

Forthcoming in the *New Palgrave Dictionary of Economics, Second Edition*, edited by Steven N. Durlauf and Lawrence E. Blume. London: Macmillan.

## Social Norms

The function of a social norm is to coordinate people's expectations in interactions that possess multiple equilibria. Norms govern a wide range of phenomena, including property rights, contracts, bargains, forms of communication, and concepts of justice. Norms impose uniformity of behavior within a given social group, but often vary substantially among groups. Over time norm shifts may occur, prompted either by changes in objective circumstances or by subjective changes in perceptions and expectations. The dynamics of this process can be modeled using evolutionary game theory, which predicts that some norms are more stable than others in the long run.

*Social norms* are customary rules of behavior that coordinate our interactions with others. Once a particular way of doing things becomes established as a rule, it continues in force because we prefer to conform to the rule given the expectation that others are going to conform (Lewis, 1969). This definition covers simple rules that are self-enforcing at a primary level, such as which hand to extend in greeting or which side of the road to drive on, and more complex rules that trigger sanctions against those who deviate from a first-order rule. (We express outrage if someone cuts in front of someone else in a queue.) The former are sometimes called *conventions* and the latter *norms* (Sugden, 1986; Coleman, 1990; Bicchieri, 2006), but in fact there are numerous gradations and levels of response to norm violation that make this dichotomy problematic. Hence I shall use the term 'norm' in its inclusive sense in what follows.

David Hume ([1739], 1978) was the first to call attention to the central role that norms play in the construction of social order. Norms define property rights,

that is, who is entitled to what. They determine what commodities are accepted as money. They shape our sense of obligation to family and community. They determine the meanings we attach to words. Indeed it is hard to think of a form of interaction that is not governed to some degree by social norms. (For book-length treatments of the subject see Lewis, 1969; Ullman-Margalit, 1977; Sugden, 1986; Young, 1998a; Posner, 2000; Hechter and Opp, 2001; Bicchieri, 2006.)

*Norms and equilibria.* Norms can be represented as equilibria of suitably defined games; indeed, Hume's analysis of norms can be viewed as one of the earliest examples of game-theoretic reasoning. Nevertheless, not every equilibrium of a game is a norm. First, the term generally applies only to games with *multiple* equilibria. People can queue for service *or* they can push. They can use gold coins *or* glass beads as money. They can drive on the left *or* on the right.

Second, even if a game has multiple equilibria, they do not necessarily qualify as norms. To illustrate the distinction, consider two individuals who get to divide a dollar provided they can agree on how to divide it. Each makes a demand, and if the demands sum to at most one the demands are met; otherwise they get nothing. This is a coordination game and it has many equilibria. For example, if one person demands 43 cents and the other demands 57 cents, the demands are in equilibrium: no one can gain from a unilateral deviation. But this is not a norm; it is an idiosyncratic equilibrium for these two individuals. Fifty-fifty division, by contrast, is a norm because it is usual and customary in games of this kind, and everyone knows it.

*Norm enforcement.* Broadly speaking there are three different mechanisms by which norms are held in place. Some are sustained by a pure coordination motive. If it is the norm to drive on the left, I adhere to the norm in order to avoid accidents. If gold is the commonly accepted currency, it would be a waste

of time to try to conduct my business with glass beads. These are “social” phenomena, because they are held in place by shared expectations about the appropriate solution to a given coordination problem, but there is no need for social enforcement.

Other norms are sustained by the threat of social disapproval or punishment for norm violations ( Sugden, 1986; Coleman, 1990). If queuing is the norm, I will be censured if I try to push my way to the front. If dueling is the proper response to an insult, I will lose status in the community if I do not challenge the one who insulted me. If I am expected to avenge the murder of my brother and fail to carry it out, I may be ostracized by other family members. (Exactly why third parties bother to express disapproval or carry out punishments is a matter of debate, but the evidence suggests that they sometimes do so even at considerable personal cost (Fehr, Fischbacher, and Gächter, 2002).)

