The Power of Norms

H. Peyton Young
Department of Economics, Johns Hopkins University, Baltimore, MD 21218, U.S.A.

ABSTRACT

Norms are basic building blocks of social and economic organization. This chapter proposes a framework for studying the evolution of norms based on the cumulative effect of many decentralized interactions by individuals. Predictions of the theory are illustrated for contractual norms between landlords and tenants in contemporary United States agriculture.

DEFINITION AND ROLE OF NORMS

A norm is a rule of behavior that is self-reinforcing: everyone wants to play their part given the expectation that others will play theirs. This definition encompasses simple rules that solve coordination problems, such as driving on a given side of the road, as well as more complex rules that involve sanctioning those who deviate from a first-order rule. (I express outrage if someone cuts in front of someone else in line; I refuse to associate with people who fail to tip in restaurants.) Norms structure our relations with others so completely that we often fail to recognize just how pervasive they are. The clothes I wear, the food I eat, the manner and time of day at which I eat it, the ways I address people, the obligations that I feel toward members of my family, the duties that I perform at work, and the amount that I earn are all determined to a significant degree by prevailing norms of behavior in the society.

Although sociologists and anthropologists have long understood the central role played by norms, economists have traditionally viewed them as being peripheral to economic decisions (notable exceptions are Schelling [1960], Akerlof [1997], and Sugden [1986]). The fact is, however, that norms influence terms of employment, the amount that people save and consume, attitudes toward debt, decisions about when to retire, and a host of other economic variables. Even property rights are governed to a considerable extent by social expectations about who is entitled to what (Hume 1739; Sugden 1986).

The social function of norms is to resolve problems of collective action and coordination (Ullman-Margalit 1977). Indeed, norms can be viewed as
equilibria of appropriately defined games. These games often have a multiplicity of equilibria, as evidenced by the fact that solutions to a given coordination problem often differ substantially from one society to another, and also within a given society over time. However, although social norms are related to the notion of equilibrium, they are not the same thing.

To illustrate the distinction, consider two individuals who can divide a dollar provided they agree on how to divide it. Each makes a demand, and if the demands sum to at most one dollar their demands are met; otherwise they get nothing. If one demands 43 cents and the other 57 cents, the demands are in equilibrium: no one can gain by unilaterally changing his demand. Yet while this is an equilibrium, it is not a norm. The equilibrium is idiosyncratic to these particular individuals. Fifty-fifty division, by contrast, is a norm because it is a usual and customary solution in games of this kind. More generally, a norm is an equilibrium behavior in a game played repeatedly by many different individuals in society where the behavior is known to be customary. Note the importance of knowledge: behavior must not only be customary, it must be known to be customary or else behaviors are not in fact self-enforcing (Lewis 1969). Note also that this definition encompasses behaviors that require no sanctions by third parties to constitute an equilibrium. The latter are sometimes referred to as conventions rather than norms, though I shall not dwell on this distinction here.

THE EVOLUTION OF NORMS

How then do norms become established and what causes them to change? In Young (1993, 1998a), I suggest a general framework for investigating this question. The fundamental idea is that norms coalesce from the decentralized, uncoordinated choices of many interacting individuals. Roughly speaking, individuals are the particles of the system, and norms are the organizational forms that bind them together. Unlike particles, however, individuals make intentional choices based on perceived constraints and opportunities. We therefore need to explain how such individual choices can lead to the emergence of society-wide norms that promote coordinated behavior.

We model this situation as follows. Consider a given type of social interaction that regularly confronts different members of a society. For simplicity we shall think of this interaction as a two-person coordination game having multiple equilibria. At any given time \( t \) various solutions will have been tried, and there will be a distribution of behaviors in society. Roughly speaking this distribution defines the state of the system at time \( t \). In the next period, a given pair (or pairs) of individuals interact with some probability. The information they have about the previous behavior of others leads them to expect some behavior (or distribution of behaviors) among their present opponents. These expectations lead them to make certain choices in the current period, which then become precedents for individuals in later periods. These choices may be purely rational,
but more likely they involve elements of conformity, experimentation, inertia, and other forms of nonrational behavior. These other elements can be incorporated into a stochastic choice model that involves mostly rational behavior with some idiosyncratic elements.

