Social norms and economic welfare

H. Peyton Young*

Department of Economics, Johns Hopkins University, Baltimore, MD 21218, USA

Abstract

A norm is an established and self-reinforcing pattern of behavior: everyone wants to play their part given the expectation that everyone else will continue to play theirs. It is, in short, an equilibrium of a game. This paper surveys some of the ways in which norms structure economic life: in the definition (and division) of property, in the terms of contracts, and in the assignment of social roles. We argue that norms can evolve from the cumulative effect of many decentralized interactions by individuals who are trying to solve a coordination problem. The theory suggests circumstances under which evolutionary forces favor norms that are efficient and more or less egalitarian in their distributive implications. © 1998 Elsevier Science B.V. All rights reserved.

JEL classification: C70; D60; D63

Keywords: Convention; Coordination; Transaction cost; Norms; Equity

1. Property rights

Social norms structure our relations with others so completely that we often fail to recognize just how pervasive they are. Even property rights, which we ordinarily think of as being fixed by the law, are in fact governed to some extent by social expectations about who is entitled to what. Zajac (1985) illustrates this

---

* E-mail: pyoung@jhu.edu.

1 This work was supported by the National Science Foundation. For an elaboration of these ideas see my forthcoming book Individual Strategy and Social Structure, Princeton University Press, 1998.
point with an incident that occurred when he and his wife were on a tour of Cuba. On arrival the members of the tour group boarded two buses. The guide on the Zajas' bus was informative and fluent in English, whereas the guide on the other bus was tiresome and spoke English with difficulty. Nobody attempted to change buses on the second day. On the third day, however, the Zajas were among the last to board the buses, when they discovered that a couple from the other bus had taken their seats. Naturally, they expressed outrage, and they were supported by others who voiced objections as well (including the guide). Rather than face the wrath of the other tour members, the interlopers eventually returned to 'their' bus. Evidently, the Zajas had established a right of possession for the rest of the trip that others recognized as being valid. While their possession was not a matter of legal entitlement, it was supported by a social convention that people were willing to defend through social disapproval. And because it was supported by convention, it amounted to a valuable property right that for all practical purposes was just as good as a legal entitlement.

This seemingly trivial example illustrates a general principle that applies to all manner of property – the right of first possession. Sugden (1989) cites an application of the principle to the gathering of driftwood by Yorkshire villagers. The tradition was that, after a storm, whoever came first onto a stretch of shore after high tide could gather the driftwood into piles without interference from later arrivals. He could also secure his claim by placing two stones atop the pile. But if the wood was not removed after two more high tides the claim lapsed. This custom did not establish a legal right of ownership, but it was widely recognized within the village, and supported by the knowledge that those who attempt to flout the custom would be subjected to strong social disapproval. It therefore amounted to a right.

These two examples illustrate the key role of convention in defining and maintaining property rights, an idea first articulated by Hume (1739). Hume recognized, moreover, that such conventions can only be sustained if they constitute an equilibrium in a suitably defined game. 'I observe that it will be for my interest to leave another in the possession of his goods, provided he will act in the same manner with regard to me. He is sensible of a like interest in the regulation of his conduct. When this common sense of interest is mutually expressed, and is known to both, it produces a suitable resolution and behavior' (Book 3, part 2, Section 2). Hume also recognized that convention is a general organizing principle of society that goes well beyond the definition of property rights. Language conveys information through the conventional meanings of words. Money has value in exchange because other people value it. Contracts are enforceable to the extent that their terms are interpreted conventionally by the courts. And so forth. It is, in fact, hard to think of an economic or social interaction that is not influenced to some degree by convention.
2. Conventions and equilibrium

But is convention the same thing as an equilibrium in a game? 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 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 this is not a convention; it is simply an idiosyncratic equilibrium selected by two particular individuals. Fifty-fifty division, by contrast, is a convention because it is a usual and customary equilibrium in games of this kind. Convention is a social construct. To capture the social dimension of convention, we could say that a convention is equilibrium behavior in a game played repeatedly by many different individuals in society, where the behaviors are widely known to be customary. Note the importance of knowledge: the behaviors must not only be customary, they must be known to be customary, or else the behaviors are not in fact self-enforcing. Lewis (1969) was the first to emphasize the importance of knowledge in sustaining equilibrium behavior; more recently, Aumann and Brandenburger (1995) have formulated precise epistemic conditions under which Nash equilibrium will be played.

