The Evolution of Social Norms

H. Peyton Young

October 1, 2014

Abstract. Social norms are patterns of behavior that are self-enforcing at the group level: everyone wants to conform when they expect everyone else to conform. There are multiple mechanisms that sustain social norms, including a desire to coordinate, fear of being sanctioned, signaling membership in the group, or simply following the lead of others. This article shows how stochastic evolutionary game theory can be used to study the dynamics of norms. We illustrate with a variety of examples drawn from economics, sociology, demography, and political science. These include bargaining norms, norms governing the terms of contracts, norms of retirement, duelling, footbinding, medical treatment, and the use of contraceptives. These cases highlight the challenges of applying the theory to empirical cases. They also show that the modern theory of norm dynamics yields insights and predictions that go beyond conventional equilibrium analysis.

Keywords: evolutionary game theory, equilibrium selection, stochastic stability

JEL: C73, A120, O10

Forthcoming in the *Annual Review of Economics*.

DOI 10.1146/annurev-economics-080614-115322

Contents

1. Overview
2. Mechanisms supporting normative behavior
3. Key features of norm dynamics
4. Evolutionary models of norm dynamics
   4.1 The basic framework
   4.2 Stochastic evolutionary game theory
   4.3 Ergodic distributions and stochastic stability
   4.4 Qualitative features of the dynamics
5. Naming
6. Bargaining
7. Contracting
8. Other cases
   8.1 Footbinding
   8.2 Dueling
   8.3 Fertility
   8.4 Medical treatment
9. Directions for future research

1. **Overview**

Social norms are rules of behavior that govern interactions with others. They are the unwritten codes and informal understandings that define what we expect of other people and what they expect of us. Norms establish standards of dress and decorum, obligations to family members, property rights, contractual relationships, conceptions of right and wrong, notions of fairness, the meanings of words. They are the building blocks of social order. Despite their importance, however, they are so embedded in our ways of thinking and acting that we often follow them unconsciously and without deliberation, hence we are sometimes unaware of how crucial they are to navigating social and economic relationships.

There is a substantial literature on social norms in philosophy, sociology, law, political science, and economics. Given space limitations it is impossible to provide a comprehensive account of this literature here.[1] Instead I shall focus on the question of how social norms evolve and how norm shifts take place using evolutionary game theory as the framework of analysis. Although this framework is relatively recent, many of the key concepts originate with Hume (1740):

> "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 express'd, and is known to both, it produces a

[1] See in particular Schelling 1960, 1978; Lewis 1969; Ullmann-Margalit 1977; Akerlof 1980; Axelrod 1984, 1986; Sugden 1986; Coleman 1987, 1990; Elster 1989a, 1989b; North 1990; Ellickson 1991; Young 1993, 1998; Bernheim 1994; Skyrms 1996, 2004; Binmore 1994, 2005; Posner 2000; Hechter and Opp 2001; Bowles 2004; Bicchieri 2006; Burke and Young 2011.

> suitable resolution and behavior. And this may properly enough be call'd a convention or agreement betwixt us, tho' without the interposition of a promise; since the actions of each of us have a reference to those of the other, and are perform'd upon the supposition, that something is to be perform'd on the other part. Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. Nor is the rule concerning the stability of possession the less deriv'd from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. On the contrary, this experience assures us still more, that the sense of interest has become common to all our fellows, and gives us a confidence of the future regularity of their conduct…In like manner are languages gradually establish'd by human conventions without any promise. In like manner do gold and silver become the common measures of exchange…"

In this remarkable passage Hume identifies three key factors in the evolution of norms: i) they are equilibria of repeated games; ii) they evolve through a dynamic learning process; iii) they underpin many forms of social and economic order. In recent years these ideas have been formalized using the tools of evolutionary game theory (Foster & Young 1990; Young 1993, 1995, 1998; Kandori et al. 1993; Skyrms 1996, 2004; Binmore 1994, 2005; Bowles 2004).[2] Although the theory is well-developed, however, there has been relatively little research that brings it into contact with empirical examples. I have therefore structured this article around seven case studies, some of which are based on experiments, some on historical narratives, and some on field data. My aim is to show how evolutionary game theory can illuminate the evolution of norms in real situations despite limitations in the data. The discussion will also highlight some shortcomings of current theory and point out directions for future research.

---

[2] Other textbook treatments of evolutionary game theory without a particular focus on norms include Vega-Redondo (1996), Samuelson (1997), Hofbauer & Sigmund (1998), and Sandholm (2010).

Let me be clear at the outset that I shall not attempt to distinguish between various categories of social norms, such as customs, conventions, conceptions of right and wrong, notions of propriety, and regularities of behavior. Instead I shall use the term 'social norm' to cover a constellation of behaviors that range from fine points of etiquette to strong conceptions of moral duty. In all of their incarnations social norms exhibit certain key features that have important implications for the ways in which they evolve. These include the following.

1. Norms are behaviors that are self-enforcing at the group level: people want to adhere to the norm if they expect others to adhere to it. The precise mechanisms and motivations that create this positive feedback loop differ substantially from one context to another, as I shall show in subsequent sections.

2. Norms typically evolve without top-down direction, through a process of trial and error, experimentation, and adaptation. They illustrate how social order is constructed through interactions of individuals rather than by design.

3. Norms can take alternative forms, that is, they govern interactions that have multiple equilibria. Consequently they are contingent on context, social group, and historical period. Consider the implications of having an illegitimate child, ignoring a challenge to a duel, binding your daughter's feet, leaving all your property to your eldest son, practicing contraception, keeping a mistress, burping at the end of a meal, or dancing at a funeral. In some societies these would represent serious norm violations; in others they are perfectly normal practice.

2. Mechanisms supporting normative behavior

Norms come in different guises, serve a variety of purposes, and are sustained as equilibria by multiple mechanisms. The literature provides a rich account of different types of norms and the social and psychological mechanisms that support them.[3] Here I shall briefly summarize some of the main mechanisms that support social norms without attempting to be exhaustive.

Coordination. Examples of pure coordination include using words with their conventional meanings, carrying standard forms of money to make purchases, and agreeing on the terms of business contracts. The motive is to coordinate with others in a particular type of interaction; there is no need to punish deviants because failure to coordinate is inherently costly.

Social pressure. Norms often prescribe behaviors that run counter to an individual's immediate self-interest, in which case they are sustained by the prospect of social disapproval, ostracism, loss of status, and other forms of social punishment. Numerous forms of cooperation are enforced in this manner, but so are dysfunctional norms, such as dueling, blood feuds, footbinding, and conspicuous consumption.[4]

Signalling and symbolism. Some norms convey intentions, aspects of personal character, or signal membership in a group. Although the behaviors themselves are of little consequence, they have important reputational implications. Dress codes are often used to signal membership in particular groups or the holding of particular preferences, such as veiling by Muslim women (Carvalho 2013), 'hanky codes' among gay men, and tattoos in criminal gangs (Gambetta 2009). A current example is the use of flag lapel pins by US politicians to signal their patriotism. Other signals connote general rather

---

[3] See in particular Homans (1950), Coleman (1987, 1990), Elster (1989a, 1989b), Hechter and Opp (2001), Bicchieri (2006).

[4] For other examples of dysfunctional norms see Edgerton (1992). There is an on-going debate about what motivates third parties to inflict costly punishments, but laboratory experiments suggest that they are in fact willing to do so (Fehr et al. 2002; Fehr and Fischbacher 2004a, 2004b)

than specific traits: observing fine points of etiquette, showing up on time, speaking in turn, and displaying appropriate degrees of deference are often taken to be signs of reliability and trustworthiness (Posner 2000). The mechanism that sustains such symbolic norms is the negative interpretation that others place on deviations.

Benchmarks and reference points. Social regularities in behavior sometimes result from the use of benchmarks or reference points to make decisions. Schelling (1960) emphasized the importance of "focal points" to solve coordination problems, such as fifty-fifty division in bargaining, but there are other types of socially constructed reference points that serve as heuristics for making individual decisions.[5] These often take the form of targets, benchmarks, or thresholds, such as what time of day to have the first drink, what fraction of one's income to give to charity, how much to save for retirement , how old one should be before getting married or how much one should weigh.[6] Such benchmarks derive their salience from the fact that they are widely used; they are a convenient way of making (or justifying) a decision that people would find difficult to solve on their own. Thus, as in the previous cases, there is a positive feedback effect between the behaviors of the group and the actions of the individual.

A particularly interesting example is the age at which people customarily retire. In the United States, the social security system was originally designed to encourage retirement at age 65. In the early 1960s the law was changed so that the expected benefits remained essentially the same for anyone retiring between 62 and 65 but many people did not take advantage of the change. Indeed the distribution of retirement ages continued to have a spike at 65 until well into the 1990s (Axtell & Epstein, 1999; Burtless 2006). It appears that once 65 became fixed in the public mind as the 'normal'

---

[5] For more on heuristics see Kahneman et al.(1982) and Gigerenzer et al. (1999).

[6] Survey data suggests that people's choice of body weight is subject to social influence, that is, it depends in part on the distribution of body weights in their social reference group (Burke & Heiland 2007; Burke, Heiland & Nadler 2010; Hammond & Ornstein 2014).

age to retire, it continued to serve as a heuristic for many years after it had ceased to have any particular economic significance.[8]

The preceding discussion does not cover all of the mechanisms that can support a social norm, nor are these mechanisms mutually exclusive. Footbinding one's daughter enhances her marriage prospects (a coordination motive) and also signals the family's status. Retiring at 65 may be largely a heuristic decision, but it may also have elements of coordination (retiring when one's friends retire). Using bad table manners signals a lack of sensitivity to social norms (symbolic), and it may preclude future profitable dealings with others at the table (coordination failure).

I have also not restricted the concept of social norm to behaviors that have a moral or injunctive character, as is common in much of the literature (see among others Homans 1950; Coleman 1990; Elster 1989a, 1989b; Bicchieri 2006). Some of the examples discussed under the previous headings are of this nature, but many of them are held in place by other mechanisms that are equally powerful. The key to the analysis of norm dynamics is the existence of a positive feedback loop between social and individual behaviors, whether or not this feedback arises from the prospect of social punishment. This feedback loop has important consequences for the evolution of norms, as we shall see in the following sections.

3. Key features of norm dynamics

As mentioned at the outset the purpose of this article is to review key elements of the theory of social norms and to discuss recent attempts to bring the theory to data. In the

---

[8] The origin of 65 as the standard retirement age can be traced to the pension system devised by Bismarck in the 1880s. In fact, the original age in the Prussian system was 70; later it was reduced to 65 (Cohen 1957; Myers 1973). Many European countries and the United States subsequently copied various features of the Prussian system in their own retirement arrangements. Thus 65 served as a convenient reference point for those designing the retirement systems, and also for the individuals within these systems once 65 became embedded in the public consciousness.

case of field data it is often difficult to identify the effects of a norm due to spurious correlation and the presence of multiple feedback channels (Manski 1993; Moffitt 2001; Brock & Durlauf 2001b). In spite of these difficulties, much can be said about the qualitative features of the dynamics that hold irrespective of the specific channels that are governing the feedback effects. These include the following.

Persistence. Norms tend to persist for long periods, and to respond very sluggishly to changes in external conditions that alter the benefits and costs of adhering to the norm. Persistence is a striking feature of certain contractual norms, such as customary shares for landowners and tenants in agricultural contracts, percentages charged by tort lawyers in malpractice and product liability cases, performance fees set by fund managers, commissions for real estate agents, and so forth (see section 7). Once established by precedent such shares can persist for many years in spite of changing economic conditions. Norms of cooperation and mutual trust, or the lack thereof, can persist for centuries once they are established in a given society, as can norms that support persistent inequalities between social classes (Banfield 1958; Putnam 1993; Tabellini 2008; Guiso et al. 2009; Acemoglu and Robinson 2012; Belloc and Bowles 2013; Hwang et al. 2014).