The combination of these assumptions leads to a stochastic dynamical system that is built around the following feedback loop:

\[ \text{Precedents} \rightarrow \text{Expectations} \rightarrow \text{Actions} \]

It can be shown that, under fairly general conditions, such a system converges to a situation that is close to being a norm in the sense that almost everyone is playing their part in a particular equilibrium, though idiosyncratic variations will also typically be present. In addition, however, the system occasionally shifts abruptly from one norm to another. These shifts can be induced either by system-wide shocks (e.g., changes in the underlying payoff structure, perhaps due to technological change) or by expectational “drift” that arises from the accumulation of idiosyncratic choices by a few individuals. Either one of these factors can push the system to a critical tipping point, beyond which expectations change and society converges to a new norm.

A central prediction of the theory is that the drift component tends, over time, to favor some norms over others (Young 1993, 1998b). In other words some norms are more durable than others in the sense that, once established, they tend to stay in place for longer periods of time. Surprisingly, the most durable norms are not necessarily efficient (i.e., Pareto optimal); in some situations, risk dominance is a better criterion of durability (Young 1993; Kandori et al. 1993).

Even when the evolutionary process does operate in favor of efficient norms, not all such norms are equally durable. In fact, the distributional properties of a norm also affect its durability to a significant extent. Specifically, in a pure coordination game between two players, the most durable norms are those that maximize the position of the worst-off party relative to the best outcome they could get. (That is, the outcome \( x \) is such that \( \min (u_1(x)/u_1^+, u_2(x)/u_2^+) \) is maximized over all \( x \), where \( u_1^+ \) and \( u_2^+ \) are the maximum utilities 1 and 2 could get under some choice of \( x \).) Thus, even when individuals always make choices that maximize their own utility, with no regard for the utility of others and no preference for fairness, the net effect of the evolutionary process is to favor norms that look as if people did have such preferences (Young 1998a).

**EMPIRICAL STUDIES OF NORM FORMATION:**

**AGRICULTURAL CONTRACTS**

In this chapter I discuss how the theory can be brought to bear on situations where norms, and norm shifts, can be verified empirically. The identification of such situations poses a number of challenges. One difficulty is that norms sometimes become codified into laws, so that even though a norm might originally
have emerged from the bottom up, it later becomes enforced from the top down, so that compliance is no longer voluntary. An example is the emergence of local ordinances in the southern U.S. requiring blacks to give up their seats to whites on public transportation. (Interestingly, even though this norm was codified, it was eventually overturned by the spontaneous [illegal] actions of a few individuals, notably Rosa Parks. This illustrates the general point that the viability of a norm ultimately rests on the shared expectation that people will conform to it, not on its legal status.)

Another difficulty in identifying empirical cases is that what appears to be a norm might be some form of spurious correlation. For example, some groups in the population smoke more than others (e.g., a much smaller proportion of black men are smokers than white women). Is this the result of different social norms operating in different groups, or is it the result of subtle differences in bodily responses to nicotine?

A study of norms by Young and Burke (2001; see also Burke and Young 2003) avoids some of these difficulties. Specifically, we studied whether social norms shape the terms of contracts between tenants and landlords in contemporary U.S. agriculture. Although this may seem like a somewhat unusual area in which to study the operation of norms, it has a number of important attributes. First, there is extensive data on contracts between tenants and landlords that extends over many years, gathered by agricultural economists and research branches of the Department of Agriculture. Second, modern U.S. agriculture is a highly competitive and sophisticated business in which both tenants and landlords have a great deal of scope for making choices. (It is quite unlike the situation where southern sharecroppers or European peasants were trapped on the land through debt and lack of alternative employment.)