3. Conventions and transaction cost

What, though, is the relationship between social convention and economic welfare? At one level the answer is simple enough: conventions reduce transaction costs by coordinating expectations and reducing uncertainty. A canonical example is rules of the road. Before the advent of traffic laws, these were largely a matter of local custom; indeed, in some European countries – including Italy and Austria – the left-hand convention prevailed in some parts of the country, and the right-hand convention in others as recently as the 1920s (Lay, 1992).

We can model this situation as a coordination game with the following payoffs:

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>0, 0</td>
<td>-1, -1</td>
</tr>
<tr>
<td>Right</td>
<td>-1, -1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

The absence of a convention means that people's choices are unpredictable, which leads to collisions. Suppose, for example, that the frequency of choosing left and right in the population is 30% Left, 70% Right. If two individuals are

---

2 See, for example, Wärneryd (1994).
approaching each other and cannot signal their intentions, the probability of a collision is 0.42. The 'transaction cost' of this state of affairs, as compared to having a convention, is $-0.42$ per interaction. Note, moreover, that the highest transaction cost occurs at 50\% Left, 50\% Right, which is also an equilibrium of the game. As game theory teaches us, not all equilibria are created equal. In particular, only pure-strategy equilibria in coordination games can be said to solve the transaction cost problem in a reasonable way.

4. Inefficient and discriminatory conventions

In the driving game both pure equilibria have the same welfare implications, so the existence of a pure-strategy convention is all that matters from the standpoint of economic welfare, not the convention itself. In general, however, pure-strategy equilibria may have different welfare implications, in which case the convention does matter. Consider, e.g., the evolution of technological standards, such as the competition between Macintosh and DOS computer operating systems, or, in an earlier era, the competition between QWERTY and DVORAK keyboard layouts on typewriters (David, 1985; Katz and Shapiro, 1985; Arthur, 1989). Abstracting away from the details, we could represent this as a coordination game with the following hypothetical payoffs:

<table>
<thead>
<tr>
<th></th>
<th>MAC</th>
<th>DOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC</td>
<td>9, 9</td>
<td>4, 4</td>
</tr>
<tr>
<td>DOS</td>
<td>4, 4</td>
<td>8, 8</td>
</tr>
</tbody>
</table>

Both standards solve the transaction cost problem, but under the assumed payoffs MAC is better from the point of view of welfare than DOS. Unfortunately, once DOS becomes entrenched, society may find it very difficult to switch to MAC. In other words, if conventions are sustained through the decentralized, uncoordinated choices of many individuals, there seems to be no reason why they would be efficient. Actually, this view is overly pessimistic: in some classes of games (including the one above) there is evolutionary pressure toward efficient conventions, as we shall show in Section 7 below.

Another point that needs to be stressed is that conventions often have important implications for the distribution of welfare. Indeed, in many societies the distribution of rights and opportunities is conditioned on such factors as race, gender, age, or social status, and these discriminatory norms persist largely because people have come to expect them. Consider again the example of bus seats. Being the first to occupy a seat does not always convey an unambiguous right of possession. If an elderly man or a pregnant woman gets on the bus and
no other seat is available, social convention dictates that the occupant yield in
their favor. And it was not long ago that, in the southern United States, blacks
were expected to yield their seats to whites (a norm that was often backed up by
legal ordinances). Thus, while social norms may reduce transactions costs, there
is no guarantee that they will do so efficiently, or that the welfare gains will be
evenly distributed.

