2) Conlisk simply posits dynamical relations between variables that describe a whole population of decision makers; we are more concerned to show how the details of individual imitation and decision-making processes lead to the dynamics of the distribution of beliefs in a population through time. As we shall see, the optimal behavior in these models is usually for individuals to mix imitation and individual decision making, depending on how the temporal dynamics work out.

We think that there are three lessons to be drawn from our theory of traditions: first, there are plausible circumstances in which it is optimal to depend nearly completely on tradition at equilibrium. Second, there are plausible genetic and cultural mechanisms that could cause people to achieve this equilibrium. Third, when people do depend largely on tradition, processes other than individual choice may have important effects on why people behave the way they do. We will begin by modeling a reference case in which people acquire their initial subjective probabilities by imitation and then modify them in accordance with their own experience in a uniform and constant environment. This model indicates that when beliefs are transmitted culturally, greater reliance on tradition always leads to higher expected utility. We will then add environmental variability to the model. When the optimal behavior varies because individuals encounter different environments, there is an optimal level of dependence on tradition. If there is a substantial chance that individuals and the people that they imitate experience the same environment, and if the information available to update priors is poor, it can be an evolutionary equilibrium to rely almost completely on tradition. In the simplest model, a population of such individuals will, on the average, behave almost as if they were perfect-information optimizers. However, in such a population other processes, which can lead to both beneficial (but poorly understood) beliefs or deleterious superstitions, may also be important. Finally, we will argue that there are cultural processes that may cause people to be characterized by an optimal reliance on tradition.

The Basic Model

In the first and simplest model there are only two processes that affect the distribution of beliefs in a population of decision makers. First, individuals use available information to update their subjective probability distributions. Second, the frequency of different beliefs is changed by the transmission of these beliefs to another generation. The model has three parts: a description of how single individuals modify their beliefs in light of their experience (a process we refer to as "individual learning"), a consideration of how individual learning affects the distribution of beliefs in a population of individuals, and a mechanism for passing one generation's beliefs to the next.

Consider the following very simple decision problem. An individual decision maker has the following utility function:

$$u(y,z) = -u_0(z-y)^2 \tag{1}$$

where z is a decision variable under his control, y is a variable that represents the state of the world, and u_0 is a constant. While the quadratic form of this utility function is unconventional in the theory of the consumer, it is a mathematically convenient representation of the usual view of individual choice. To see this, consider the following example: suppose that the decision maker is a young professional just beginning his or her career and that z represents the amount of time devoted to career advancement. The remainder of the young professional's time, t, is devoted to family and recreation. Then t and z are arguments of a personal "production function," which gives amounts of various "commodities," for example, income and marital happiness, produced for each combination of t and z. The consumption of these commodities in turn generates utility. By using the constraint that total time is fixed and assuming that the young professional's personal production and utility functions have the appropriate convexity properties, one could derive a unimodal function giving utility as a function of z. The optimum value of this function, y, would depend on the properties of the personal production function, which in turn will depend on the state of the world. For example, the relationship between time devoted to work and income might depend on what kind of firm the young professional has entered. While the utility function so derived is unlikely to be exactly quadratic, this functional form is a reasonable caricature of a more general unimodal function. In fact, one could think of it as the first two terms of a Taylor's series expansion of an arbitrary utility function in the neighborhood of the optimum. Because we have not specified how commodities map onto utilities, this model can represent any degree of risk preference.

The individual does not know the value of y with certainty, but his or her beliefs about the likelihood that y takes on various values conform to a normal probability distribution with mean $\hat{y}$ and variance L. Note that y is not a random variable; in a given environment there is an optimum amount of time devoted to career. The probability distribution describes the decision maker's subjective beliefs about what value of z is optimum.

Before making his or her choice, the decision maker has the opportunity to review a certain amount of evidence about the state of the world. For example, by observing the effects of time devoted to work on career advancement and home life, the young professional could get an estimate of the optimal amount of time to devote to work. Because our young professional's initial rate of advancement and domestic satisfaction might depend on a variety of factors other than the amount of time devoted to work, this estimate will be imperfect. Suppose that this evidence can be quantified by the variable x. The decision maker believes (correctly) that the value of x is normally distributed with mean y

and variance V_e. After using this evidence and Bayes's law, the decision maker's updated subjective probability distribution is normal with mean $\hat{y}'$ where

$$\hat{y}' = \frac{V_e\hat{y} + Lx}{V_e + L} \tag{2}$$

To simplify the development here, assume that the decision maker does not update the variance of his or her subjective probability distribution.

The decision maker uses the updated distribution to calculate his or her expected utility as a function of z:

$$E\{u(z,y)|\hat{y}',x\} = -u_0[(z-\hat{y}')^2 + L] \tag{3}$$

and, thus, the value of z that maximizes his or her expected utility, z^* is the following:

$$z^* = \hat{y}' \tag{4}$$

That is, the optimal behavior is the individual's posterior estimate of the most likely state of the environment.