A third enforcement mechanism arises through the internalization of norms of proper conduct. If it is the norm not to litter, I will avoid littering even in situations where no one can see me. If I eat a meal in a foreign city and fail to tip the waiter, I need not fear the consequences because there is no continuing relationship; nevertheless I may think the worse of myself for having done it. More generally, norms often take on the character of virtuous or right action (Hume [1739], 1978), and departures from a norm can trigger emotions of shame or guilt even when third party enforcement is absent (Coleman, 1990; Elster, 1989, 1999). This fact is especially useful in large-scale societies, where it may be difficult to monitor compliance with equilibrium behavior that entails sanctions by third parties.

*Norms and efficiency.* It remains to be explained why a dictionary on economics, in contrast to sociology or law, should bother with an entry on social norms.

What economic purpose do they serve? The answer is that norms coordinate expectations, and thereby reduce transaction costs in interactions that possess multiple equilibria (Wärneryd, 1994).

This point seems clear enough intuitively; it can also be demonstrated experimentally (Roth, 1985). Consider a game in which two players can divide a pile of chips in any way they like, but if they fail to agree on a division within a specified period of time they forfeit all of them. When all the chips can be cashed in for the same amount of money, the norm is to divide the chips equally; moreover the great majority of players do in fact coordinate in this manner. Notice, however, that *any* division of the chips, not just fifty-fifty, can constitute an equilibrium of the one-shot game if both players expect that it will be played. Now consider a variant in which the players get to cash in the chips for different amounts of money (which is publicly known). In this case there are two potential focal solutions -- divide the chips equally and divide the money-value of the chips equally -- but there is no norm to steer the players' expectations toward one or the other. As a result, the frequency of disagreement rises substantially.

More generally, a norm has economic value if it creates a uniquely salient or focal solution to a coordination problem, thus reducing the risk of coordination failure. In this sense norms are a form of social capital (Coleman, 1987).

This does not mean, however, that norms are invariably efficiency-enhancing. Indeed, some norms would appear not to have *any* direct welfare implications. Consider norms of etiquette, such as the fine points of table manners. The welfare consequences are so trivial that it is hard to see why anyone bothers with them. No one is harmed, for example, if I wear a hat to dinner or eat peas with my fingers. The fact is, however, that such indiscretions may do serious harm to my reputation. In particular, they signal that I am a person who does not care

about social norms, which may lead others to doubt my reliability in more important interactions (Posner, 2000). Complex social rituals allow people to signal their sensitivity to *norms in general*; they also provide a training ground for learning to follow norms, and for disciplining those who fail to do so.

Even when norms do have direct welfare implications, one cannot conclude that societies will opt for efficient norms. It is doubtful, for example, that norms of retribution are efficient, or that pushing is superior to queuing. Yet these are the operative norms in quite a few cases. The problem is that norms are not “chosen;” they arise from historical accident and the accumulation of precedent. Once expectations converge on an inefficient norm, it can be very difficult to dislodge.

Over the longer run, it is conceivable that societies could somehow extricate themselves from inefficient outcomes. One way that this could happen is that societies with superior norms simply displace societies with inferior norms, through growth, conquest, or migration. Another possibility is that societies with inferior norms imitate the practices of more successful ones (Robson and Vega-Redondo, 1996; Boyd and Richerson, 2002). Yet a third possibility is that norm change comes from within, the result of gradual and almost imperceptible changes in expectations that “tip” the society into a new way of doing things without anyone intending it. I shall discuss this possibility in more detail below.

*Excess uniformity.* I have argued that some norms may be inferior from the standpoint of welfare, yet stay in place for long periods of time. Others may have little direct effect on welfare but serve an important signaling function. Another way in which norms can affect economic welfare is by imposing excess uniformity on behavior. People might be better off if they adapted their actions to their particular circumstances. To illustrate, consider a situation in which a

principal and an agent are bargaining over the terms of a contract. To be concrete, think of a landlord bargaining with a prospective tenant over the terms for renting a plot of land. In theory, the optimal contract will depend on a variety of factors that may be idiosyncratic to the contracting parties, including information asymmetries, monitoring costs, and attitudes toward risk (Cheung, 1969). Yet, in practice, contracts often exhibit a high degree of uniformity, and employ 'usual and customary' terms that mask idiosyncratic differences (Bardhan, 1984). In late twentieth-century Illinois agriculture, for example, over fifty percent of the contracts specified fixed shares between tenant and landlord, and of these over ninety percent specified the shares  $1/2-1/2$ ,  $2/5-3/5$ , or  $1/3-2/3$  for landlord and tenant respectively (Young and Burke, 2001).