Tipping. When norm shifts do occur, the transition tends to be sudden rather than incremental (Schelling 1960, 1978). Once a crucial threshold is crossed and a sufficient number of people have made the change, positive feedback reinforces the new way of doing things and the transition is completed rapidly.

These two effects – persistence and tipping -- create a characteristic signature in the historical evolution of norms: there are long periods of no change punctuated by occasional bursts of activity in which an old norm is rapidly displaced by a new one.

This is the *punctuated equilibrium effect* (Young 1998).[9] The footbinding of women was a norm in China for over a thousand years; yet it disappeared almost entirely within a single generation (see section 8.1 below). In our own times attitudes toward homosexual behavior have undergone a radical shift within the space of a few decades.

Compression. An important implication of positive feedback is that individual choices tend to exhibit less variation than would otherwise be expected. Burke and Young (2011) refer to this effect as "conformity warp"; here I shall call it *compression*. For example, if people did not use a socially constructed retirement age as a heuristic, the distribution of retirement ages would be less concentrated than it actually is. If there were no focal bonus rate in the hedge fund industry there would be a greater range of negotiated fees. If landowners and tenants did not gravitate toward normative contracts, such as the fifty-fifty division of output, one would see more diversity in contract terms, depending on underlying fundamentals such as the quality of land and labor inputs, risk aversion, and the value of the parties' outside options (see section 7).

*Local conformity/global diversity***.** Norms are population-level equilibria in interactions that typically have multiple equilibria. They evolve through a process in which chance events play an important role, hence the norms in a given community at a given point in time are inherently unpredictable. When communities do not interact, or interact only occasionally, they may follow different evolutionary trajectories and thus have different norms even though in other respects they are similar (Young 1998). In the US, for example, the standard treatment for certain medical conditions differs markedly from one region to another even when there are no significant differences in the cost of performing these procedures (see section 8.4).

---

[9] In biology this term has a more specific interpretation; here I use it to describe the characteristic look of a dynamical path.

4. Evolutionary models of norm dynamics

4.1. The basic framework

Consider a population of players who interact over time, where each interaction entails playing a normal form coordination game G. For expositional simplicity I shall usually assume that G is a two-person coordination game, but much of the theory extends to the n-person case. Depending on the context I shall sometimes assume that G is symmetric and that pairs of players are drawn at random from a single population to play the game. At other times I shall assume that G is asymmetric and players are matched at random from disjoint populations that occupy different roles in the game. It is assumed that the populations are sufficiently large that the actions of any single individual will typically have a small impact on the dynamics of the process as whole. More generally I shall make the following background assumptions.

i) People do not have full information about what is going on in society at large. Their information is limited to their own prior experience and a sample of experiences of others in their social group.

ii) People typically choose myopic best responses given their information. Myopia is consistent with rationality given our assumption that a given player cannot influence the future course of the process.

iii) People sometimes deviate from myopic best responses for a variety of reasons, including inattention, experimentation, idiosyncratic beliefs, or unobserved payoff shocks.

iv) People interact at random with others in their social group, usually with some bias toward their geographical or social "neighbors."

This framework occupies a middle ground in the level of rationality attributed to the agents. It presumes that people are purposeful and that they adapt their behavior based on expectations of others' behavior. It does not presume that they know the overall process in which they are embedded, or that they attempt to manipulate the process through strategic, forward-looking behavior. In a large population where people interact at random, it would be rare for a given individual to be in a position to alter the dynamics in a reasonable length of time. Thus, in such a setting, forward-looking rational agents will often act *as if* they were myopic (Ellison, 1997).

There are exceptions however. When people interact in small close-knit groups, such as villages, local religious communities, and small professional organizations, a single individual may be able to steer the evolutionary process toward a new norm within a fairly short period of time. This is particularly true if the individual is in a position to set a public example, such as a religious leader or village elder. Such 'norm entrepreneurs' stand to enhance their status if the new norm is beneficial to the community (Sunstein 1996; Posner 1998; Ellickson 2001; Acemoglu & Jackson 2012). It should be noted, however, that being prominent is often a liability because such individuals have a lot to lose should their efforts fail: norm entrepreneurship is a risky undertaking for the well-connected. For this reason, politicians, religious leaders, and village elders may be among the least willing to induce norm shifts even though they are the ones most capable of doing it. (It has been argued that this is one of the reasons that dueling persisted for so long; see section 8.2.)

In any event we do not need to know the preferences, beliefs, and degree of rationality of every individual in the model. Since we are dealing with a population of agents, we can view departures from myopic best response as the result of unobserved heterogeneity. As long as these departures have a relatively low probability the method of analysis still applies. Moreover if we know more about the distribution of behavioral types in the population, we can obtain a more exact representation of the dynamics though still with some residual error that arises from unmodeled behavioral characteristics. For example, heterogeneity in risk aversion, heterogeneity in information, and heterogeneity in the degree of rationality attributed to other agents (k-level reasoning) can be treated within the same general framework (Young 1993b, 1998; Hurkens, 1995; Saez-Marti & Weibull 1999).

Before leaving the topic of rationality I should remark that there is an important branch of the evolutionary games literature that attributes *less* rationality to the agents. This is the biological model of the evolution of cooperation due to Axelrod (1984, 1986). In this framework people do not consciously optimize or form beliefs; rather they are endowed with particular strategies whose reproductive success hinges on how well they do in competition with alternative strategies. Given space limitations I shall not elaborate on this approach here, except to note that the stochastic evolutionary framework can also be applied to this case (Foster and Young 1991).

In the remainder of this survey I shall use the perturbed best response framework to examine the following questions. First, under what conditions do the interactions of many dispersed individuals eventually coalesce into a social norm, that is, to a situation in which people's actions and expectations constitute an equilibrium at the level of the group? Second, are some norms more likely to emerge than others? Third, what are the characteristic features of the dynamics in the intermediate and in the long run? Fourth, what are the welfare implications of these dynamics?

4.2. Stochastic evolutionary game theory

To address these questions we need to formalize the process by which individuals form expectations and update their behavior. To simplify the exposition we begin by considering a symmetric two-person game G and a single population of players who interact in pairs and play G. Later we shall see how to extend the framework to asymmetric games and multiple populations.

The set of available actions to each player will be denoted by X. Given a pair of actions (x, x'), let $u(x, x')$ denote the payoff to the first player and $u(x', x)$ the payoff to the second player. Thus the payoffs define a symmetric two-person game G. In most applications the players will be embedded in a social or geographical space that determines the probability with which each pair will interact, and also the importance or 'salience' of their interaction. To accommodate these differences, let us assume that each pair of players (i, j) has an importance weight $w_{ij} \geq 0$. For simplicity we shall assume that the weights are symmetric, that is, $w_{ij} = w_{ji}$ for all $i \neq j$. This holds for example if w_{ij} measures the *proximity* of *i* and *j* in a suitably defined space. If the weights are binary, say $w_{ij} \in \{0,1\}$, the pairs $\{i, j\}$ such that $w_{ij} = 1$ can be interpreted as the edges of a network.

The dynamics can be modelled as follows. There is a population consisting of n players with identical utility functions. Time periods are discrete and denoted by $t = 1, 2, 3, ...$ At the end of period t, the *state* is an n-vector $\boldsymbol{x}(t)$, where $x_i(t) \in X$ is the current choice of action by each player i, $1 \leq i \leq n$. The *state space* will be denoted by $\boldsymbol{X} = X^n$. Assume that the players update their strategies asynchronously: at the start of period $t+1$, one agent, say *i*, is selected uniformly at random to update. Given the current choices of everyone else, $\boldsymbol{x}_{-i}(t)$, define the utility of agent i from choosing action x to be

$$U_i(x, \boldsymbol{x}_{-i}(t)) = \sum_{i,j} w_{ij} u(x, x_j(t)) . \qquad (1)$$

This expression can be interpreted in several ways. Suppose for example that in each period one pair $\{i, j\}$ is drawn from the set of all pairs with probability $p_{ij} = w_{ij} / \sum_{1 \le h < k \le n} w_{hk}$. In this case $U_i(x, \boldsymbol{x}_{-i}(t))$ is i's expected utility from adopting action x given the actions of the other members of the population. Thus the functions $U_i(x, \boldsymbol{x}_{-i}(t))$ define an n-person game, which we shall call the *social game* induced by G and the weights w_{ij} .

An alternative interpretation of (1) is that $U_i(x, \boldsymbol{x}_{-i}(t))$ is not a von Neumann Morgenstern utility but rather a *propensity* to adopt various actions given the actions of others. This interpretation can accommodate a variety of 'spillover effects' on individual behavior due to social influences; the influence does not necessarily arise from 'playing games' with one's neighbors. (For example, the decision about what to wear to a dinner party is influenced by one's expectations about what others will wear; there is no need to model the situation as a coordination game played with other guests at the party.)

Assume that when agent i updates his choice, he chooses a new action $x_i(t+1) = x$ with a probability that is increasing in $U_i(x, \boldsymbol{x}_{-i}(t))$. In particular we shall usually assume that *i's* choice maximizes $U_i(x, \boldsymbol{x}_{-i}(t))$ with high probability, and that he chooses other actions with low probability. We shall call this a *perturbed best response model*.

In the evolutionary games literature there are two benchmark models of perturbed best response behavior. The *uniform error model* posits a small error rate $\varepsilon \in [0,1]$ such that with probability 1 – ε agent i chooses an action that maximizes $U_i(x, \boldsymbol{x}_{-i}(t))$, and with probability ε he chooses an action uniformly at random. (If there are several actions that maximize $U_i(x, \boldsymbol{x}_{-i}(t))$ they are chosen with equal probability.) An alternative approach is the *log-linear response model* (Blume 1993, 1995; Durlauf 1997; Blume & Durlauf 2001, 2003; Brock & Durlauf 2001a, 2001b). Let β be a nonnegative real number and suppose that *i*'s probability of choosing action *x* at time *t + 1* be given by the expression

$$P[x_i = x] = e^{\beta U_i(x, \boldsymbol{x}_{-i}(t))} / \sum_{x' \in S} e^{\beta U_i(x', \boldsymbol{x}_{-i}(t))} . \qquad (2)$$

This is a perturbed best response model in which the probability of deviating from best response decreases the greater the corresponding loss in utility. As β becomes large the probability of choosing a best response approaches one. In the limiting case ($\beta = \infty$) a best response is chosen with probability one. [11]

The preceding framework can be generalized to include the history of actions taken over many periods, not just those taken in the immediately preceding period. We can also allow for the case of asymmetric games by supposing that players are drawn from multiple populations. To illustrate these extensions, consider a two-person game *G* with action space *X* for the row players and a possibly different action space *Y* for the column players. In each period prospective row and column players are drawn from two disjoint populations *R* and *C* to play the game *G*. Let $x^t \in X^R$ be the vector of

[11] An alternative interpretation is that the agent chooses a best response with certainty and his utility is subject to an unobserved payoff shock. When these shocks are i.i.d. extreme-value distributed the resulting probabilities are given by expression (2). This is a standard model in the discrete choice literature (McFadden 1974).

actions taken by the row players at time *t*, and let $y^t \in Y^C$ be the vector of actions taken by the column players at time *t.* Let δ_{ij}^t be an indicator function such that $\delta_{ij}^t = \delta_{ji}^t = 1$ if the pair {i, j} interacted at time t, and $\delta_{ij}^t = \delta_{ji}^t = 0$ otherwise. The history through period *t* can be represented by the sequence

$$h^t = (x^1, y^1, \delta^1), (x^2, y^2, \delta^2) \ldots (x^t, y^t, \delta^t) . \qquad (3)$$

In practice people tend to put more weight on recent events than on distant ones. This *recency effect* can be modelled by discounting prior actions at an exponential rate. Alternatively one can truncate the history to some finite number of periods and assume that all actions taken within this time frame have equal weight.