Now, suppose that there is a large population of decision makers. The individuals who make up this population differ in only two respects: (1) they have different prior beliefs about the most likely state of the world, and (2) they are exposed to different evidence about the state of the world. To formalize the first assumption, we assume that the frequency distribution of $\hat{y}$ *in the population* before the subjective probability distributions have been updated, $Q_t(\hat{y})$, is normal with mean M_t and variance B_t. Notice that this is a description of the population, not a probability density. To formalize the second assumption, we assume that the value of x experienced by each different individual is an independent random variable with the density $p(y)$, which has a mean equal to the true state of the world, y, and variance V_e. Otherwise, all individuals are identical; in particular, they all have the same utility function and their subjective probability distribution is characterized by the same value of L.

Let us now consider how the use of Bayes's law by individuals to modify their beliefs changes the frequency distribution of $\hat{y}$ in the population. The distribution of $\hat{y}$ in the population of decision makers after updating, Q'_t, is as follows:

$$Q'_t(\hat{y}) = \iint h(\hat{y}|\hat{y}',x)Q_t(\hat{y}')p(x)d\hat{y}'dx \tag{5}$$

where $h(\hat{y}|\hat{y}',x)$ is the conditional density of an individual's belief after updating, given that the individual had beliefs characterized by $\hat{y}'$ before updating and observed x. Then $Q'_t(\hat{y})$ is normal with this mean:

$$M'_t = \frac{M_tV_e + yL}{V_e + L} \tag{6}$$

and variance:

$$B'_t = \frac{B_tV_e^2 + V_eL^2}{(V_e + L)^2} \tag{7}$$

Thus, after updating, the mean value of $\hat{y}$ moves closer to the correct value, y; the variance may either increase or decrease depending on the magnitudes of B_t, V_e, and L.

So far, we have followed the usual practice of taking the decision maker's initial subjective probabilities as given. We are now in a position to consider the effect of the transmission of these beliefs to another "generation" of decision makers by imitation. For example, suppose that the young professionals advance in their firm and are eventually replaced by a new cohort of entry-level professionals, who form a new population of decision makers and face the same decision problem that their predecessors faced. Initially the individuals in this second "generation" are naive; they have no beliefs of any kind about how much time should be devoted to work. However, each naive individual has been able to observe n models of behavior of the previous generation of professionals. Based on the behavior of their models, naive individuals are able to infer what each model believes about how much time should be devoted to one's profession. Then each of the naive individuals adopts the mean of the n inferred values of $\hat{y}$ that characterize their models as the mean of their own subjective probability distributions. We assume that the variance, L, remains constant at the same value as in the previous generation.

With these assumptions the distribution of $\hat{y}$ in the population just before updating in generation, $t+1, Q_{t+1}(\hat{y})$, is normal with mean, $M_{t+1} = M'_t$, and variance, $B_{t+1} = (1/n)B'_t$. Because the distribution of $\hat{y}$ remains normal, the state of the population of decision makers at any time can be specified by the mean and variance of $\hat{y}$. If the environment remains constant, the values of the mean and variance in the population will eventually reach a unique stable equilibrium, $\hat{M}$ and $\hat{B}$, where

$$\hat{M} = y \tag{8}$$

and

$$\hat{B} = \frac{V_e L^2}{n(V_e + L)^2 - V_e^2} \tag{9}$$

Equations (6) and (8) say that the effect of the repeated application of Bayesian inference and accurate imitation on the mean value of $\hat{y}$ is unambiguous: the average of the best guesses about the state of the environment in the population converges monotonically to the actual state of the environment. According to (7) and (9), however, the variance of $\hat{y}$ is affected by competing processes. New variation is introduced each generation by errors in individual learning; this process acts to increase $\hat{B}$. On average, however, inference causes beliefs about the environment to become more accurate, and this decreases $\hat{B}$. Finally, if $n > 1$, imitation itself acts to decrease the variance of $\hat{B}$ in the population.

The Evolutionary Stable Amount of Tradition

The relative importance of tradition and individual learning is determined by the relative magnitudes of the width of each individual's initial prior probability distribution (L) and the quality of the information available to individuals (V_e). If L is small compared to V_e, young professionals' work habits will be mostly determined by the beliefs that they acquire by imitation. If L is large, the information that individuals gather for themselves will be more important.

In this section we determine the evolutionary stable, or ESS, value of L. To do this, we find the value of L that when common in a population has higher expected utility than slightly different values of L. One way to justify the ESS approach is to assume that L is a genetically variable character and that utility is monotonically related to fitness. The ESS value of L is the value that prevents the rare genotypes from invading under the influence of natural selection. Some models of cultural transmission have very similar properties to genetic ones, and for our immediate purposes, we can think of L as evolving under the influence of either process. Clearly, cultural and genetic transmission also differ in important ways, for example, in the timescale over which they are relevant. Variations in reliance on tradition among contemporary societies likely require a cultural explanation, while a genetic model would be appropriate for studying the evolution of humans from apes. The penultimate section of the chapter will address several explicitly cultural mechanisms that can lead to the ESS.