The logic of 'usual and customary' contractual terms is that they create a focal solution in a situation that has many possible solutions, thereby reducing transaction costs. Such norms are not unique to agricultural contracts: for example, building contractors and architects get customary mark-ups over cost, franchisees pay standard percentages to their parent companies, real estate agents receive customary commissions on house sales, and so forth. While such norms may reduce transaction costs, however, the uniformity imposed by the norm may prevent the contracting parties from fully wringing out all of the potential gains in their particular situation. Thus, in evaluating the efficiency of a norm, one must consider both the savings in transaction costs and the costs imposed by excess uniformity.

*The evolution of norms.* If a norm merely represents one equilibrium out of many, how does society settle on a particular one starting from out-of-equilibrium conditions? We may distinguish three ways in which norms become established and change over time: top-down influences, including official edicts and role models; bottom-up influences in which local customs and practices coalesce into

norms; and lateral influences in which established norms from one type of interaction are transferred to related types of interactions. The law, for example, operates partly from the top down: statutes and judicial rulings identify norms of acceptable behavior in people's relations with others. At the same time, the boundary between acceptable and unacceptable behavior is constantly in flux due to variations in the way that individual cases are resolved by individual courts (a bottom-up effect). And precedents in one domain can be transferred by analogy to other domains (a lateral effect). An example of the latter is the extension of laws regarding persons to those involving corporations. (For more on the interplay between norms and the law see Ellickson, 1991, and Posner, 2000).

As these examples suggest, the evolution of norms is a complex process that involves the interplay of many different forces. One may nevertheless gain insight into the process by examining how small variations in behavior at the individual level can trigger major norm shifts at the societal level. Consider a symmetric, two-person coordination game  $G$  that is played by pairs of agents drawn from a large population. Assume for simplicity that the total number of agents is even, and that in each period everyone is paired with someone else through a random matching process. Each matched pair chooses actions simultaneously and receives the corresponding payoffs in  $G$ . Assume that each agent makes a 'trembled best response' to the distribution of choices in the previous period. Specifically, suppose that for some  $0 < \varepsilon, \lambda < 1$ , each agent chooses a best response with probability  $\lambda(1 - \varepsilon)$ , trembles with probability  $\lambda\varepsilon$  (in which case he chooses an action uniformly at random), and chooses the same action as before with probability  $1 - \lambda$ , due to inertia (Kandori, Mailath, and Rob, 1993; Young, 1993a).

This type of learning process has rather striking implications for the social norms that are most likely to emerge and remain in place for long periods of time. To illustrate, consider a competition between alternative forms of money. Suppose there are  $m$  different commodities, indexed  $1 \leq k \leq m$ . Assume that in a given pairwise interaction, each player's payoff is  $a_k$  if both adopt the  $k^{\text{th}}$  form of money, whereas their payoffs are zero if they adopt different forms. It can be shown that this trembled best-response process selects the efficient equilibrium with high probability, that is, when  $\varepsilon$  is small the probability is high that, in the long run, almost everyone in the population will be using the form of money with the highest payoff (Kandori and Rob, 1995).