The next step is to model the players' updating process. For example, in each period a player might choose a best response to the actions taken by the players with whom he was matched. Or he might choose a best response with probability $1 - \varepsilon$ and choose randomly with probability ε. This is the *uniform error model based on personal experience*. Similarly one can define the *log-linear response model based on personal experience*.

In some situations players are only matched occasionally, hence personal experience will not provide much information on which to base a decision. (Dueling is an example.) One way to accommodate this is to suppose that when a player is selected to play, he draws a random sample from the history of actions taken by prior players in the opposite population, and chooses a trembled best response to the resulting sample frequency distribution, augmented by those actions he observed from personal experience.

4.3. Ergodic distributions and stochastic stability

The models described above can be analyzed using the theory of large deviations (Freidlin & Wentzell 1984). Here I shall restrict attention to the case where the state space is finite, in particular the action spaces are finite, historical memory is finite, and the populations are finite. In this case we obtain an irreducible Markov chain that has a unique ergodic distribution. When the departures from best response have very low probability, this distribution will typically be concentrated on a unique state (or in case of ties on a small subset of states). These are called the *stochastically stable states* of the evolutionary process (Foster & Young 1990; Young 1993a; Kandori et al. 1993).

There are various techniques for computing the stochastically stable states (Young, 1993a; Kandori et al. 1993; Blume 1995; Ellison 2000). An especially tractable case arises when: i) players respond only to the prior-period actions of others using a log-linear response function as in (1) and (2); ii) the game G is a two-person symmetric potential game. In general, a potential game has the property that there is a function $\rho : X^2 \to R$ such that

$$\forall x, x', y \in X, \rho(x', y) - \rho(x, y) = u(x', y) - u(x, y). \tag{4}$$

In other words, a unilateral change of action leads to a change in payoff that exactly equals the change in the potential function ρ (Monderer & Shapley 1996). A potential function for G can be used to define a potential function for the n-person game with payoff functions $U_i(\boldsymbol{x}) = \sum_{i,j} w_{ij} u(x_i, x_j)$ as follows:

$$\tilde{\rho}(\boldsymbol{x}) = \sum_{i,j} w_{ij} \rho(x_i, x_j). \tag{5}$$

This framework can be extended to include heterogeneous preferences for taking different actions. Suppose that the payoff of each player *i*'s consists of two parts: i's expected payoff from playing a game *G* against others, $U_i(\boldsymbol{x}) = \sum_{i,j} w_{ij} u(x_i, x_j)$, and *i*'s enjoyment of the action itself, say $v_i(x_i)$. Then i's total utility is

$$U_i^*(\boldsymbol{x}) = \sum_{i,j} w_{ij} u(x_i, x_j) + v_i(x_i). \tag{6}$$

It is straightforward to show that if *G* has potential function *ρ*, then (6) defines a game with potential function

$$\rho^*(\boldsymbol{x}) = \sum_{i,j} w_{ij} \rho(x_i, x_j) + \sum_i v_i(x_i). \tag{7}$$

Theorem 1. *The evolutionary process with the potential function in (7) has a unique ergodic distribution, where the long-run probability of each state* $\boldsymbol{x}$ *is given by*

$$\mu(\boldsymbol{x}) = e^{\beta\rho^*(\boldsymbol{x})} / \sum_{\boldsymbol{x}' \in \boldsymbol{X}} e^{\beta\rho^*(\boldsymbol{x}')}. \tag{8}$$

Corollary 1.1. *The stochastically stable states of this process are precisely those states* $\boldsymbol{x}$ *that maximize the potential function* $\rho^*(\boldsymbol{x})$.

We shall show how this model can be applied to an empirical case in section 7.

4.4. Qualitative features of the dynamics

Although the models discussed above differ in their details, they have similar qualitative properties. First, all of them exhibit the dichotomy of persistence and tipping: there are long periods in which people follow a norm (with some idiosyncratic variation), but every so often these tranquil periods are interrupted by a burst of activity in which one norm is displaced by another (the *punctuated equilibrium effect*). Second, the existence of a norm creates excess uniformity within a given community; in other words people make less diverse choices than they otherwise would (the *compression effect*). Third, the evolutionary process is inherently stochastic and unpredictable, hence communities that do not interact may operate under different norms even though they are exposed to the same influences and have similar socioeconomic characteristics. In short, there will be greater diversity between communities than within them after controlling for fundamentals (the *local conformity/global diversity effect*).

In the remainder of the paper I shall illustrate these ideas through a series of seven case studies. Two of them are based on historical narratives (Chinese footbinding and dueling), two on laboratory experiments (naming and bargaining games), and three on field data (norms of fertility, norms of medical practice, and contractual norms in agriculture). I treat three of these cases in some detail (naming, bargaining, and contracts) and the remaining four in synoptic form. There are of course many other empirical studies of social norms, some of which I list in the concluding section. I have chosen to concentrate on these few in order to highlight the complexities that must be dealt with in applying the theory to actual cases, and to illustrate how one can obtain useful insights in spite of data limitations. The cases also point to ways in which current theory can be extended.

5. **Naming**

Our first case study is an experiment designed to show how naming conventions can arise in a large population through a process of trial and error (Centola and Baroncelli 2014; see also Baroncelli et al. 2006). The basic set-up is a pure coordination game: two people are shown a picture of a face and they simultaneously and independently suggest names for it. If they provide the same name, they earn a reward; otherwise they pay a small penalty. There is no restriction on the names that subjects can provide – this is left up to their imagination.

Each trial consists of 25 rounds in which the same face is shown to all subjects in all rounds. The number of subjects in each trial ranged from 24 to 96. At the start of a given round, each subject is paired anonymously with another subject and they are given 20 seconds in which to provide a name. Pairs are drawn at random from the edges of a fixed, undirected network. The subjects have no information about the structure of the network or the identities of the people they were paired with; they know only the names that their partners provided in previous rounds. Thus they accumulate information about the names that are currently popular among the people they are being paired with. This fact permits naming conventions to emerge spontaneously. Of particular interest is the effect the network structure has on the evolutionary process.

The dynamics can be modelled in the manner discussed in section 4. Each subject is located at the node of a fixed network (which is not known to the subjects). In each round one edge of the network is selected at random and the corresponding two subjects play the name game. At the end of the round each learns the name chosen by his or her partner. The history through round t is a sequence of form $h^t = (x^1, \delta^1), (x^2, \delta^2) \ldots (x^t, \delta^t)$, where x_i^t is the name player i offered in round t, $\delta_{ij}^t = 1$ if players i and j were paired in round t, and $\delta_{ij}^t = 0$ otherwise. In particular at the end of round t player i knows the names provided by all partners with whom he was paired in rounds $1 \leq t' \leq t$.

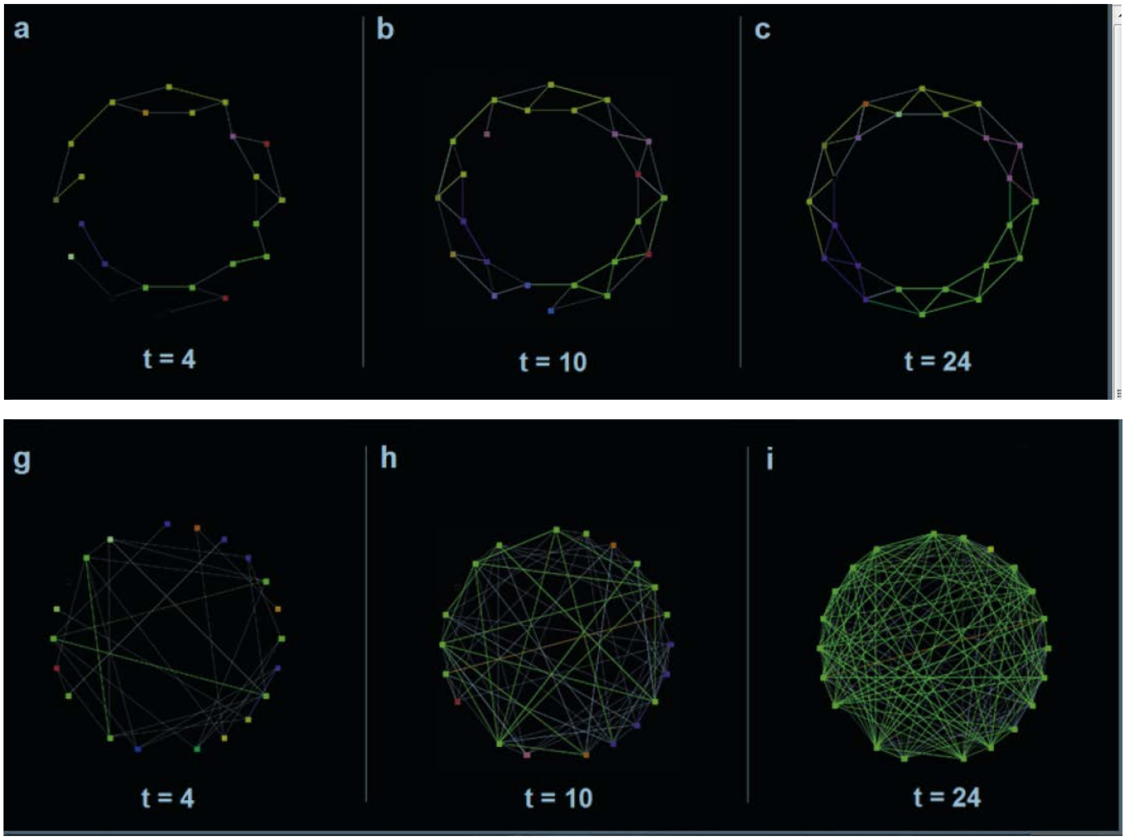

Figure 1. The Name Game. An edge is colored if the two subjects proposed the same name, and is white otherwise. Different colors represent different names. Top panel: results from the ring network. Bottom panel: results from the complete network. Source: Centola & Baroncelli 2014.

Let us compare the results from two different types of network: i) a ring in which each player is adjacent to his nearest four neighbors; ii) the complete graph in which everyone is adjacent to everyone else. The top panel of Figure 1 shows a representative evolutionary path for the ring network. By round 10 several distinct local conventions have emerged in which small groups of nearby agents choose the same name. There are five local conventions altogether with contested border regions separating them. The bottom panel of Figure 1 shows a representative path for the complete graph.

Compared to the ring network, there are more failures to coordinate in the initial rounds, but by round 10 a dominant convention has become established, and by round 24 this convention has displaced all of the others. These results accord with theory, which predicts that clustering accelerates the formation of local norms, which can then persist for a substantial length of time before merging into a global norm. Note that these patterns emerge without the subjects knowing anything about the overall structure of the network or the state of the system.