5. The evolution of conventions

So far I have examined the relationship between social norms and economic
welfare, but have said nothing about how such norms come into being. Here
I shall outline a simple model of how norms can evolve from the cumulative
impact of many individual decisions, and then show that the theory has rather
striking implications for their efficiency and distributive properties. To make
the discussion concrete, consider a situation in which people can enter into pairwise
contractual relationships with one another (Young, 1998a). There are two
classes of agents: tenants and landlords, employers and employees, bankers and
borrowers, men and women, etc. At the beginning of each time period, every
agent is matched at random with someone from the opposite population. We
shall assume that all matchings are equally likely, although this is not essential
for the following results. Each matched pair of agents must agree on a contract,
i.e., on the terms that will govern their relationship over the current period.
There are $m$ possible contracts. The terms of each contract are enforceable, and
there is no renegotiation over the period.

A negotiation between a pair of agents will be modeled as a take-it-or-leave-it
game. Each side names a contract. They enter into a relationship if they name
the same contract; otherwise they do not enter into a relationship and are
unattached for the period. We make no assumption that the agents select only
optimal contracts (this will be derived as a consequence of the model). We do
assume, however, that all contracts are desirable (i.e., individually rational) in the
sense that the expected utility for each side is greater than the utility of being
unattached.

Altogether there are $m + 1$ states that an individual can occupy in any given
period – be a party to one of the $m$ contracts, or be unattached. Assume for
simplicity that all row players have the same utility for a given contract, and
similarly for the column players. Let $(a_i, b_i)$ be the expected utility to the row and
column player, respectively, from being in contract $i$, and let $(a_0, b_0)$ the expected
utility from being unattached. There is no loss of generality in assigning a utility
of zero to the unattached state for each player, and we shall assume this
henceforth. We therefore obtain a pure coordination game in which the diagonal
payoffs are strictly positive and the off-diagonal payoffs are zero. Call this type
of coordination game a contracting game.
The terms that a player demands are governed by his or her expectations about the terms that the other side will demand, and these expectations are derived from observations about the terms that agents in that population actually did demand in previous periods. The evolutionary process is therefore driven by a feedback loop in which precedent shapes present expectations, which determine present behavior, which become future precedents. We further assume that the process is buffeted by small stochastic shocks that represent idiosyncratic aspects of individual behavior and other random events that are not explicitly modeled.

Specifically, the state at the end of period $t$ is a pair of integer-valued vectors $z^t = (x^t, y^t)$ where $x^t_i$ is the number of row players who proposed contract $i$ in period $t$, and $y^t_i$ is the number of column players who proposed contract $i$ in period $t$. We shall assume that each population contains $N$ individuals; thus in every period $\sum x^t_i = \sum y^t_i = N$.

At the beginning of period $t + 1$, the individuals from the two populations are matched at random. Each member of a matched pair proposes a contract, i.e., the terms or conditions of their prospective relationship. Each person decides what terms to demand by consulting the frequency distribution of demands made by members of the other population in the previous period. He uses this distribution to predict the likelihood that his current partner will make various demands now. These assumptions obviously sacrifice some degree of realism; nevertheless they have considerable justification in a large-population setting. The myopic best reply of a row player, given the state $z^t = (x^t, y^t)$, is to demand a contract $i$ that maximizes $a_i y^t_i$, $1 \leq i \leq m$. Similarly, the myopic best reply of a column player is to maximize $b_i x^t_i$, $1 \leq i \leq m$. If two or more contracts constitute a best reply for some agent, we can suppose that he chooses among them according to a fixed probability distribution that has full support.

Individuals can deviate from best reply in two ways: through inertia and random error. Inertia means that the player adopts the same choice that he (or his predecessor) did in the previous period. Error encompasses random shocks to utility, experimentation, and various idiosyncratic aspects of behavior that we do not model explicitly. Let $\nu \in (0, 1)$ be the inertia probability and let $\varepsilon \in [0, 1)$ be the error probability. In each period, a player sticks to his previous choice with probability $\nu$, chooses a best reply with probability $(1 - \nu)(1 - \varepsilon)$, and chooses a contract at random with probability $(1 - \nu)\varepsilon$. These behavioral rules define a Markov chain $P^{N, \nu, \varepsilon}$ on the finite state space $Z$ consisting of all pairs $z = (x, y) \in \mathbb{R}_+^m \times \mathbb{R}_+^m$, where $\sum x_i = \sum y_i = N$, and $x_i, y_i$ are nonnegative integers.