Unfortunately, efficiency may fail when the game does not have the structure of a pure coordination game. Suppose, for example, that a potential form of money generates payoffs in two ways: as a medium of exchange and as a form of adornment. In each pairwise interaction, let  $a_k > 0$  be the payoff in each period from using  $k$  as a medium of exchange (assuming the other player also uses  $k$ ), and let  $b_k > 0$  be the payoff from using it instead as jewelry. When there are just two commodities one obtains the symmetric payoff matrix

$$\begin{array}{cc} a_1 + b_1, a_1 + b_1 & b_1, b_2 \\ b_2, b_1 & a_2 + b_2, a_2 + b_2 \end{array}$$

Assume that this is a coordination game:  $a_1 + b_1 > b_2$  and  $a_2 + b_2 > b_1$ . Commodity  $k$  is *efficient* if it maximizes  $a_k + b_k$ . However, the evolutionary process defined above selects the *risk-dominant* commodity, that is, the commodity  $k$  that maximizes  $a_k + 2b_k$  (Young, 1998b). The latter criterion gives twice as much weight to the payoffs from adornment (which do not require coordination) as to the payoffs that arise from using the same medium of



exchange (which do require coordination). Moreover, this conclusion holds under a wide range of assumptions about the nature of the trembled best response process (Blume, 2003).

*Evolution and fairness norms.* The evolutionary framework outlined above has implications not only for the efficiency of norms but for their distributive properties as well. Consider a situation in which a principal and an agent must agree on the form of contract that will govern their relationship. Different types of contracts will have different distributive implications, some favoring the principal, others favoring the agent. Let us restrict attention to just those contracts that leave both parties better off than they would be under their outside options. To illustrate, suppose that just three contract forms are available: A, B, C. Assume a one-shot, take-it-or-leave-it bargaining process in which a given contract is adopted if and only if both parties simultaneously agree to it; otherwise they fall back on their outside options. Without loss of generality one can assume that the outside options have zero utility for both players. Consider the following example:

|            |   | Agents |      |      |
|------------|---|--------|------|------|
|            |   | A      | B    | C    |
| Principals | A | 5, 1   | 0, 0 | 0, 0 |
|            | B | 0, 0   | 3, 3 | 0, 0 |
|            | C | 0, 0   | 0, 0 | 1, 5 |

Contract A favors the principal, C favors the agent, and B is a compromise between A and C. (Of course in reality there may be many more contract forms, but this does not change the analysis in any fundamental way.) Consider an evolutionary process like the one for money conventions, but with two distinct populations – one of principals, the other of agents – that are randomly matched

in each period. Assume, as in the previous model, that they play trembled best responses with inertia, where 'best response' is defined relative to the frequency of play of the opposite population in the previous period.

A *contractual norm* is a situation in which the same contract is agreed to by everyone. In this example all three norms are efficient: none of them Pareto dominates another. It can be shown, however, that the evolutionary process favors exactly one of them, namely the compromise contract B. More generally, in evolutionary processes of this sort there tends to be a selection bias toward outcomes that represent a compromise for the two parties, and against extreme outcomes that lie near the boundary of the payoff-possibility set (Young, 1998b). This suggests that *norms of fairness* may result from evolutionary forces, an idea that is explored by Binmore (1994, 2005) and Young (1998a).

*General implications.* Although evolutionary accounts of norm formation vary in their details, they have several qualitative implications that hold quite generally. One is that different societies often employ different norms for solving the same type of coordination problem. This follows from the fact that norms represent alternative equilibria that can become established through different sequences of chance events. This is known as the *local conformity/global diversity* effect (Young, 1998a). It has been documented in a variety of settings, including agricultural contracting (Young and Burke, 2001), and the manner in which subjects divide payoffs in experimental situations (Henrich et al., 2004).

A second implication is that, due to stochastic perturbations, norms occasionally shift, and these shifts tend to be quite rapid compared to the long periods of stasis when a given norm is in place. This is the *tipping* or *punctuated equilibrium* effect (Young, 1998a).

A third implication is that some norms are inherently more stable or durable than others: once established they tend to remain in place for long periods of time even when buffeted by stochastic shocks. These *stochastically stable norms* depend on the payoff structure of the underlying game, and also on the nature of the stochastic perturbations (Foster and Young, 1990; Young, 1993a; Kandori, Mailath and Rob, 1993; Samuelson, 1997). Irrespective of these details, the important point is that some norms are remarkably resilient under changing circumstances. Due to their longevity, such norms may come to be seen as right and necessary, though in fact they are the product of chance and contingency, and are sustained simply because they coordinate people's expectations about how to interact with one another.