6. Bargaining

The next case is an experimental study of the evolution of bargaining norms (Gallo, 2014). There is a population of subjects called "buyers" who repeatedly play the Nash demand game against a "Seller." In each period one buyer is matched with the Seller and they demand shares of a fixed pie. If their demands do not exceed the total size of the pie, they receive rewards proportional to their demands; otherwise they get nothing (Nash, 1950). Before making his demand, the buyer receives information about the demands made by the Seller in previous encounters with *other* buyers. The Seller is programmed to play a perturbed best response strategy to a sample of previous demands of the buyers, and the buyers know this.[12]

A key feature of the experiment is to vary the network through which buyers obtain information from other buyers, and to see what effect the network topology has on the trajectory of play. Specifically the experiment permits an examination of the following questions. How often does the process converge to a norm, that is, a state in which almost all buyers demand the same amount irrespective of their position in the

[12] This is a variant of the evolutionary bargaining model proposed in (Young, 1993b). For other extensions and variations of this model see Saez-Marti & Weibull 1999; Binmore et al. 2003; Bowles et al. 2010.

network? How many rounds on average does it take for convergence to occur? Does the network topology affect the resulting distribution of norms?

We now describe the experimental set-up in greater detail. A given trial involves a fixed group of 6 buyers plus the Seller.[13] The buyers are located at the nodes of a fixed network, which determines the channels of communication between them. Each trial consists of 50 rounds of play. In every round, each buyer is matched once with the Seller in the Nash demand game. The buyer learns only whether his demand was compatible with the Seller's; he is not told how much the Seller demanded. Demands are constrained to be nonnegative integers and the total size of the pie is 17 units. Hence a pair of demands (x, y) is compatible if and only if x, y are integers, $x, y \geq 0$, and $x + y \leq 17$. The number 17 was chosen to reduce the focal qualities of certain solutions; in particular simple fractions such as ½ and 1/3 cannot be implemented in whole numbers.

When a given buyer *b* is about to make a demand he is told what demands were made by the Seller in a random sample of prior matches with *b*'s neighbors in the previous six rounds. The idea is that each buyer learns about some of the Seller's prior demands through his network of contacts. The size of the buyer's sample is $2d_b$ where d_b is the number of neighbors to whom *b* is connected. From this information the buyer can draw inferences about the demands the Seller is currently making; he also knows from prior matches which of his own demands led to successful outcomes. When matched against any given buyer, the Seller samples from the prior demands made by all buyers within the last six rounds; she then chooses a perturbed best response. Specifically, the Seller is programmed to choose a myopic best response to the frequency distribution of

[13] The experiments were conducted at the Center for Experimental Social Science at Nuffield College, University of Oxford. Subjects consisted of both undergraduate and graduate students from the university.

demands in her sample with 95% probability, and to make a demand at random with 5% probability.[14]

Define a *bargaining norm* to be a situation in which at least 5 out of 6 subjects (i.e., buyers) demand the same amount for at least four consecutive periods. Consider the situation where the network is regular of degree 4 (see figure 2, left panel).

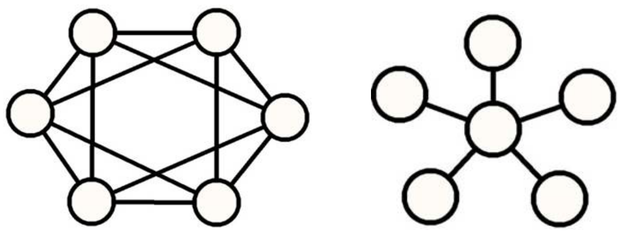

Figure 2. Regular network (left panel) and star network (right panel)

Fix a particular buyer *b*. At the start of a given round $t \geq 7$, b receives a random sample of demands made by the Seller in interactions with *b*'s neighbors over the previous six periods. There were 24 such interactions in all, and the size of the sample is $2d_b$ = 8. Given this information the buyer then chooses a demand. Simultaneously the Seller *S* generates a demand as follows: *S* chooses a random sample of size 8 from all demands made by all buyers in the last six periods (6 x 6 = 36). With probability 95% *S* chooses a best response to the sample frequency distribution, and with probability 5% *S* chooses a demand uniformly at random from the set of integers {1, 2, …,17}. Note that before making a demand of each buyer, *S* resamples from the data, hence the demands will generally differ across buyers.

[14] The Seller follows this procedure separately for each buyer, that is, she resamples and chooses a perturbed best response. The best response is computed under the assumption of risk neutrality.

The results are displayed in Figure 3, left panel. Convergence occurred in 11 out of 13 trials, and the resulting norms ranged from 9 to 14. A particularly interesting finding is that convergence typically occurred quite rapidly. Indeed, in the 11 out of 13 cases where convergence took place, the average expected waiting time was 33 rounds. The stochastic evolutionary models discussed in section 4 do not speak directly to the question of how rapidly convergence to a norm will occur in practice. This depends on the amount of 'noise' in the subjects' responses.[15] The results of Gallo's experiment indicate that subjects frequently deviated from best response behavior in the early rounds of the experiment, but that the rate of deviation subsided substantially as an equilibrium was approached.

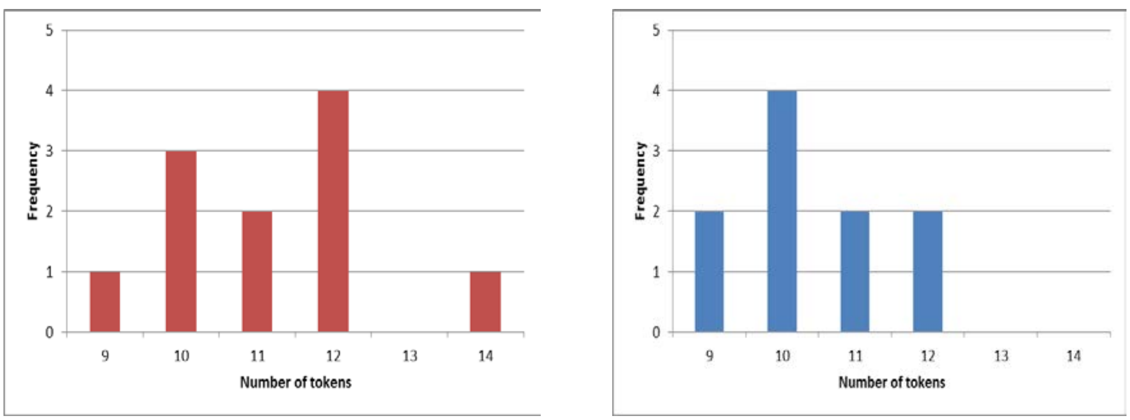

Figure 3. Frequency distribution of bargaining norms expressed in number of tokens for the buyers. Regular network (left panel), star network (right panel).

One of the aims of this experiment was to examine whether the structure of the buyers' network influences the norms to which they converge. The stochastic evolutionary

---

[15] The noise was measured as the frequency with which subjects deviated from a myopic best response (under risk neutrality) to the information they had about the Seller's previous demands. As noted earlier, deviations may arise from heterogeneity in beliefs, risk aversion, discount factors and other parameters that are not observed by the experimenter. They do not necessarily represent departures from best response from the subject's perspective.

model in section 4 makes a specific prediction about this. Namely, when subjects within a population have different sample sizes (and thus different amounts of information about the other side's actions), the stochastically stable norm is particularly sensitive to the *minimum* sample size. The smaller the minimum sample size within a given population, the less that population can expect to get, all else being equal. The reason is the following. The long-run stability of a norm depends on how easily it can be dislodged through random perturbations that cause a shift in expectations by one or both populations. An agent with a small sample size can be induced to lower his expectations given only a few instances of higher demands by the opposite population; changing the expectations of an agent with a large sample size takes more high demands by the other side, which is less likely to occur.[16]

The results of Gallo's experiment agree with this prediction. Consider the star network shown in the right panel of Figure 2. There are five agents with one neighbor each, hence in each period they receive a sample of size two. The central agent has degree five and her sample size is 10. The Seller's sample size remains as before (8). The theory suggests that, as a group, the buyers in this situation are less advantaged than in the regular network of degree 4. This is confirmed by the data: in the 8 trials where the process converges, the modal norm is 10 instead of 12, and on average the buyers receive .81 less than before (significant at the 1% level).

These results suggest a number of directions in which the theory could be usefully extended. The framework outlined in section 4 assumes that error rates are constant, and the predictions of long run equilibrium selection depend importantly on this assumption. However Gallo's experiment suggests that subjects' error rates decline quite noticeably as the process converges towards equilibrium, which makes sense intuitively. We need a theory that can accommodate the time-varying error rates that

[16] This prediction continues to hold even when agents are heterogeneous in their degree of risk aversion (Young, 1993b; 1998).

we see in practice. Given that the observed error rates are quite large, we also need a theory that treats the dynamics when the error rates are bounded away from zero. These issues are closely related to the question of how long it takes the evolutionary process to reach equilibrium. For recent work along these lines see Kreindler & Young (2013, 2014), Bowles et al. (2014), and Ellison et al.(2014).

7. Contracting

The next case uses field data to examine contractual norms in agriculture, which is a leading example of principal-agent contracting. The logic of a share contract is that it spreads the risk between the contracting parties. Theory predicts that the terms of such a contract will depend on such factors as the agent's inherent skill, the principal's ability to monitor the agent's effort, their attitudes toward risk, and the value of their outside options (Cheung 1969; Stiglitz 1974; Bolton & Dewatripont 2005). To the extent that there is heterogeneity in these factors, there should be variation in the shares that the parties negotiate.

In practice, however, share contracts often exhibit a high degree of uniformity, a phenomenon that has attracted the attention of economists since the foundation of the discipline (Smith 1776; Mill 1848). This 'excess uniformity' is a feature not only of agricultural contracts but of many other principal-agent contracts that are based on 'usual and customary' shares, such as real estate commissions (Fisher & Yavas 2010), and contingency fees for tort lawyers (Kritzer 2004).[17] The case of performance fees in the hedge fund industry is particularly well-documented (Mallaby 2010). Until very recently the standard in the industry was to charge 'two and twenty': 2% per year in

[17] Based on a survey of contingency fees in Wisconsin, Kritzer (2004 p. 39) finds that in about 60% of the cases the shares were one-third for the lawyer, two-thirds for the client (except for certain where the fees are regulated by statute).

management fees, and 20% of the profits as a performance fee irrespective of the size of the fund. The 20% rate can be traced to Alfred Winslow Jones, who originated the concept of hedge funds in the 1940s, and claimed as justification that this was the rate charged by Phoenician merchants in ancient times (Mallaby 2010, p. 30).

The evolutionary theory of norms provides a framework for understanding this phenomenon. The thesis is that, once a particular way of sharing the output becomes established in a given business and a given locale, it provides a powerful focal point that suppresses much of the variation that would arise from idiosyncratic factors. Here I shall discuss the application of this idea to agricultural share contracts in the midwestern United States, for which remarkably good data are available.

An *agricultural share contract* is an arrangement in which a landowner and a tenant split the gross proceeds of each year's harvest in fixed proportions or shares. A key advantage of such a contract is that it shares the risk of an uncertain outcome while offering the tenant an incentive to increase the expected value of that outcome. When contracts are competitively negotiated, one would expect the size of the share to reflect such factors as the expected yield per acre, the risk aversion of the parties, their outside options, and so forth. In practice, however, the shares tend to cluster around 'usual and customary' levels even when there are substantial and observable differences in the quality (expected yield) of different parcels of land, and there are different outside options for the tenants. Furthermore these contractual norms tend to be anchored at prominent focal points, such as 1/2-1/2, 2/5-3/5, or 1/3-2/3. The importance of such focal points, and the fact that they tend to be specific to particular regions, has been documented in many parts of the world, including India and Africa (Bardhan & Rudra 1980; Bardhan 1984; Robertson 1987).