Third, there is little reason to think that norms are the spurious by-product of associational preferences: farmers do not move to an area because they like the other farmers there, but in most cases because they were born there. (This applies to landowners as well as tenants.) In other words, interactions are determined by geography and happenstance of birth; they are not contaminated by the endogenous sorting of people into like-minded groups — a difficulty that has bedevilled other norm studies (see Manski 1993).

Fourth, there is extensive data on the underlying quality of factor inputs, so that the standard refuge of the skeptic — the existence of common unobservables — can be eliminated by appropriate statistical tests. Finally, the choice of contract is a purely voluntary act between principal and agent that is not constrained by law. Hence, standard competitive market forces should, in theory, determine the outcome.

PREDICTIONS OF STANDARD THEORY

Let us quickly review what standard competitive theory would predict in this situation. Each year a given landowner (the principal) and a prospective tenant
farmer (the agent) negotiate a contract for the coming year. The main factors governing the outcome of the bargain are: quality of the land, quality of the labor, the cost of monitoring the tenant's performance, the parties' attitude toward risk, and their opportunity costs, that is, the value of their next-best alternatives. (For the tenant this includes the option of alternative employment, say in a nearby factory.) This problem has been extensively studied from a theoretical point of view (Cheung 1969; Stiglitz 1974; Hayami and Otsuka 1993). Basically the monitoring cost and attitudes toward risk determine the form of the contract, that is, whether the parties prefer a share contract, a cash lease, a wage contract, or some hybrid form. In our study we restrict attention to the subset of farms that adopted the share format, thus effectively sidestepping the impact of monitoring costs and risk attitudes, which we cannot observe in any event. (Theory says that those agents who choose share contracts tend to be those who are more risk averse and for whom monitoring costs are low.)

Among those agents who opt for share contracts, the empirical question is what division of the crop they negotiate. In principle they could agree to different divisions of each crop (corn, wheat, and soybeans) as well as to different divisions of each input (e.g., fertilizer, seed, and equipment). In practice, however, almost all share contracts are expressed in terms of a single share for all the outputs and a single share for all inputs except equipment, which is the sole responsibility of the tenant.

Assume for the moment that all labor is similar in quality and that the reservation wage in a given area is fixed, say, by the going wage in factory employment. Then, in competitive equilibrium, laborers will earn the reservation wage and the residual surplus will go to the landowner as pure rent. In particular, soils of a given quality should earn the same rent across different farms, and higher-quality soils should command higher rents. This implies that the negotiated share to the tenant should be lower on higher-quality farms. Furthermore, the two parties know how to adjust the share appropriately because the quality of soils on a given farm is rated according to a scheme that gives expected productivity per acre of different crops, holding labor and other inputs constant, and both parties know the ratings, which are a matter of public record.

We turn now to the data to see if these predictions hold. These data come from a sample survey conducted by the Illinois Cooperative Agricultural Extension Service (1995). Of the 1704 responses in the 1995 survey, cropsharing contracts were the most frequent (55%) and land rent contracts the next most frequent (41%); all other contract forms (mostly livestock and pasture leases) constituted less than 4%. We restrict ourselves to an analysis of the 935 cropshare contracts.

Figure 20.1 shows the frequency distribution of shares of corn output, which is virtually the same as the frequency distribution of shares for soybeans and wheat. Note that 1/2-1/2 is by far the most common division, and virtually all contracts use either 1/2-1/2, 3/5-2/5, or 2/3-1/3. (Here and elsewhere we list the share to the tenant first.) These data are difficult to reconcile with the standard
competitive account, because they show so little variation in contract terms. (It is also quite suspicious that the shares concentrate solely on fractions with small denominators.)