Consider a population in which most individuals have a learning rule characterized by the parameter value, L, and that has reached the associated equilibrium values $\hat{M}$ and $\hat{B}$. The expected utility of an individual whose learning rule is characterized by parameter L' is the following:

$$E\{u(\hat{y},x)\} = -U_0 \frac{V_e^2}{(V_e + L)^2}[(y=\hat{M})^2 + \hat{B}] + \frac{L^2 V_e}{(V_e + L)^2} \tag{10}$$

One can show that this expression for expected utility is concave with a global maximum at the value of L, $L^\dagger$,

$$L^\dagger = (y - \hat{M})^2 + \hat{B} \tag{11}$$

The term $(y - \hat{M})^2 + \hat{B}$ measures the closeness of the population's beliefs about the state of the world to its actual state; V_e measures the accuracy of the information gained by each individual through his own experience. Relation (11) (together with [1]) says individuals should rely on imitation in proportion to the accuracy of the distribution of beliefs. If $(y - \hat{M})^2 + \hat{B}$ is large compared to V_e, individuals should rely mainly on their own experience; if $(y - \hat{M})^2 + \hat{B}$ is small compared to V_e, then it is optimal to depend mainly on imitation. This expression does not depend on the assumption that the population is in equilibrium nor that the environment is constant.

Now, suppose that natural selection, or an analogous cultural process, favors L, which increases expected utility. Then because $\hat{B}$ is a function of L, the population will eventually reach an ESS value of L, L^*, such that $L^* = \hat{B}(L^*)$. Using the expression for $\hat{B}$ given in equation (9), one can show that the ESS amount of imitation is $L^* = 0$. At equilibrium, individuals will depend completely on tradition and totally disregard the evidence presented by one's own experience.

This result has an intuitive explanation. At equilibrium, the relative merit of tradition and learning depends on the relative "noisiness" of the two sources of information. Learning has two effects on the variance in the population. On average, learning causes individual's estimates of y to move toward the correct value and thus acts to reduce the variation in the population. However, errors made during learning increase the variation of the population. Once the population reaches equilibrium in a constant environment, the net effect of learning is to maintain erroneous beliefs in the population. Decreasing L always decreases $\hat{B}$. Thus, any process that acts to change L so as to increase expected utility will reduce L until experience plays no role in determining individual beliefs.

Heterogeneous Environments

There are good reasons to doubt the robustness of the conclusion of the previous section. So far, we have assumed that (1) every member of the population experienced the same state of the world, (2) the state of the world did not vary from generation to generation, and (3) all individuals had the same utility function. Relaxing any one of these assumptions reduces the usefulness of tradition. For example, consider a heterogeneous environment in which different individuals experience different states of the world, but in which there is some chance that individuals in one environment draw models from other environments. In a given environment, people's beliefs will tend toward the optimum in that environment, but drawing models from diverse environments will reduce the likelihood that an individual acquires beliefs that are appropriate to its own environment. The models in this section show that a substantial reliance on tradition may still be evolutionarily stable in a heterogeneous environment or in a population in which utility functions vary. We have shown elsewhere that this conclusion also holds true in an environment that changes through time (Boyd and Richerson, 1983, 1985: ch. 4).

The essential feature of a heterogeneous environment is that different individuals in the population experience different states of the world, formalized in terms of the value of y. Such variation might arise for many reasons. For example, different young professionals might work in different firms, practice different professions, or live in different regions. We will model heterogeneous environments by assuming that the probability that an individual in the population experiences the environment specified by the value y is given by a normal density function, $f(y)$, with mean 0 and variance H. Setting the mean to 0 can be done without loss of generality since it sets only the origin from which different environments are measured. The variance, H, is a measure of the amount of environmental variation.

Suppose that in the environment characterized by the value y, the frequency of individuals with a subjective probability distribution characterized by a mean $\hat{y}$ before updating is normal with mean $M_t(y)$ and variance $B_t(y)$. Then the mean and variance after updating in that environment are given by equations (6) and (7) with the appropriate value of y. Further, suppose that there is a probability $1 - m$ that given models experience the same environment that their naive imitators will experience and a probability, m, that models are drawn at random from the population as a whole. Thus, for example, some of a particular young professional's models might be drawn from another firm in which more (or less) dedication is required to succeed. This model also applies to a population of individuals who live in a uniform environment but whose utility functions have different optima.

With these assumptions, one can derive recursions for the mean and variance of the distribution of prior beliefs in each environment. One can show that the equilibrium mean in habitat y is shown here:

$$\hat{M}(y) = \frac{(1-m)yL}{mV_e + L} \tag{12}$$

Equation (12) says that in a heterogeneous environment on average individuals have incorrect beliefs about their environment. The mean value of $\hat{y}$ in any environment y results from the balance of two forces. The Bayesian learning process tends to move the mean toward the correct value for that environment, but the exposure to models drawn from other environments moves the mean toward the mean for the entire population, 0. To find the equilibrium variance, we proceed exactly as in the previous section.