Here I shall summarize a study of share contracts in the state of Illinois (Young & Burke, 2001). This study is based on survey data collected by the Illinois Cooperative

Extension Service (1995), which provide detailed information on the terms of contracts on several thousand farms in different parts of the state. This includes information on the size of the farm, the terms on which all inputs and outputs are shared, gross output, and the net incomes to tenant and landlord. A key additional feature is a measure of each farm's inherent productivity (i.e., expected yield per acre) based on the soil types found on that farm.[18] There are three striking features of these data. First, over 98% of the share contracts were based on 1/2-1/2, 2/5-3/5, or 1/3-2/3. Second, there are strong regional differences in their frequency of use: in the northern part of the state the customary share is 1/2-1/2, whereas in the southern part of the state the customary share is either 1/3-2/3 or 2/5-3/5.[19] Furthermore there is a high degree of uniformity *within* each region despite the fact that there are substantial differences in the productivities of farms in the region that would seem to call for different shares. This is an example of the *local conformity/global diversity* effect discussed in section 2. Third, there are many instances in which farms have essentially the same land quality, but when they are located in different regions they often operate under substantially different shares that reflect differences regional norms. *Compression* prevents the contractual shares from fully reflecting the underlying heterogeneity among farms.

These features are captured in Figure 4, which shows the distribution of shares in two representative counties (one in the north and one in the south), as a function of the soil quality of the farms in these counties.

---

[18] Each farm is assigned a soil quality index, which measures the expected yield per acre under standard management conditions.

[19] This north-south division corresponds roughly to the southern boundary of the last major glaciation. In both regions, farming techniques are similar and the same crops are grown -- mainly corn, soybeans, and wheat. In the north the land tends to be flatter and more productive than in the south, though there is substantial variability within each of the regions.

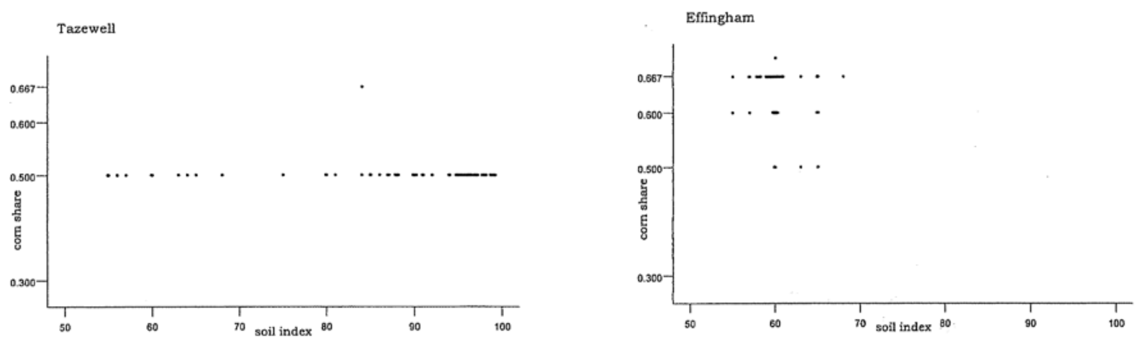

Figure 4. Distribution of shares in a northern county (Tazewell, left panel) and a southern county (Effingham, right panel). *Source*: Young and Burke (2001).

The evolutionary model discussed in section 3 can shed some light on these apparent anomalies. Let us identify each farm i with the vertex of a graph. Assume that the contract adopted at a given farm depends on three factors: its inherent productivity (output per acre), local wages from nonagricultural employment, and the frequency with which the contract is used by others in the vicinity. Let s_i denote the expected output of farm i under a standard management regime, and let w_i denote the annual wage available to the tenant from alternative employment in the vicinity of farm i. A *share contract* at i specifies a proportional share x_i for the tenant, and the complementary share $1-x_i$ for the landlord, where $x_i \in [0,1]$. Thus the tenant's expected annual income on farm i is $x_i s_i$. The *incentive-compatible* contracts are those that induce the tenant to accept farm employment instead of the alternative wage w_i, that is, they satisfy the constraint $x_i s_i \geq w_i$. The *rent* to the landlord is $v_i(x_i) = (1-x_i)s_i$.[20]

To model the impact of local custom, let us suppose that there are n farms situated at the nodes of a graph Γ. Let E denote the set of edges in Γ. Two farms are said to be

[20] This is an oversimplification: a share contract also specifies the share of the various inputs that the parties are responsible for providing. In practice, inputs such as fertilizer, seed, and herbicides are often shared in the same proportions as the output.

neighbors if there is an edge joining them. Assume for simplicity that all edge-weights $w_{ij} = 1$, that is, all neighbors of a given farm *i* have an equal degree of social influence on the contracting parties at i. For analytical convenience we shall assume that the set of feasible contracts $X \subset (0,1)$ is finite. The subset $X_i = \{x \in X : x_i s_i \geq w_i\}$ consists of the incentive-compatible contracts at *i*.

A *state* of the process is a vector $x(t) \in \prod_i X_i$ that specifies the contract in force at each farm at a given point in time t. For simplicity we shall omit the dependence on t in the following discussion. Given a state $x \in \prod_i X_i$, let $\delta_{ij}(\boldsymbol{x}) = 1$ if i and j are neighbors and $x_i = x_j$; otherwise let $\delta_{ij}(\boldsymbol{x}) = 0$. Thus $\delta_{ij}(\boldsymbol{x})$ identifies the pairs of neighbors that are coordinated on the same contract in state x. Let us posit that the *propensity* to adopt a particular contract x_i at *i* is an increasing function of its expected rent to the landlord and the degree to which it conforms to other contracts in *i*'s neighborhood subject to the incentive-compatibility constraint. Define the *adoption propensity* function as follows

$$\forall x_i \in X_i, \quad \pi_i(x_i, x_{-i}) = (1 - x_i)s_i + \gamma \sum_j \delta_{ij}(x_i, x_{-i}) \quad . \tag{9}$$

The number γ is a *social conformity parameter* that can be estimated from data. This captures theidea that the terms of a given contract will tend to be pulled in the direction of customary practice in the area.

Note that the model specifies that conformity must be precise to have any benefit; in other words, if two contracts offer slightly different shares (say 48% versus 50%) the conformity payoff is nil. An alternative approach is to suppose that the disutility from deviating from the neighbors' choices is a continuous and increasing function of the deviation, say $f_i(x) = \sum_{j \in N_i} (x_i - x_j)^2$ (see for example Durlauf 1997; Blume & Durlauf

(2001)). Under this specification, however, one would expect to see small deviations around the norm in any given region. This is not consistent with the data for Illinois, which show that over 99% of the shares are concentrated on *exactly* distinct fractional values. This is consistent with the hypothesis that these fractions are serving as focal points. Indeed focal points have the peculiar property that nearby solutions are among the *least* focal (Schelling 1960, pp. 111-112). Thus the focal point hypothesis suggests that the distribution of outcomes will be sharply concentrated on particular values; the distribution will not be smooth as would be expected if they arise from unobserved heterogeneity.

Consider the following evolutionary dynamic: from time to time contracts come up for renegotiation on each farm. Assume that these renegotiation times are described by Poisson arrival processes that are i.i.d. among farms. Thus, with probability one, no two renegotiations occur simultaneously, and we can think of the process as evolving in discrete time periods with a single renegotiation occurring in each period. Let x^t denote the state of the process in period t, and suppose that the next renegotiation occurs at farm i at the start of period t + 1. We shall suppose that the probability of adopting different contracts at i is a log-linear function of the propensities to adopt, that is, for some $\beta > 0$,

$$\forall x \in X_i,\ P[x_i^{t+1} = x] = e^{\beta \pi_i (x, x_{-i}^t)} / \sum_{x' \in X_i} e^{\beta \pi_i (x', x_{-i}^t)} . \qquad (10)$$

This process has the following potential function

$$\forall x,\quad \rho(x) = \sum_i (1 - x_i)s_i + (\gamma / 2)\sum_{i,j} \delta_{ij}(x) \ .^{21} \tag{11}$$

The term $\sum_i (1 - x_i)s_i$ represents the total *rent to land*, which we shall abbreviate by $r(\boldsymbol{x})$. The second term is γ times the quantity $c(\boldsymbol{x}) = (1/2)\sum_{i,j} \delta_{ij}(\boldsymbol{x})$, which represents the total number of edges coordinated on the same contract in state $\boldsymbol{x}$. Thus the potential function can be written in the compact form

$$\rho(\boldsymbol{x}) = r(\boldsymbol{x}) + \gamma c(\boldsymbol{x})\ . \tag{12}$$

It follows that the ergodic distribution has the Gibbs-Boltzmann form, that is, the long-run probability $\mu(x)$ of each state x is given by

$$\mu(x) \propto e^{\beta[r(\boldsymbol{x}) + \gamma c(\boldsymbol{x})]}\ . \tag{13}$$

In particular the log probability of each state $\boldsymbol{x}$ *is a linear function of the total rent to land plus the degree of local conformity.*

Given sufficiently rich event data one could compute the probability that a given contract is adopted at a given farm i, conditional on the frequency of contracts in force at other nearby farms (say farms within a given radius of i). One could then estimate γ and β from a linear regression of form

[21] To verify that this is a potential function, it suffices to check that whenever an agent i undertakes a unilateral change of action, the change in i's propensity exactly equals the change in potential.

$$\log P[x_i = y] = \beta[\pi_i(x, x_{-i}) - \pi_i(y, x_{-i}) + \varepsilon_i]$$

$$= \beta s_i(y - x) + \beta\gamma(\sum_j (\delta_{ij}(x, x_{-i}) - \delta_{ij}(y, x_{-i})) + \beta\varepsilon_i \quad (14)$$

Note that this framework does not *presume* that conformity matters, but if $\beta\gamma > 0$ at a high level of significance this would provide support for the hypothesis that it does matter. Young & Burke (2001) did not attempt such an estimation due to data limitations. However, the model makes other predictions that are corroborated by the Illinois data. One of these predictions is the local conformity/global diversity effect discussed in section 2. Specifically the model predicts that, for a wide range of parameter values, the most likely state is one in which customs vary from one region to another based on the average soil quality in each region.

What is the evidence that social norms are playing a role in the choice of share contracts, as opposed to a common unobserved factor? This issue bedevils all of the studies that rely on cross-sectional data. Complete identification in such cases is extremely difficult, and essentially impossible if one relies only on linear regression models (Manski 1993; Moffitt 2001; Brock & Durlauf 2001b). However there are several features of the data that strongly suggest that social norms are present. First, the distribution is concentrated on a small number of inherently focal fractions. If this concentration were due to a common factor, such as the reservation wage in a given region, one would expect to see a smooth distribution around each of these fractions, but this is not supported by the data. Second, using cross-sectional data one can test for compression: if the same share applies to farms of different quality, there should be an upward bias in the net income of tenants who work on high quality land. This prediction is confirmed by the Illinois data (Burke 2004). It is of course possible that the market equilibrates through assortative matching rather than through variation in contract terms, that is, high quality farms attract high quality tenants. If this is the case, total output per acre on high quality farms should be higher than is predicted by their soil quality index, which was calibrated using a fixed level of capital and labor inputs. The data do not support this hypothesis either: it appears that tenants on high quality land are able to capture a portion of the land rent without a concomitant increase in labor productivity (Burke 2004).

The analytical framework described above is quite general and could be applied to many other types of contractual relationships -- such as real estate commissions, lawyers' contingency fees, and hedge fund managers' bonuses -- where norms may well play a role in determining contract terms.