The situation is further illuminated by looking at the distribution of shares in different parts of the state. Illinois exhibits considerable variation in its soil characteristics. In the north, the land is mostly flat and the soils are on average highly productive, whereas in the south the land tends to be hillier, the topsoil is not as thick, and on average it is less productive. (This north-south division corresponds roughly to the southern boundary of the last major glaciation.) When we compare contract frequencies in the northern and southern parts of the state, substantial differences appear (see Figure 20.2). In the north, contract terms are almost exclusively 1/2–1/2, whereas in the south the predominant contracts are 2/3–1/3 and 3/5–2/5.

These differences make sense from an economic point of view. Because the land in the south is, on average, inherently less productive than the land in the north, the share for the tenant must be higher in the south if net returns to the tenants in the two regions are to be comparable. Viewed in this way, the data seem to vindicate standard competitive theory.

The data only vindicate standard theory in the crudest sense, however. In the first place, there should not be three shares but a spectrum of shares reflecting the soil qualities of the farms in question. Are we to believe, for example, that all of the farms in the north have the same soil quality, which justifies a share of 1/2, while all the farms in the south have one of two soil qualities, namely, those that justify 2/3 or 3/5 to the tenant?

This hypothesis seems absurd on the face of it. Moreover, it is completely contradicted by the actual distribution of soil qualities. In both the north and the south there is a wide range of qualities, and the highest produces over twice as much per acre as the lowest (holding labor and other inputs constant). In fact,
Figure 20.2  Distribution of share contracts by region. Illinois Cooperative Agricultural Extension Service (1995).

this is true within virtually each individual county in the north and the south. According to standard theory the negotiated shares should reflect these differences. Instead, the shares in a given county cluster around a small number of values independently of the soil quality. Figure 20.3 illustrates this effect for two representative counties, one in the north and one in the south.

**PREDICTIONS OF A NORMS-BASED THEORY**

To understand a model of norm formation that can account for these phenomena, three facts need to be explained. First, we have seen that there is a tendency for contract terms to cluster on a few simple fractions that have a priori focal power. This is the quantum effect. Second, there is much less heterogeneity locally than competitive theory would predict. This is the local conformity effect. Notice that the second does not follow from the first: even though the quantum effect limits the number of distinct contracts that might be observed, contract terms could still vary substantially from one farm to the next, resulting in a great deal of local diversity. This is not confirmed by the data even though local differences in fundamentals might call for it. The third fact to be explained is that the contractual norm differs between the two regions in a way that is broadly consistent with the competitive model. This is the regional diversity effect (Young 1999b).

Consider the following dynamic model. In each period, landlords propose contracts to tenants. In proposing a contract, the landlord takes into account both economic and psychological factors. First, he obviously cares about the expected returns from the contract. Second, he wants to conclude an agreement expeditiously and without a lot of haggling, which argues for keeping the terms simple. Third, he values his relationship with the tenant and his standing among his neighbors. Hence he wants to conclude an agreement that is perceived by the tenant to be fair and that adheres to general standards within the local community. In particular, if he were to offer the tenant less than the going share, he risks antagonizing someone he must work with and he may be seen as greedy or exploitative by his neighbors. If he offers the tenant more than the going share, the tenant may be happy but the neighboring landlords will not be; moreover, this goes against his interest in maximizing returns. The tradeoffs between these
Figure 20.3  Distribution of shares in (a) Tazewell Co. (north) and (b) Effingham Co. (south).

Considerations are difficult to estimate and will no doubt vary substantially among regions and among particular principal-agent pairs. Here we shall posit a model that could, in principle, be estimated empirically from event histories that specify the temporal sequence of contract adoptions by location. (The Illinois data do not allow us to make this estimation because they do not contain a sufficiently large panel of farms, and the locations are aggregated so as not to reveal the identity of particular respondents.)