By averaging the expressions for the equilibrium mean and variance over all habitats, and using the expression for the ESS value of L given by equation (11), one can calculate L^* in a heterogeneous environment. The results of this calculation are shown in figure 18.1, which plots the relative importance of imitation in determining behavior, $V_e/(L^* + V_e)$, as a function of V_e for several

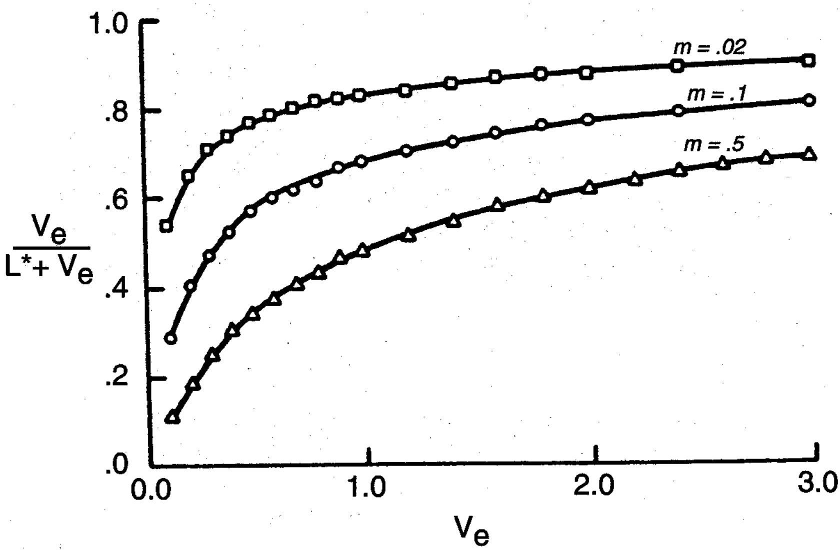

Figure 18.1. Plot of the fractional importance of tradition in determining behavior when the propensity to rely on tradition is at its equilibrium value, $V_e/(L^* + V_e)$, as a function of the quality of information available to individuals (V_e) assuming a heterogeneous environment, $n = 1$ and $H = 1.0$. Increasing values of m represent increasing amounts of mixing of models among different environments.

values of m. This figure indicates that the equilibrium optimum amount of imitation increases as the quality of the information available through individual experience declines and as the probability that models are drawn from foreign environments decreases.

These results make sense. The amount of imitation favored by evolutionary processes depends on the relative quality of two sources of information, the information available to individuals through their own experience and through observing the behavior of their models. As V_e increases, the quality of the information available to individuals through experience declines. As m decreases, the probability that an individual's models will exhibit behavior that is appropriate in the local environment increases. Thus, both increasing V_e and decreasing m cause the equilibrium value of L to increase.

These results suggest that the conclusions of the first section are not entirely misleading. When the amount of mixing between environments is not too large and information is of low quality, individuals achieve the highest expected utility by relying mainly on tradition. We think that this combination of circumstances is not uncommon. The world is complicated and poorly understood and the effects of many decisions are experienced over the course of a lifetime. In deciding how much time to devote to their families, young professionals must estimate not only the immediate effect on their careers and homelives but also the long-run effects on the development of their children's adolescent behavior. In such cases the information available to individuals may be very poor indeed, and it is plausible that they are best off relying almost entirely on traditional beliefs. Also notice that figure 18.1 is a worst case for tradition because it assumes that there is only one model ($n = 1$). As n increases, the equilibrium variance within environments decreases, and, therefore, tradition is relatively more reliable.

It is important to note that even when the amount of individual learning is small, it plays an important role in the evolutionary dynamics of the population. Some individual learning is necessary if traditional beliefs are to remain utilitarian in local environments in the face of imitation of experienced individuals from other environments. However, a relatively small amount of individual learning is sufficient to keep traditional behaviors on average reasonably near utilitarian optima, so long as mixing between heterogeneous environments is not too great.

Biased Imitation

To this point, we have assumed that individuals adopt a simple unbiased average of the beliefs of the models to which they are exposed. This may not be the most sensible procedure. It would seem better to preferentially imitate models whose behavior has been successful. Young professionals might imitate models who are particularly accomplished in their work and content in their private lives. More generally, naive individuals may imitate prosperous models, contented models, prestigious models, or devout models. By doing this, naive individuals will be more likely to acquire beliefs that lead to prosperity, devotion, contentment, or prestige. In this section we show how this form of biased cultural transmission can increase the frequency of correct beliefs in a population, even when individuals do not understand the causal connection between beliefs and their consequences.