8. Other cases

This section provides a brief overview of four case studies of norm dynamics that further illustrate how the evolutionary theory of norms can be brought to bear on empirical cases. They differ in the types of data they use – historical, cross-sectional, experimental – and in the phenomena that they emphasize. Mackie's study of footbinding and infibulation focuses on the types of interventions that have successfully induced norm shifts, as well as top-down attempts that have failed to induce such shifts. Jindani (2014) proposes an evolutionary model of dueling showing how norms respond to changes in objective costs (in this case mortality rates) and changing social attitudes that weaken the force of sanctions. Munshi & Myaux (2006) examine the effect of group norms on women's willingness to practice contraception in rural India. They attempt to control for unobservable common factors by studying the difference in the rate of contraception between Hindus and Muslims who live in the same villages. The fourth case is concerned with regional differences in standards of medical practice in the United States (Wennberg & Gittelsohn 1973, 1982; Burke et al. 2010; Chandra & Staiger 2007). Although these differences are substantial, there is an on-going debate about whether they result from different norms that become established in different medical communities, or whether they result from local productivity spillover effects.

Both mechanisms produce similar dynamic feedback effects, but the implications for welfare are quite different.

8.1. Footbinding

Footbinding is an ancient Chinese practice in which a young girl's toes are bent backwards toward the heel and her feet are wrapped in tight bandages to prevent them from growing to normal size.[22] After 5-10 years of painful treatment the result is a pair of tiny bowed feet that are only three to four inches long. A woman with bound feet cannot engage in most forms of manual labor and cannot walk very far. The custom appears to have originated nearly a thousand years ago under the Sung dynasty (960-1279). It was initially applied to concubines in the imperial court, it then spread to the upper classes as a sign of gentility. By the Ming Dynasty (1368-1644) it had become common practice except among the lower classes where women were needed as laborers. In addition to being a sign of gentility, the practice was thought to promote female fidelity, because a woman with bound feet was also housebound. From the standpoint of the girl's family, however, the key consideration was that it enhanced her marriage prospects. Footbinding was the equilibrium of a game in which boys' families preferred them to marry girls with bound feet, and the girls' families dared not give up the practice for fear that they would not find good husbands.

This social norm remained in place for over a thousand years, yet it disappeared almost entirely within a single generation. Mackie (1996) argues that the norm unraveled due to a deliberate campaign by reformers to eradicate the practice. A key part of the story is that reformers did not rely on top-down edicts prohibiting the practice; these had been tried repeatedly and had come to naught. Instead they organized 'natural foot societies' in which families pledged not to bind their daughters and not to allow their

[22] This discussion is based on Mackie (1996).

sons to marry bound women. In addition they conducted public campaigns explaining the adverse consequences for health, mobility, and employability. In other words, they recognized that the key problem was to simultaneously shift the expectations of a group of interacting families, so that it would become rational for them not to continue the practice given its harmful effects and its decreased benefit in the local marriage market. Mackie argues that this approach could serve as a template for displacing other harmful norms, including the practice of female genital cutting in sub-Saharan Africa.

8.2. Dueling

From the Renaissance to the nineteenth century dueling was a common practice among the upper classes in Europe (Nye 1993; Hopton 2007). A gentleman risked losing his honor if he failed to challenge someone for making offensive statements, or if he failed to accept such a challenge. Thus both parties were willing to risk their lives rather than face the prospective damage to their reputations by violating the norm. It is a stark example of how an extremely costly norm can be held in place by social pressure. Although dueling was practiced initially by the nobility, by the nineteenth century it was also common among members of the professional classes, especially politicians and journalists, who were very much in the public eye.

The norm of dueling is particularly interesting for several reasons. First, it is a highly complex norm with many interlocking parts, each of which can be construed as a subnorm operating within a larger normative framework. Once a challenge had been made, seconds were named whose function was first to try to reconcile the parties. Failing that they negotiated a meeting time and place, the choice of weapons, and various other conditions (such as the number of paces if pistols were used). Second, the practice persisted despite repeated attempts to ban it. According to Hopton (2007) there were nine separate attempts to enforce a ban on dueling in France during the nineteenth century, none of which was successful. Thus it is an excellent example of how difficult it can be to change a norm – even a very costly one – from the top down. Indeed the very people who were best positioned to exercise leadership – legislators and opinion makers – were also highly susceptible to public pressure and had much to lose by attempting to change the status quo.

Third, the costliness of the norm (i.e., the fatality rate) was affected by changes in weapons technology. Before 1500 it was customary in England and France to use slashing broadswords and to protect oneself with bucklers. Although the swords caused injuries, they were mostly superficial flesh wounds and the fatality rate was quite low. By the late 1500s, broadswords had been replaced by rapiers, which were more deadly, and the fatality rate increased substantially. Then in the eighteenth century rapiers were replaced by pistols, which further increased the fatality rate.

A particularly interesting feature of the dynamics is that in France the norm adjusted by a return to rapiers together with rules of engagement that limited the risk of serious injury, whereas pistols continued to be the weapon of choice in England and the United States. Jindani (2014) proposes an evolutionary model in which norm shifts are driven in part by changes in their *objective cost*, and in part by broader changes in society that determine the *social cost* of flouting the norm. The model suggests that the differential evolution of the norm in England and France can be attributed to differences in objective costs (lower fatality rates in France), and to differences in social attitudes, which in France continued to support the gentleman's code of honor that was a vestige of the ancien regime (Nye 1993, Hopton 2007).

8.3. Fertility

Birthrates in developing countries have been steadily declining due to the greater availability of contraceptives, family planning clinics, and enhanced educational and economic opportunities for women. These changes can be understood as a rational response by individuals to the perceived benefits and costs of contraception. Nevertheless the timing and pace of the fertility transition in different communities has led many demographers to conclude that other factors are at work, including differences in social norms (Bongaarts and Watkins 1996; Montgomery & Casterline 1996; Entwhisle et al. 1996; Kohler 1997, 2000, 2001; Kohler et al. 2001; Behrman et al. 2002; Billari et al. 2009). In practice however it is extremely difficult to disentangle the impact of community norms from other factors such as unobserved common effects, learning from others, and pure imitation.

Here I shall briefly outline the results of several recent papers that address these issues. For example, Munshi & Myaux (2006) examine the rate of contraceptive use by groups that are located within the same community and that have similar socioeconomic characteristics, but that differ in their religious affiliation. The data report contraceptive use over an eleven-year period by women in 70 villages in rural Bangladesh, where each village is composed of significant numbers of both Muslims and Hindus. Munshi & Myaux regress each individual's decision to practice contraception as a function of: i) the proportion currently using contraception in their *own* religious group, and ii) the proportion currently using contraception in the *other* religious group, controlling for individual characteristics such as age, education level, spouse's occupation, spouse's education level, and household assets. Their key finding is that an individual's contraception decision is positively associated with its prevalence within her own religious group, but there is no statistically significant relationship with its prevalence in the other group.

Although this result is consistent with a social norms explanation there are other plausible possibilities. For example, the absence of cross-group effects could arise from a social learning process in which women gain information about the benefits and costs from the experiences of other women in their social network. Even though family planning clinics and contraceptives were available to all women in a given village at the same time, there could be differences in the rate of uptake due to the stochastic nature of the adoption process, which may "take off" at different times in different groups due purely to random effects.

Kohler et al. (2001) suggest a way of distinguishing between the 'social learning' and alternative mechanisms by examining the effect of *network structure* on the rate of contraceptive use. They define a woman's network as the set of women with whom she informally discusses family planning issues. The *density* of the individual's network is the proportion of links that members have with each other. Thus if a network contains n individuals, the maximum possible number of links between them is $n(n - 1)/2$. Kohler et al. hypothesize that a sparse network conveys more information about the benefits and costs of contraception than does a dense network (conditional on the proportion of users being the same in the two cases), because the absence of links in the sparse network reduces the redundancy of the information being provided. They further conjecture that a dense network provides greater scope for applying social pressure.

These hypotheses lead to two predictions: i) women in sparse networks will be more likely to act on a given amount of information (as measured by the proportion of adopters in the network), and ii) women in dense networks will show a greater tendency to conform to the dominant practice in their group (the compression effect).

The authors test these predictions using a detailed survey of some 500 women living in rural Kenya. They find that there is a significant difference in behavior between regions with isolated villages and low market activity (Owich, Kawadhgone, Wakula South) and another region (Obisa) where women have access to a large and active market. Figure 5 shows the probability that a given woman is an adopter conditional on the proportion in her network who are adopters and on the density of the network. In all regions, sparser networks are associated with higher probabilities of adoption when the proportion of adopters is less than 2/3. This is consistent with the first hypothesis, namely, that a sparse network provides more independent information than does a dense network. In the nonmarket regions, however, a reversal occurs when the proportion of adopters exceeds a critical threshold that is slightly above 2/3 (the estimated value is 71%). Above this value, higher densities *increase* the probability that a given individual will adopt. In other words, dense networks initially inhibit the adoption of contraception, but beyond a certain threshold they solidify support for it.

Of course, this analysis does not rule out other factors, such as social learning, nor does it rule out the possibility that dense networks result from a tendency of individuals to interact with others like themselves (homophily). Stronger inferences could be drawn from longitudinal data that provide the timing of individual adoption decisions conditional on adoptions by members of one's social network and controlling for individual fixed effects, but such data are difficult to obtain on a large scale.[23]

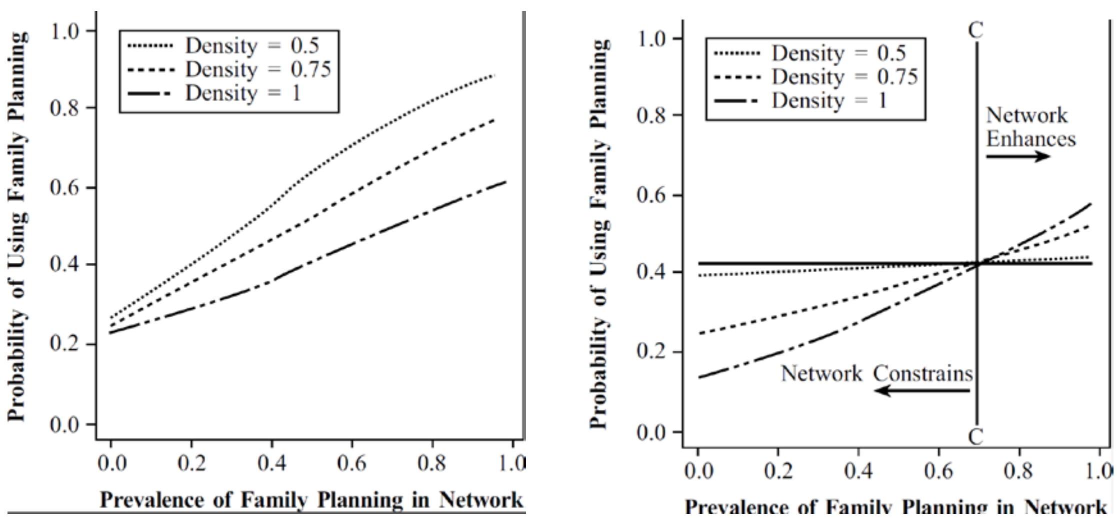

Figure 5. Effect of contraceptive prevalence on the probability of adoption as a function of network densities. Left panel: Obisa. Right panel: Owich, Kawadhgone, Wakula South

[23] Behrman et al. (2002) analyze longitudinal data on contraceptive use in Kenya. They show that social interactions matter after controlling for individual fixed effects, but they do not have network density data as in the study by Kohler et al. (2001). General methods for distinguishing between social learning and social conformity are discussed in Young (2009) and Brown (2013).