Assume that each landlord chooses from a menu of simple contracts that apply a single, easy-to-calculate fraction to all inputs and outputs. The probability of choosing a given contract in the menu is given by a logit function in which the log probability of a contract increases linearly in its expected returns and the degree to which it conforms with contracts used on nearby farms. The relative weighting between these two factors is a parameter that can be estimated. This type of setup is standard in discrete choice analysis (McFadden 1974; Brock and Durlauf 2001) and is relatively easy to estimate from microlevel data (if we had such data).
The basic predictions of the model are as follows. Consider any distribution of soil qualities that is heterogeneous locally and exhibits substantial shifts in average quality between geographic regions. It can be shown that, for a wide range of parameter values, the model converges with high probability to a situation characterized by regional customs. That is, there is near-uniformity of the contractual custom within each region, but substantial differences between regions that reflect underlying differences in average quality. Boundaries between regional customs form endogenously and are quite sharp: just over the border, people do things differently. However, there may be nothing that marks the boundary line per se: it suffices that average land quality be somewhat different on opposite sides. This is just what we see in the Illinois data.

The existence of these regional norms has important implications for the economic returns realized by labor and land. To see this, suppose that the share to the tenant is uniform in a given region, irrespective of soil quality. Then, on soils of higher quality the tenant will earn higher net returns because he captures a fixed fraction of the increased yield per acre that the higher-quality soil produces. Thus, unless the landlord enforces a higher level of labor input per acre on the higher-quality land, labor will effectively capture a portion of the land rent because of the rigidity of the contract.

It can be shown, however, that labor input per acre does not increase on higher-quality land (Burke and Young 2001). (This exploits the detailed knowledge we have of the productivity of different soils when inputs remain fixed.) In other words, tenants succeed in capturing part of the rent that should accrue to land. Moreover, the amount of rent capture is quite sizable. Our estimate is that the tenant captures about one-third of the rent that ought to accrue to higher soil quality, and this is attributable mainly to the rigidities induced by contractual custom. On a farm of several hundred acres, the rent capture by the tenant may amount to several thousand dollars per year, which is a nonnegligible fraction of his income.

Of course, this argument would be undermined if we could identify hidden costs to the tenant or hidden benefits to the landlord that justify the higher payments to labor on higher-quality land. One possibility, for example, is that yields on higher-quality soils are more variable than on lower-quality soils. This would imply that risk-averse tenants must receive higher expected payments to be willing to farm high-quality soils. (As pointed out earlier, there is reason to believe that the tenants in our sample are risk averse because they opted for a share contract instead of a land rent contract.) In fact, however, the data show that variability in yield is not significantly related to soil quality.

A second possibility is that higher returns to the tenant result in lower turnover rates, thus reducing transaction costs for the landlord. Although the data do not allow us to test for this possibility directly, the rate of tenant turnover is so low that only a small portion of the labor premium could be explained by this consideration (if in fact it explains any of the premium).
A third possibility is that higher-quality tenants migrate to higher-quality land. In other words, equilibrium is achieved in the labor market through assortative matching rather than by adjusting contract terms. If this explanation is to have any force, however, then higher-quality tenants must generate an increase in yield that goes beyond the increase attributable to higher land quality alone, holding other inputs fixed. The evidence strongly suggests that this is not the case; indeed, if anything, there appears to be some slacking off in the quality or quantity of labor input as land quality rises.

Finally, we consider the possibility that the increase in tenant net income is the result of spurious correlation effects. It could be, for example, that reservation wages happen to be higher in regions with higher land quality. In this case, the return to tenant labor would rise with land quality, but the relationship would be purely spurious. We test for this and other local fixed effects and find that they do not account for the observed increase in tenant income.

The question remains as to how this kind of behavior can be sustained in a competitive market environment. Are norms really so powerful that they can distort economic returns to this extent? I conjecture that this phenomenon is quite general and that similar distortions would be uncovered in many other parts of the economy if people would only look for them. Among the prime candidates are decisions about when to retire, how much to save for retirement, how much it is prudent to borrow, how much to pay your lawyer to defend against a lawsuit, how much to pay senior faculty in comparison to junior faculty, how many stock options to award the CEO of a corporation, and so on. In all of these cases, my guess is that social norms substantially alter the decisions that would be predicted by the standard competitive model and that economists need to start taking these effects seriously.

REFERENCES