Suppose that instead of simply averaging the beliefs of their models, naive individuals weight models according to their utility-models achieving higher utility having a greater influence on a naive individual's initial belief than individuals with lower utility. There are many plausible observable correlates of utility, such as level of consumption. It seems likely that by imitating individuals with higher levels of consumption, naive individuals might increase their chances of acquiring beliefs that lead to higher utility. In particular, suppose that the initial value of $\hat{y}$ acquired by a naive individual exposed to models with the utilities $u_1, \ldots, u_n$, and beliefs $\hat{y}_1, \ldots, \hat{y}_n$, is this expression:

$$\hat{y} = \frac{\sum_{i=1}^{n} \hat{y}_i(1 + bu_i)}{\sum_{i=1}^{n} (1 + bu_i)} \tag{13}$$

where b is a positive constant small enough that terms of order b^2 can be ignored.

With this assumption, it can be shown (Boyd and Richerson, 1985) the mean in the population after transmission is shown here:

$$M_{t+1} = M'_t + (1 - 1/n)B'_t E\,\{\text{Reg}[\hat{y}, u(\hat{y})]\} \tag{14}$$

where $E\{\text{Reg}(\hat{y}, u(\hat{y})\}$ is the regression of utility on $\hat{y}$ averaged over all possible sets of models. According to equation (14), the change in the mean due to biased transmission depends on two factors: the amount of variability within sets of models $[(1 - 1/n)B'_t]$ and the extent to which beliefs about the world are predictably related to utility $[E\{\text{Reg}[\hat{y}, u(\hat{y})]\}]$. Variability within sets of models is important because biased transmission is a culling process that works because some models are more attractive than others. If all models are identical, biased transmission can have no effect. The regression of utility on $\hat{y}$ is a measure of the average effect of a change in an individual's beliefs on his or her utility. If it is positive, individuals with larger values of $\hat{y}$ will have higher utility and, therefore, be more likely to be imitated. This will cause the mean value of $\hat{y}$ in the population to increase. Both the sign and the magnitude of $E\{\text{Reg}[\hat{y}, u(\hat{y})]\}$ depend on the distribution of $\hat{y}$ in the population. If M_t is less than the optimum value (y), larger values of $\hat{y}$ will on average lead to higher utility, and the regression will be positive. The reverse will occur if $M_t < y$. This means that biased transmission will leave the mean unchanged only if it is at the optimal value.

Biased transmission is of interest because it can explain the existence of "folk wisdom," beneficial but poorly understood customs. The preferential imitation of successful people will tend to increase beliefs and practices that lead to success; there is no need for individuals to understand the causal connection between traditional practice and success, even on the part of the individuals who invent the practices.

Natural Selection

So far we have assumed that the probability that a naive individual is exposed to models who are characterized by given beliefs (i.e., a given value of $\hat{y}$) is equal to the frequency of that kind of individual in the previous generation. There is good reason to suppose that this assumption is often violated. For example, the probability that young professionals are advanced in their firm is likely to depend on how much time they devote to work. Underachievers are likely to be fired and overachievers to be promoted. Thus, models who are available for imitation within a firm may represent a biased sample of the original population. More generally, if the behaviors that are shaped by the beliefs acquired by imitation are important, they may affect many aspects of individuals' lives: whom they meet, how long they live, how many children they have, or whether they get tenure. All of these factors could affect the probability that an individual becomes available as a model for others. This means that individuals characterized by some values of $\hat{y}$ will end up being more likely to be imitated than individuals with other values. All other things being equal, it is intuitive that this process, which we will term "natural selection" because of its close resemblance to the biological process, will increase the frequency of the variants most likely to "survive" to enter the pool of models. For a more extensive discussion of the natural selection of culturally transmitted behaviors, see Boyd and Richerson (1985:173–203).

To formalize this idea, we suppose that the probability that an individual who chooses behavior z becomes available as a model, $W(z)$, is the following:

$$W(z) = \exp\{-(z-w)^2/2K\} \qquad (15)$$

where w is behavior that maximizes the probability of being in the model pool and $1/K$ is a measure of the intensity of the selection process. Note w need not equal y; for example, individuals who devote more than the utility maximizing amount of time to their work may be more likely to be promoted within the firm.

Using (15) one can show that the mean value of $\hat{y}$ in the population of models (after selection), M_t'', in this equation:

$$M_t'' = \frac{M_t'K + wB_t'}{B_t' + K} \qquad (16)$$

Thus, selection moves the mean value of $\hat{y}$ in the population toward the value that maximizes the probability of entering the pool of models, w. One can also show that it reduces the variance of $\hat{y}$ in the population. The strength of both these effects is proportional to the variance in $\hat{y}$ in the population and the intensity of the selection process.

Natural selection is important because it explains how a reliance on tradition can lead to erroneous or deleterious beliefs. Many social and economic processes affect the kinds of individuals available as models. Some of these processes act on the level of the individual, as in the case of the young professional. Others affect whole firms or institutions. For example, firms composed of overachievers may be more likely to survive and expand than firms composed of utility maximizers. When culturally acquired beliefs are important in determining people's behavior, these selective processes will affect what kinds of people are available for imitation and therefore what beliefs will characterize the population. Since there is no reason to believe that such selective processes always favor utility maximizing behavior, selection may cause the most common beliefs in a population to be deleterious. Nonetheless, if information is imperfect and costly to acquire, it may still be sensible to rely on tradition; a modest systematic error may be preferable to a larger random error.