8.4. Medical treatment

A long-standing puzzle for public health researchers in the US is that treatments for certain medical conditions vary substantially between states, and even between counties within the *same* state. Differences have been documented for a wide variety of treatments, including Caeserian sections, tonsillectomies, beta-blockers, and stents (Wennberg & Gittelsohn 1973, 1982; Chandra & Staiger 2007; Burke et al. 2010). These variations remain substantial even after controlling for differences in the availability and cost of medical technologies as well as other factors that might predict differences in treatment intensity (Phelps & Mooney 1993).

It has been suggested that these local norms of medical practice are due in part to peer effects (Burke et al. 2007). Physicians tend to respond more readily to practice recommendations made by local "opinion leaders" than to impersonal recommendations such as professional practice guidelines and results of clinical trials (see also Soumerai et al. 1998; Bhandari et al. 2003). Burke et al. (2010) incorporate these peer effects into a dynamic model of treatment choice as follows. They assume that a physician's choice of treatment for a given patient is guided in part by an assessment of its efficacy for patients of that type (depending on age, sex, medical condition) and in part by its frequency of use by others in the local medical community. In this framework, regional treatment norms emerge. Moreover the norm in each region will typically reflect the (objectively) best treatment for the dominant patient type in that region. For example, if treatment A is better suited to older patients than treatment B, one would expect to see A as the norm in regions with a high percentage of older people, and B in regions with a high proportion of younger people. The logic is analogous to the case of contractual norms discussed in section 7: the normal share for the tenant is higher in regions where the land quality is poor on average, and lower where the land quality is high on average. Similarly in the case of medical treatments, old people who live in a young region may not receive the most appropriate treatment (and vice versa). This is the compression effect: the norm engenders excess uniformity in treatment.

It should be stressed that local conformity does not necessarily lead to lower welfare. In the case of medical practice, positive feedbacks may result from learning and productivity spillovers, in which case the welfare effects of local treatment norms may be positive (Chandra & Staiger 2007; Wennberg & Gittelsohn 1973, 1982). As physicians become particularly adept in the dominant local treatment, some patients, already well-suited to the treatment, benefit from their doctors' added expertise. However, patients who are better suited to an alternative treatment face the prospect of receiving either a substandard application of that alternative or an expert application of the ill-suited, dominant treatment. Unless these patients can be transferred to a location where there are experts in the alternative treatment, as a second-best they are better off receiving the dominant local treatment.

More generally, this case illustrates the importance of disentangling the various channels through which positive feedback effects may be operating. It would not be surprising if both social conformity and productivity spillovers are factors in the evolution of regional norms of medical practice. A challenge for future research is to tease out the relative strength of these two effects, and to examine their welfare implications.

9. Directions for future research

My aim in this article has been to outline the main ideas in evolutionary game theory and how they relate to the evolution of social norms. This theory provides a comprehensive framework for analyzing the dynamics that result when individuals make choices based on conventional economic factors as well as on the norms of behavior within their social group. The dynamics exhibit certain features that are not captured by conventional equilibrium models that omit social feedback effects. One of these is *persistence*, that is, the tendency of a norm to stay in place in spite of exogenous changes in incentives. A second key feature is *local conformity/global diversity*: similar communities may exhibit quite different norms due to chance events. Norms of fertility and norms of medical practice offer concrete illustrations of this phenomenon. Third, the very existence of a norm diminishes the variation in behaviors that would otherwise arise. This *compression* effect can be tested with sufficiently rich cross-sectional data, as we saw in the case of agricultural contracts. A similar analysis could be carried out for other types of contracts based on customary percentages, such as the fees earned by fund managers, real estate agents, and tort lawyers.

The theory shows how social norms can coalesce from out-of-equilibrium conditions with no top-down direction. It also demonstrates that the outcome need not be optimal: evolutionary forces can lead to dysfunctional norms that persist for long periods of time, a prediction that is confirmed by historical evidence. On a more positive note, the theory suggests that certain kinds of interventions will be more effective than others for promoting norm shifts. Traditional instruments, such as the blanket application of taxes or subsidies, are often insufficient to overcome the social penalties for violating a prevailing norm. A more effective strategy is to alter the behaviors of a few key actors in small close-knit groups, thus leveraging the positive feedback effects from social interactions.

Although the cases provide considerable support for the predictions of theory, they also point to areas where the theory needs further refinement. For example, current models generally assume that the error rates are held at a low and constant level. Experiments suggest that, in fact, error rates change endogenously depending on how successful people are in coordinating, and are quite high initially. These effects – large and endogenously changing error rates – have important implications for the rate at which the evolutionary process approaches equilibrium and for the equilibria that are most likely to be selected. Another shortcoming of current theory is that it focuses almost exclusively on long-run behavior (for which powerful selection results exist) and tends to neglect the intermediate run dynamics which may be of even greater interest.

I have selected a few cases illustrating how evolutionary game theory can be brought to bear on empirical and experimental evidence, but there are many other examples. These include norms stigmatizing welfare and unemployment (Lindbeck et al. 1999; Brügger et al. 2010), norms stigmatizing out-of-wedlock births (Akerlof et al. 1996), norms of revenge (Elster 1990), norms of tax compliance (Wenzel 2004, 2005; Nicolaides 2012), corruption (Fisman & Miguel 2006); veiling (Carvalho 2013), littering (Cialdini et al. 1990), executive compensation (Brown & Young 2014), body weight (Burke & Heiland 2007; Burke, Heiland & Nadler 2010; Hammond & Ornstein 2014), number of children (Leifbroer & Billari 2009), property rights (Platteau 2000; Bowles & Choi 2013), and marriage (Kanazawa & Still 2001).

Acknowledgments. I thank Sam Bowles, Mary Burke, Jean-Paul Carvalho, Damon Centola, Edoardo Gallo, and Sam Jindani for very helpful comments on an earlier draft.

References

Acemoglu D, Jackson MO. 2012. History, Expectations and Leadership in the Evolution of Social Norms. Working Paper, Dept. Economics, MIT.

Acemoglu D, Robinson JA. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.

Akerlof GA. 1980. A theory of social custom, of which unemployment may be one consequence. Quarterly Journal of Economics 94: 749-775.

Akerlof GA, Yellen JL, Katz L. 1996. An analysis of out-of-wedlock childbearing in the United States. Quarterly Journal of Economics 111: 277–317.

Axelrod R. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod R. 1986. An evolutionary approach to norms. American Political Science Review 80: 1095-1111.

Axtell R, Epstein JM. 1999. Coordination in transient social networks: an agent-based computational model of the timing of retirement. In *Behavioral Dimensions of Retirement Economics*, ed. HJ Aaron. Washington DC: The Brookings Institution.

Banfield EC. 1958. *The Moral Basis of a Backward Society*. Chicago: The Free Press.

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

Bardhan P, Rudra A. 1980. Terms and conditions of sharecropping contracts: an analysis of village survey data in India. Journal of Development Studies 16: 287-302.

Barnes CA. 1983. *Journey from Jim Crow: The Desegregation of Southern Transit.* New York: Columbia University Press.

Baronchelli A, Loreto V, Steels L. 2008. In-depth analysis of the Naming Game dynamics: the homogeneous mixing case. International Journal of Modern Physics C 19: 785.

Behrman JR, Kohler H-P, Watkins SC. 2002. Social networks and changes in contraceptive use over time: evidence from a longitudinal study in rural Kenya. Demography 39: 713-738.

Belloc M, Bowles S. 2013. The persistence of inferior cultural-institutional conventions. American Economic Review: Papers and Proceedings 103: 1-7.

Belloc M, Bowles S.2014. Cultural-institutional persistence under autarchy, international trade, and factor mobility. Working Paper, Santa Fe Institute.

Bernheim D. 1994. A theory of conformity. Journal of Political Economy 102: 841-877.

Bhandari M, Devereaux PJ, Swiontkowski MF et al. 2003. A randomized trial of opinion leader endorsement in a survey of orthopedic surgeons: effect on primary response rates. International Journal of Epidemiology 32: 634-36.

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

Billari F, Philipov D, Testa MR. 2009. Attitudes, norms and perceived behavioural control: explaining fertility intentions in Bulgaria. European Journal of Population.

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

Binmore K. 2005. *Natural Justice*. Oxford UK: Oxford University Press.

Binmore K, Samuelson L, Young HP. 2003. Equilibrium selection in bargaining models. Games and Economic Behavior 45: 296-328.

Blume LE. 1995. The statistical mechanics of best-response strategy revision. Games and Economic Behavior 11: 111-45.

Blume LE, Durlauf, SN. 2001. The interactions-based approach to socioeconomic behavior. In *Social Dynamics*, chapter 2, ed. SN Durlauf, HP Young, pp. 15-44. Cambridge: MIT Press and Washington DC: The Brookings Institution Press.

Blume LE, Durlauf, SN. 2003. Equilibrium concepts for social interaction models. International Game Theory Review 5: 193-209.

Bolton P, Dewatripont M. 2005. *Contract Theory*. Cambridge MA: MIT Press.

Bongaarts J, Watkins SC. 1996. Social interactions and the contemporary fertility transitions. Population and Development Review 22: 639-682.

Bowles S. 2004. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton NJ: Princeton University Press.

Bowles S, Choi J-K. 2013. Coevolution of farming and private property during the Early Holocene. Proceedings of the National Academy of Sciences 110: 8830-8735.

Bowles S, Hwang H-S, Naidu S. 2010. Evolutionary bargaining with intentional idiosyncratic play. Economics Letters 109: 31-33.

Boyd R, Richerson PJ. 2002. Group beneficial norms can spread rapidly in a structured population. Journal of Theoretical Biology 215, 287-296.

Brock, William and Steven N. Durlauf. 2001b. Discrete choice with social interactions. Review of Economic Studies 68: 235-260.

Brown LM. 2013. The diffusion of innovations: empirical and experimental evidence. DPhil dissertation, University of Oxford.

Brown LM, Young HP. 2014. The diffusion of a social innovation: executive stock options from 1936-2005. Working Paper, Dept Economics, University of Oxford.

Brügger, B, Lalive R, Zweimüller J. 2009. Does culture affect unemployment? Evidence from the Röstigraben. Discussion Paper 4283. IZA, Bonn.

Burke MA. 2004. On the distributional effects of contractual norms: the case of cropshare contracts. Working Paper, Dept Economics, Florida State University.

Burke MA, Fournier GM, Prasad K. 2007. The Diffusion of a Medical Innovation: Is Success in the Stars? Southern Economic Journal 73: 588-603.

Burke, MA, Fournier GM, Prasad K. 2010. Geographic Variations in a Model of Physician Treatment Choice with Social Interactions. Journal of Economic Behavior and Organization 73: 418-32.

Burke MA, Heiland F. 2007. Social dynamics of obesity. Economic Inquiry 45: 571-591.

Burke MA, Heiland F, Nadler, CM. 2010. From 'overweight' to 'about right': evidence of a generational shift in body weight norms. Obesity 18: 1226-34.

Burke MA, Young, HP. Social norms. In *Handbook of Social Economics, vol 1A*, eds. A Bisin, J Benhabib, MO Jackson. Amsterdam: Elsevier.

Burtless G. 2006. Social norms, rules of thumb, and retirement: evidence for rationality in retirement planning. In *Social Structures, Aging, and Self-Regulation in the Elderly*, eds. KW Schaie, LL Carstensen. New York: Springer.

Carvalho J-P. 2013. Veiling. *Quarterly Journal of Economics* 128: 337-370.

Centola D, Baronchelli A. 2014. The emergence of social conventions: an experimental study. Work. Pap. Annenberg School of Communication, Univ. Pennsylvania.