## References

Bardhan, Pranab. 1984. Land, Labor, and Rural Poverty: Essays in Development Economics. New York: Columbia University Press.

Bicchieri, Cristina. 2006. The Grammar of Society: The Nature and Dynamics of Social Norms. New York: Cambridge University Press.

Binmore, Ken. 1994. Game Theory and the Social Contract, Vol. I. Playing Fair. Cambridge MA: MIT Press.

Binmore, Ken. 2005. Natural Justice. Oxford: Oxford University Press.

Blume, Lawrence E. 2003. How noise matters. Games and Economic Behavior 44, 251-271.

Boyd, Robert and Peter J. Richerson. 2002. Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology* 215, 287-296.

Cheung, Stephen N.S. *The Theory of Share Tenancy*. Chicago: University of Chicago Press.

Coleman, James S. 1987. Norms as social capital. In *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*, G. Radnitzky and P. Bernholz, eds. New York: Paragon House.

Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge MA: Harvard University Press.

Ellickson, Robert C. 1991. *Order Without Law: How Neighbors Settle Disputes*. Cambridge MA: Harvard University Press.

Elster, Jon. 1989. Social norms and economic theory. *Journal of Economic Perspectives*, 3, no. 4, 99-117.

Elster, Jon. (1999). *Alchemies of the Mind*. Cambridge UK: Cambridge University Press.

Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13, 1-25.

Foster, Dean P., and H. Peyton Young. 1990. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38, 219-232.

Harsanyi, John and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge MA: MIT Press.

Hechter, Michael and Karl-Dieter Opp, eds. 2001. *Social Norms*. New York: Russell Sage Foundation.

Henrich, Joe, Robert Boyd, Samuel Bowles, Ernst Fehr, and Herbert Gintis, eds. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence in 15 Small-Scale Societies*. Oxford: Oxford University Press.

Hume, [1739]. 1978. *A Treatise of Human Nature*. Oxford: Oxford University Press.

Kandori, Michihiro, George Mailath, and Rafael Rob. 1993. Learning, mutation, and long-run equilibria in games. *Econometrica* 61, 29-56.

Kandori, Michihiro, and Rafael Rob. 1995. Evolution of equilibria in the long run: a general theory and applications. *Journal of Economic Theory* 65, 29-56.

Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge MA: Harvard University Press.

Posner, Eric. 2000. *Law and Social Norms*. Cambridge MA: Harvard University Press.

Robson, Arthur and Fernando Vega-Redondo. 1996. Efficient equilibrium selection in evolutionary games with random matching. *Journal of Economic Theory* 70, 65-92.

Roth, Alvin E. 1985. Toward a focal point theory of bargaining. In *Game-Theoretic Models of Bargaining*, Alvin E. Roth, ed. Cambridge UK and New York: Cambridge University Press.

Samuelson, Larry. 1997. *Evolutionary Games and Equilibrium Selection*. Cambridge MA: MIT Press.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge MA: Harvard University Press.

Sugden, Robert. 1986. *The Economics of Rights, Cooperation and Welfare*. Oxford: Basil Blackwell.

Ullmann-Margalit, Edna. 1977. *The Emergence of Norms*. Oxford: Oxford University Press.

Wärneryd, Karl. 1994. Transaction cost, institutions, and evolution. *Journal of Economic Behavior and Organization* 25, 219-239.

Young, H. Peyton. 1993a. The evolution of conventions. *Econometrica* 61, 57-84.

Young, H. Peyton. 1993b. An evolutionary model of bargaining. *Journal of Economic Theory* 59, 145-168.

Young, H. Peyton. 1998a. *Individual Strategy and Social Structure*. Princeton NJ: Princeton University Press.

Young, H. Peyton. 1998b. Conventional contracts. *Review of Economic Studies*, 65, 773-792.

Young, H. Peyton and Mary A. Burke. 2001. Competition and custom in economic contracts: a case study of Illinois agriculture. *American Economic Review* 91, 559-573.