As an aside, we could also interpret the case of a naive manager being socialized by overachievers as the acquisition of a new utility function by considering that preferences are transmitted by tradition and modified by evolutionary processes such as selection. Such a model would allow a more general account of the relationship between learning and tradition than the Bayesian framework used here permits in order to reflect other models of the decision-making process (e.g., Nelson and Winter's, 1982, evolving "routines"). To enlarge on these problems is, however, outside the scope of this chapter. Here we want to emphasize that the standard, and normatively appropriate, Bayesian model is incomplete without a theory of tradition.

Cultural Mechanisms Leading to the ESS Amount of Imitation

So far we have assumed that natural selection acting on genetic variation or an analogous cultural process causes the value of L to change in the direction of increasing expected utility. In this section we consider such cultural processes in more detail. Suppose that the relative dependence on tradition versus one's own experience itself is a culturally transmitted trait. Then each of the three mechanisms we have just studied can, under the right circumstances, act like natural selection to change L in the direction that increases expected utility.

First, however, it is important to clarify why, within the context of the model outlined so far, it is not possible for individuals to choose directly the appropriate value of L. An essential assumption of this chapter is that the information available to individuals is limited; they know the results of their own direct experience and the observable behavior of the individuals whom they had available to imitate, but they do not know the optimum behavior, y. From equation (11), the optimal amount of imitation is given by the term $\hat{B} + (\hat{M} - y)^2$. Individuals can estimate $\hat{B}$ and $\hat{M}$ from their sample of models, and under some circumstances this information might be sensibly used to modify L. They cannot choose the optimum value of L, however, because that value depends on how close the mean belief in the population is to the optimum, y.

How do people acquire their attitudes toward tradition? Assume that people acquire their value of L by imitation during an earlier episode of social learning. With this assumption, any of the processes that change the frequency of a culturally transmitted trait could affect the evolution of the mean value of L in the population:

1. *Ordinary learning*. Individuals might acquire an initial value of L by imitation or teaching and then modify it in accordance with their experience. For example, during enculturation, individuals must acquire many different beliefs and behaviors. They might experiment

with different values of *L* during early episodes of learning, retaining the value that seems to yield the best results. This process would change the mean value of *L* among members of the population in the direction that increased average utility.

2. *Biased transmission.* Suppose that available models are variable, some of them relying on tradition to a greater degree than others. Moreover, suppose that naive individuals can observe some behavior of their models that serves as a useful index of the model's utility. Then if naive individuals are predisposed to imitate successful models, the mean value of *L* in the population will move toward the optimum. Notice that this can be true even if, as we have assumed, individuals have no understanding of why certain beliefs lead to higher utilities.
3. *Natural selection.* Once again assume that individuals vary in their attitudes toward tradition. Individuals with different values of *L* will, on average, behave differently. If an individual's behavior affects the probability that he or she becomes a model, natural selection will change the mean value of *L* in the direction that increases the chance of acquiring behaviors that make an individual likely to become a model. To the extent that there is a correlation between the utility associated with a behavior and the probability that an individual with the same behavior will become a model, natural selection would modify *L* in a utility maximizing direction.

To see how these processes might work, consider how attitudes toward tradition might change as a society undergoes industrialization. It is often thought that in pre-industrial agricultural societies people rely heavily on tradition. If one supposes that in such societies information is costly, then their reliance on tradition is sensible according to our model. Now, suppose that during industrialization, technical and institutional change makes information less costly. According to the model, people would be better off if they relied more on their own experience and less on tradition. This might come about by any of the three processes mentioned. To some extent, individuals might have been able to infer from their own experience that a lower reliance on tradition improved their lot. More plausibly, during industrialization people with a tendency to rely more on their own experience and less on traditional beliefs might more readily acquire nontraditional skills that lead to wealth and other kinds of observable markers of success. If successful individuals are more likely to be imitated, biased transmission would decrease average reliance on tradition. Or less traditional individuals might simply be more successful at becoming teachers, managers, and bureaucrats in modernizing societies. The natural selection mechanism could have favored a reduced dependence on tradition through differential achievement of roles that are important in socialization.

Invoking processes that affect earlier episodes of imitation to understand the nature of a subsequent episode clearly creates a problem of explanatory regress. Each of the three processes mentioned depends on some aspect of the imitation process, which then must be explained. In the case of ordinary learning, individuals must have some way of weighting the importance of the value of *L* that they acquired by imitation against the value that their experience indicates is best. Do they rely on their experience or on imitation? In the case of biased transmission, individuals must have some criteria of success—do they imitate wealthy individuals? Content individuals? Even natural selection will differ in its effects depending on whom naive individuals are prone to imitate. Are they disproportionately affected by their parents, or are other individuals important?