When $\varepsilon$ is positive, there is a positive probability of moving from any state $z$ to any other state $z'$ in one period, because everyone could make an error (i.e., choose a contract at random). Hence, the process is irreducible. It is a standard result that a finite, irreducible Markov chain has a unique stationary distribution $\mu^{N, \nu, \varepsilon}(z)$. Over almost all realizations of the process, $\mu^{N, \nu, \varepsilon}(z)$ is the relative frequency with which the process visits state $z$ (Kemeny and Snell, 1960).
A convention is a state in which all players demand the same contract, which we shall call a conventional contract. Such a state has the form \( z^i = (N \bar{e}^i, N \bar{e}^i) \), where \( \bar{e}^i \) is the unit vector with 1 in position \( i \) and 0 elsewhere. An absorbing state is a state \( z \) such that, if the process is in state \( z \) in some period, it remains in that state forever. When \( \varepsilon = 0 \), every convention is an absorbing state; moreover, they are the only absorbing states (assuming the inertia probability is strictly less than one).

When the disturbance term \( \varepsilon \) is small but positive, it can be shown that the stationary distribution \( \mu^{N,\varepsilon}(\cdot) \) puts almost all of the probability on one or more conventions \( z^i \). To be precise, there is a unique, nonempty set of conventions \( Z^0 \subseteq \{ z^1, z^2, \ldots, z^m \} \) such that

\[
\lim_{\varepsilon \to 0} \mu^{N,\varepsilon}(z) > 0 \quad \text{if and only if } z \in Z^0. \tag{1}
\]

Equivalently, \( Z^0 \) is the minimal set of states such that, given any \( p < 1 \),

\[
\sum_{z \in Z^0} \mu^{N,\varepsilon}(z) \geq p \quad \text{for all sufficiently small } \varepsilon > 0. \tag{2}
\]

The states satisfying Eq. (1) are said to be stochastically stable (Foster and Young, 1990). The meaning of stochastic stability is that the process is very likely to be in such a state when the perturbation probability \( \varepsilon \) is small but nonvanishing. In other words, when conventions are occasionally but persistently challenged by innovations, the stable one(s) are more likely to be observed than the others over the long run.

Of course, since the state space is determined by the population size \( N \), it is conceivable that stochastic stability also depends on \( N \). Fortunately, this is not an issue when \( N \) is large: for generic contracting games there exists a unique contract \( i \) such that the corresponding norm \( z^i \) is stochastically stable for all sufficiently large \( N \). Even for nongeneric contracting games, there exists a unique set of contracts that are stochastically stable for arbitrarily large values of \( N \). In general, we say that contract \( i \) is stable if for every \( \nu \in (0, 1) \) there are arbitrarily large values of \( N \) such that \( \lim_{\varepsilon \to 0} \mu^{N,\varepsilon}(z^i) > 0. \)

6. An example: Marriage contracts

We illustrate these ideas with an example involving marriage contracts. Let the two populations consist of men and women. In each period, every man is tentatively matched with one woman. A matched pair gets married if (and only if) they can agree on control of the marital property. For simplicity assume that there are only three possible contracts: the husband has full control over the property, the wife has full control over the property, or they control it jointly. If
they name the same contract they get married, if they name different ones they break up. Let the payoffs be as follows:

<table>
<thead>
<tr>
<th>Marriage contracting game</th>
<th>men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>woman</td>
</tr>
<tr>
<td>woman controls</td>
<td>5, 1</td>
</tr>
<tr>
<td>women controls</td>
<td>0, 0</td>
</tr>
<tr>
<td>woman controls</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

Beginning from an arbitrary initial state, the process eventually settles down into a pattern with the following characteristics. First, at any given time, it is likely that society will be close to one of the three possible conventions, though there will also be some variation around the convention. Second, the current convention occasionally shifts due to an accumulation of stochastic shocks that causes people's expectations to change. Third, one of the conventions – joint control – is considerably more likely than the others (it is the stochastically stable convention).