Chandra A, Staiger D. 2007. "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks." *Journal of Political Economy* 115(1): 103-140.

Cheung SNS. 1969. *The Theory of Share Tenancy*. Chicago: University of Chicago Press.

Cialdini R, Kallgren C, Reno R. 1990. A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior. Advances in Experimental Social Psychology 24: 201-234.

Cohen WJ. 1957. *Retirement Policies Under Social Security*. Berkeley: University of California Press.

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

Coleman JS. 1990. *Foundations of Social Theory*. Cambridge MA: Harvard University Press.

Durlauf SN. 1997. "Statistical mechanical approaches to socioeconomic behavior," In *The Economy as a Complex Evolving System, vol. II,* eds. WB Arthur, SN Durlauf, D Lane. Menlo Park CA: Addison-Wesley.

Durlauf SN, Young HP, eds. 2001. *Social Dynamics.* Cambridge MA: MIT Press and Washington DC: The Brookings Institution Press.

Edgerton RB. 1992. *Sick Societies: Challenging the Myth of Primitive Harmony*. New York: The Free Press.

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

Ellickson RC. 2001. The evolution of social norms: a perspective from the legal academy. In *Social Norms*, ed. M Hechter, K-D Opp, pp. 35-75. New York: Russell Sage.

Ellison G. 1993. Learning, local interaction and coordination. Econometrica 61: 1047-1071.

Ellison G. 1997. Learning from personal experience: one rational guy and the justification of myopia. Games and Economic Behavior 19: 180-210.

Ellison G. 2000. Basins of attraction, long run stochastic stability and the speed of step-by-step evolution. Review of Economic Studies 67: 17-45.

Ellison G, Fudenberg D, Imhof LA. 2014. Fast convergence in evolutionary models: a Lyapunov approach. Working Paper, Department of Economics, Harvard University.

Elster J. 1989a. Social norms and economic theory. Journal of Economic Perspectives 3: 99-117.

Elster J. 1989b. *The Cement of Society*. Cambridge UK: Cambridge University Press.

Elster J. 1990. Norms of revenge. Ethics 100: 862-85.

Entwhisle B, Rindfuss RD, Guilkey A, Chamratrithirong A, Curran SR, Sawangdee Y. 1996. Community and contraceptive choice in rural Thailand: a case study of Nang Rong. Demography 33: 1-11.

Fehr E, Fischbacher U. 2004a. Third party punishment and social norms. Evolution and Human Behavior 25: 63-87.

Fehr E, Fischbacher U. 2004b. Social norms and human cooperation. Trends in Cognitive Sciences 8: 185-190.

Fehr E, Fischbacher U, Gächter S. 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. Human Nature 13, 1-25.

Fisher LM, Yavas A. 2010. A case for percentage commission contracts: the impact of a "race" among agents. Journal of Real Estate Finance 40:1-13.

Fisman R, Miguel E. 2006. Cultures of corruption: evidence from diplomatic parking tickets. NBER Working Paper 12312. Cambridge MA.

Foster DP, Young HP. 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38: 219-232.

Foster DP, Young HP. 1990. Cooperation in the short and in the long run. Games and Economic Behavior 3: 145-156.

Freidlin M, Wentzell A. 1984. *Random Perturbations of Dynamical Systems*. Berlin: Springer-Verlag.

Gallo E. 2014. Communication networks in markets. Working Papers in Economics 1431, University of Cambridge.

Gambetta D. 2009. *Codes of the Underworld: How Criminals Communicate.* Princeton: Princeton University Press.

Gigerenzer G, Todd PM, ABC Research Group. 1999. *Simple Heuristics That Make Us Smart.* Oxford: Oxford University Press.

Gottlieb S, McCarter R, Vogel R. 1998. "Effect of Beta-Blockade on Mortality among high-risk and low-risk patients after myocardial infarction." *New England Journal of Medicine* 339 (August 20): 489-497.

Guiso L, Sapienza P, Zingales L. 2009. Long term persistence. NBER Working Paper 14278, Cambridge MA.

Hammond R, Ornstein J. 2014. A model of social influence on body weight. Annals of the New York Academy of Sciences, forthcoming.

Hechter M, Opp K-D, eds. 2001. *Social Norms.* New York: Russell Sage Foundation.

Hofbauer J, Sigmund K. 1998. *Evolutionary Games and Population Dynamics*. Cambridge and New York: Cambridge University Press.

Homans GC. 1950. *The Human Group*. New York: Harcourt Brace

Hopton R. 2007. *Pistols at Dawn: A History of Duelling*. London: Portrait.

Hwang S-H, Naidu S, Bowles S. 2014. Social conflict and the evolution of unequal conventions. Working Paper.

Hume D. 1739. *A Treatise of Human Nature*: *Being an Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects*. London: John Noon.

Hurkens S. 1995. Learning by forgetful players. Games and Economic Behavior 11: 304-329.

Illinois Cooperative Extension Service. 1995. Cooperative Extension Service Farm Leasing Survey. Dept of Agricultural and Consumer Economics, Cooperative Extension Service, University of Illinois Champaign-Urbana.

Jencks SF et al. 2000. "Quality of Medical Care Delivered to Medicare Beneficiaries: A Profile at State and National Levels." *Journal of the American Medical Association* 284(13), 1670-1676.

Jindani S. 2014. A game-theoretic approach to dueling and other social norms. MPhil Thesis, Dept. Economics, Oxford University.

Kahneman D, Tversky A, Slovic P, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge UK: Cambridge University Press.

Kanazawa S, Still MC. 2006. The emergence of marriage norms: an evolutionary psychological perspective. In *Social Norms*, ed. M. Hechter, K-D Opp, pp. 274-304. New York: Russell Sage Foundation.

Kandori M, Mailath G, Rob R. 1993. Learning, mutation, and long-run equilibria in games. Econometrica 61: 29-56.

Kandori M. 1992. Social norms and community enforcement. Review of Economic Studies 59: 63-80.

Kohler H-P. 1997. Learning in social networks and contraceptive choice. Demography 34:369-383.

Kohler H-P. 2000. Fertility decline as a coordination problem. Journal of Development Economics 63: 231-163.

Kohler H-P, Bereman JR,Watkins SC. 2001. The density of social networks and fertility decisions: evidence from Sounth Nyanza District, Kenya. *Demography* 38: 43-58.

Kreindler GE, Young HP. 2013. Fast convergence in evolutionary equilibrium selection. *Games and Economic Behavior* 80: 39-67.

Kreindler GE, Young HP. 2014. Rapid innovation diffusion in social networks. *Proceedings of the National Academy of Sciences*, forthcoming.

Kritzer HM. 2004. *Risks, Reputations, and Rewards: Contingency Fee Legal Practice in the United States*. Palo Alto: Stanford University Press.

Leifbroer AC, Billari FC. 2009. Bringing norms back in: a theoretical and empirical discussion of their importance for understanding demographic behavior. Population, Space, and Place. DOI : 10.1002/psp.

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

Lindbeck A, Nyberg S, Weibull, J. 1999. Social norms and economic incentives in the welfare state. Quarterly Journal of Economics 114: 1-35.

Mackie G. 1996. Ending footbinding and infibulation: a convention account. American Sociological Review 61: 999-1017.

Mallaby S. 2010. *More Money than God*. London: Bloomsbury Publishing.

Manski C. 1993. Identification of endogenous social effects: the reflection problem. Review of Economic Studies 60: 531-542.

McFadden D. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, ed. P Zarembka. New York: Academic Press.

Mill JS. 1848. *Principles of Political Economy*. London: Longmans Green.

Moffitt RA. 2001. Policy interventions, low-level equilibria, and social interactions. In *Social Dynamics*, chap. 3, ed. SN Durlauf & HP Young. Cambridge: MIT Press and Washington DC: The Brookings Institution.

Monderer D, Shapley LS. 1996. Potential games. Games and Economic Behavior 14:124-43.

Montgomery MR, Casterline JB. 1996. Social influence, social learning, and new models of fertility. Population and Development Review 22: 151 -175.

Myers RJ. 1973. Bismarck and the retirement age. Letters, The Actuary, April 1973.

Munshi K, Myaux J. 2006. Social norms and the fertility transition. J. Develop. Economics 80: 1-38.

Nash J. 1950. The bargaining problem. Econometrica 18: 155-62.

Nicolaides P. 2012. Personal values and social norms in tax compliance: a model of dynamic belief formation in a network of individual taxpayers. MPhil Thesis, Dept of Economics, Oxford University.

Nye RA. 1993. *Masculinity and Codes of Male Honor in Modern France*. Oxford: Oxford University Press.

Opp K-D. 1982. The evolutionary emergence of norms. British Journal of Social Psychology 21: 139-49.

Phelps CE, Mooney C. 1993. Variations in medical practice use: causes and consequences. In *Competitive Approaches to Health Care Reform*, ed. RJ Arnould, RF Rich, W White. Washington, DC: The Urban Institute Press.

Platteau J-P. 2000. *Institutions, Social Norms, and Economic Development*. London and New York: Routledge.

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

Putnam RD. 1993. *Making Democracy Work – Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

Robertson AF. 1987. *The Dynamics of Productive Relationships: African Share Contracts in Comparative Perspective*. Cambridge: Cambridge University Press.

Saez-Marti M, Weibull JW. 1999. Clever agents in Young's bargaining model. Journal of Economic Theory 86: 268-279.

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

Sandholm W. 2010. *Population Games and Evolutionary Dynamics*. Cambridge: MIT Press.

Schelling TC. 1960. *The Strategy of Conflict.* Cambridge MA: Harvard University Press.

Schelling TC. 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.

Skyrms B. 1996. *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Skyrms B. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.

Smith A. 1776. *An Inquiry Into the Nature and Causes of the Wealth of Nations*. London: Strahan and Cadell.

Soumerai SB et al. 1998. Effect of Local Medical Opinion Leaders on Quality of Care for Acute Myocardial Infarction: A Randomized Controlled Trial. *Journal of the American Medical Association* 279(17), 1358-1363.

Stiglitz JE. 1974. Incentives and risk sharing in sharecropping. Review of Economic Studies 41: 219-55.

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

Sunstein CR. 1996. Social norms and social roles. Columbia Law Review 96: 903-68.

Tabellini G. 2008. Institutions and culture. Journal of the European Economic Association 6: 255-294.

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

Vega-Redondo F. 1996. *Evolution, Games, and Economic Behavior*. New York: Oxford University Press.

Wennberg J, Gittelsohn A. 1973. Small area variations in health care delivery. Science 182: 1102-1108.

Wennberg J, Gittelsohn A. 1982. Variations in Medical Care Among Small Areas. Scientific American 182: 1102-1108.

Wenzel M. 2004. An analysis of norm processes in tax compliance. Journal of Economic Psychology 25: 213-228.

Wenzel M. 2005. Motivation or rationalization?: causal relations between ethics, norms, and tax compliance. Journal of Economic Psychology 26: 491-508.

Young HP. 1993a. The evolution of conventions. Econometrica 61: 57-84.

Young HP. 1993b. An evolutionary model of bargaining. Journal of Economic Theory 59: 145-168.

Young HP. 1995. The economics of convention. Journal of Economic Perspectives 10:105-22.

Young HP. 1998a. *Individual Strategy and Social Structure*: *An Evolutionary Theory of Institutions*. Princeton NJ: Princeton University Press.

Young, HP. 1998b. Conventional contracts. Review of Economic Studies 65: 773-792.

Young HP. 2009. Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. American Economic Review 99: 1899-1924.

Young HP, Burke MA. 2001. Competition and custom in economic contracts: a case study of Illinois agriculture. American Economic Review 91: 559-573.