Ultimately, these are questions about human nature. The answers must be sought in the long-run processes that govern the interactions of cultural and genetic evolution in our species. This topic has been discussed at length by us (Boyd and Richerson, 1985) and others (Pulliam and Dunford, 1980; Lumsden and Wilson, 1981; Durham, 1978). Our work supports two generalizations that are relevant here:

1. If there is genetic variation that affects the tendency of people to imitate, natural selection will tend to modify this tendency so that it maximizes genetic fitness. Thus, to the extent that people prefer fitness-enhancing outcomes, selection would increase average utility.
2. There are a variety of conditions in which the fitness-maximizing values of L are near 1. Thus, it is plausible that even the earliest episodes of imitation are not directly subject to genetic influences.

Discussion

The economic theory of rational choice under uncertainty is incomplete because it is silent about the source of people's initial beliefs about the world. People are not immortal; sometime between birth and adulthood they acquire a set of beliefs about the world. Because rational behavior, including the rational response to new information, depends on the nature of an individual's prior beliefs, virtually any behavior can be rational, and therefore explicable, given some set of prior beliefs. A peasant's initial resistance to a beneficial innovation is explicable if one supposes that he believes that traditional ways are superior to modern ones. His ultimate rejection of modern practices may also be rational if his beliefs are described by "tight" priors.

In this chapter we have extended the economic theory of choice under uncertainty by assuming that individuals acquire their initial subjective probability distribution by imitation. In particular, we supposed that each naive individual observes the behavior of a number of experienced models sampled from a larger population, induces the belief that led to the observed behavior, and then adopts an average of those beliefs as his own initial beliefs. Then to understand why people acquire the initial beliefs that they do, we must understand why the population is characterized by a particular distribution of beliefs. This means that models that allow for imitation must account for all of the processes that will arise from individual learning and decision making, while others result from social and economic processes that have different effects on people with different beliefs.

This amendment to economic theory is not proposed as a behavioral alternative to the usual assumption that people are rational optimizers. Whether they are optimizers or not, mortal individuals must acquire their initial beliefs from others. It well may be that the particular model of imitation we have chosen is incorrect, that Bayesian optimizing is a poor model of how humans make choices, or that genetic inheritance is important in determining people's behavioral predispositions. In any case, we believe that a complete theory of human behavior would have a similar structure to the models outlined here; it would keep track of the dynamics of a population of decision makers by accounting for the processes that change the distribution of beliefs or other predispositions in the population. Some of these processes will result from people's attempts at improving their lot, while others will result from what happens to them because they hold the beliefs that they do.

There are two lessons that can be drawn from the models presented here: first, they suggest that a strong reliance on tradition may indeed be sensible. At equilibrium, individuals may rely almost entirely on traditional knowledge and ignore any other information that may be available to them. When (1) the quality of information available to individuals is low and improving it is costly, (2) there is a good chance that the individuals' models experienced the same environment that they experience. Traditional solutions to problems may be much closer to the optimal behavior, on the average, than the solutions that individuals could devise on their own.

The theory also suggests, however, that when traditions are substantially more important in determining people's beliefs than their own experience, a variety of processes other than individual learning may affect the commonness of different beliefs. When tradition is important, it acts like a system of inheritance to create heritable variation within and among groups. Processes like biased transmission and natural selection can then affect the frequency of different beliefs by making it more likely that some beliefs will be transmitted from one generation to the next. When the effect of individual experience is small, it is plausible that such processes may have an important effect on the way that people behave.

Some of these processes, such as biased transmission, may increase the frequency of utility-enhancing behaviors. This fact is of interest because it may explain "folk wisdom," that is, the fact that people hold beneficial traditions that they do not understand. The most striking examples of folk wisdom come from anthropological research. For example, in many parts of the New World native peoples treated maize as a strong base to produce foods such as hominy or masa as part of their traditional cuisine. Katz, Hediger, and Valleroy (1974) have shown that such treatment makes more of the amino acid lysine available (lysine is the least plentiful amino acid in maize). They have also shown that there was a strong negative correlation between the use of alkali treatment and the availability of protein from sources other than maize. Given that many factors influence nutrition, and that only small, uncontrolled samples were available, it is difficult to see how individuals in these cultures could have detected the effect of the treatment. Indeed, although Africans have been using maize as a staple for a few centuries, alkali cooking has not yet developed there. It seems more likely that it could spread because eating treated maize made people more successful or more likely to survive and, therefore, more likely to be imitated. Folk wisdom also plays a role in economic thinking. Hayek (1978) argues that traditional beliefs and institutional arrangements reflect wisdom beyond the ken of any individual, and he bases many political and economic prescriptions on this view. Similarly, proponents of an evolutionary view of the firm (e.g., Alchian, 1950; Nelson and Winter, 1982) argue that inherited decision rules that determine a firm's response to market conditions may be sensible in ways that nobody in the firm understands.