7. A welfare theorem on conventions

We now state a general theorem about the welfare properties of stochastically stable contracts. Consider a set of \( m \) contracts with expected payoffs \( (a_i, b_i) > (0, 0), 1 \leq i \leq m \). A contract is efficient if there exists no other contract that yields strictly higher payoffs to both parties. Let \( a^+ = \max\{a_i; 1 \leq i \leq m\} \) denote the maximum feasible payoff to the row players under some contract; similarly let \( b^+ = \max\{b_i; 1 \leq i \leq m\} \) denote the maximum feasible payoff to the column players. Define the welfare index of contract \( i \) to be the smaller of \( a_i/a^+ \) and \( b_i/b^+ \), i.e.

\[
w_i = a_i/a^+ \wedge b_i/b^+
\]

and let

\[
w^+ = \max_i w_i.
\]

A contract with welfare index \( w^+ \) is a maximin contract. A maximin contract favors the least-favored class in the sense that no other contract improves the position of the least-favored class relative to its most-preferred contract.

Let \( a^- \) denote the minimum payoff to the row player among all contracts in which the column players get their maximum, i.e., \( a^- = \min\{a_i; b_i = b^+\} \). Similarly, let \( b^- \) be the minimum payoff to the column player among all contracts in
which the row players get their maximum. Define

$$w^- = a^- \lor b^-.$$  \hspace{1cm} (5)

Normally, we would expect $w^-$ to be small, because eking out large gains for one class will generally impose a cost on the other class due to substitution possibilities. We claim that the smaller $w^-$ is, the closer is the stochastically stable outcome to the maximin solution. Specifically, define the distortion parameter $\alpha$ as follows:

$$\alpha = \frac{w^-(1 - (w^+)^2)}{(1 + w^-)(w^+ + w^-)}.$$  \hspace{1cm} (6)

Note that $\alpha$ is small when $w^-$ is small and/or $w^+$ is close to one.

**Theorem 1** (Young, 1998b). Every stable conventional contract is efficient and approximately maximin in the sense that its welfare index is at least $(w^+ - \alpha)/(1 + \alpha)$.

Although the proof of this result is somewhat involved, the intuitive explanation for it is straightforward. In a convention with payoffs near the extremes of the payoff possibility set, one or the other group has an expected payoff close to zero, i.e., its expected payoff is not much higher than the payoff from being unattached. People in this group do not have much to lose by trying something different. When change is in the air, this group will find it rational to demand a different contract. By contrast, conventions whose payoffs are centrally located in the feasible payoff set are harder to dislodge, because it takes more uncertainty (a greater accumulation of stochastic shocks) to induce any group to change. The net effect, over the long run, is to push society away from the boundaries of the feasible payoff set and toward the middle of the efficiency frontier.

We should, of course, be careful not to read too much into these results, since they are based on a stylized model of individual behavior. Moreover, social change occurs for a complex variety of reasons. It is driven to some degree by the opinions and actions of influential people (role models), who precipitate change because they are widely noticed and imitated. Furthermore, changes in one sphere of interaction may have important spillover effects on other spheres, and are related to broad underlying trends in society. Conventions about seating on buses reflect attitudes about race, norms in marriage have much to do with the relative status of women and men in society, and so forth. In this paper we have abstracted away from such complications. Instead, we adopted the hypothesis that changes in convention are driven by the cumulative effect of many small variations in behaviors and expectations, just as biological change is driven by variations in genetic structure. If this hypothesis is correct, the model outlined above gives us grounds for (cautious) optimism. It suggests that, in the long run,
evolutionary forces tend to favor forms of contracts that are economically efficient, and that are more or less egalitarian in their welfare implications.

References