However, for other processes that affect the frequency of alternative beliefs in a population, such as natural selection, there is no guarantee that utility-maximizing behaviors will be favored. This may explain the existence of behavior that seems paradoxical under the usual assumption of individual rationality. In our example of natural selection on behaviors transmitted in the workplace, people could come to work harder than they would desire. Such behaviors could remain in a population because on average the traditions transmitted within a firm are more useful than alternative behaviors individuals could acquire by their own efforts. In other words, a reliance on tradition causes individuals to trade systematically suboptimal behaviors transmitted within the firm for the randomly suboptimal ones that can be discovered by individual effort. Elsewhere we show that processes other than natural selection can have this general effect (Boyd and Richerson, 1985).

Finally, models of the kind described here may also be useful in clarifying the relationship between human evolution and contemporary human behavior. Hirshleifer (1977) has argued that one of the attractive features of sociobiological theory is that it provides an independent way to derive utility functions; namely, human preferences have been shaped by natural selection so that, at least in the context of a hunter-gatherer society, they enhanced genetic fitness. While we are sympathetic to this general approach, we have argued (Boyd and Richerson, 1985) that many human preferences are difficult to explain on this basis. For example, many contemporary professionals seem to sacrifice genetic fitness by delaying marriage, reducing family size, and limiting time devoted to child care in order to gain professional success. Such behavior is explicable, however, if one imagines that individuals who value professional accomplishment for its own sake are more likely to rise to positions of influence than those with more "sociobiological" values. To take another example, humans cooperate in large groups of unrelated individuals to provide public goods (such as victory in warfare) in a way that seems difficult to reconcile with individual fitness maximization. In the work cited, we have shown how some forms of cultural transmission, permitting selection on culture at the level of groups, can arise from attempt to use traditions to enhance the ends of genetic fitness. To take advantage of the economies of information acquisition that tradition offers requires a measure of blind trust of traditional wisdom. Such weak rational control on tradition by its users may be sensible but at the same time allows culture to respond to blind evolutionary processes unique to the cultural system of inheritance. These processes may ultimately have important effects on what individuals prefer as well as on what they believe.

NOTE

We thank Robert Brandon, John Conlisk, Jack Hirshleifer, Richard Nelson, Eric A. Smith, John Staddon, Robert Seyfarth, Joan Silk, Michael Wade, and John Wiley for providing comments on an earlier version of this chapter; we also thank John Gillespie and Ron Pullman for crucial insights about modeling environmental variation and learning, respectively. As tradition dictates, we stipulate that any errors are our own.

REFERENCES

Alchian, A. A. 1950. Uncertainty, evolution and economic theory. *American Economic Review* 58:219–238.

Bandura, A. 1977. *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Boyd, R., & P. J. Richerson. 1983. The cultural transmission of acquired variation: Effect on genetic fitness. *Journal of Theoretical Biology* 100:567–596.

Boyd, R., & P. J. Richerson. 1985. *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Cavalli-Sforza, L. L., & M. W. Feldman. 1981. *Cultural transmission and evolution*. Princeton, NJ: Princeton University Press.

Conlisk, J. 1980. Costly optimizers versus cheap imitators. *Journal of Economic Behavior and Organization* 1:275–293.

Durham, W. H. 1978. Toward a coevolutionary theory of human biology and culture. In: *The sociobiology debate*, A. Caplan, ed. New York: Harper & Row.

Edgerton, R. B. (with W. Goldschmidt). 1971. *The individual in cultural adaptation: A study of four East African Peoples*. Berkeley: University of California Press.

Fuller, J. L., & R. W. Thompson. 1960. *Behavior genetics*. New York: Wiley.

Hayek, F. A. 1978. *The three sources of human values*. L. T. Hobhouse Memorial Trust Lecture 44. London: London School of Economics and Political Science.

Hirshleifer, J. 1977. Economics from a biological viewpoint. *Journal of Law and Economics*. 20:1–52.

Hirshleifer, J., & J. Riley. 1979. The analytics of uncertainty and information: An expository survey. *Journal of Economic Literature* 17:1375–1421.

Katz, S., M. Hediger, & L. Valleroy. 1974. Traditional maize processing techniques in the new world. *Science* 184:765–773.

Lumsden, C., & E. O. Wilson. 1981. *Genes, mind, and culture*. Cambridge, MA: Harvard University Press.

Nelson, R. R., & S. G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Pulliam, R., & C. Dunford. 1980. *Programmed to learn*. New York: Columbia University Press.

Rogers, A. 1989. Does biology constrain culture. *American Anthropologist* 90:819–831.

Rosenthal, T., & B. Zimmerman. 1978. *Social learning and cognition*. New York: Academic Press.

Scarr, S., & R. A. Weinberg. 1976. I.Q. performance of black children adopted by white families. *American Psychology* 31:726–739.

Smith, S. S. 1975. *The battered child syndrome*. London: Butterworths.

Van Maanen, J., & E. H. Schein. 1979. Toward a theory of organizational socialization. *Research on Organization and Behavior* 1:209–263.