

Ensminger, Jean and Joseph Henrich
2014 *Experimenting with Social Norms: Fairness and Punishment in Cross-Cultural Perspective*. Jean Ensminger and Joseph Henrich, eds. New York: Russell Sage Foundation.

Theoretical Foundations: The Coevolution of Social Norms, Intrinsic Motivation, Markets, and the Institutions of Complex Societies

Joseph Henrich and Jean Ensminger

Classical scholars long ago proposed a positive relationship between developed market economies and prosocial or fair-minded motivations in impersonal interactions (for an overview, see Hirschman 1982). One of the first and best-known scholars to write in this vein was none other than Adam Smith (1759/2000), whose position was consistent with the findings we present in this volume. Even before Smith, however, Montesquieu (1749/1900, 319) was explicit on this subject: “The spirit of trade produces in the mind of a man a certain sense of exact justice, opposite, on the one hand, to robbery, and on the other to those moral virtues which forbid our always adhering rigidly to the rules of private interest, and suffer us to neglect this for the advantage of others.”

Within the context of societal evolution, Montesquieu anticipated the specifics of the trade-off between narrow economic self-interest and other-regarding behavior that we examine throughout this volume. We compare societies with more and less commerce (market integration), and he anticipated the direction of our findings: there is more fair-minded behavior in more market-integrated societies. In the same vein, David Hume (1751/2006, 25) described the early social evolution of prosocial norms and institutions and anticipated the relationship of prosocial behavior and scale that we address in this chapter.

Suppose that several families unite together into one society, which is totally disjoined from all others, the rules, which preserve peace and order, enlarge themselves to the utmost extent of that society; but becoming then entirely useless, lose their force when carried one step farther. But again suppose, that several distinct societies maintain a kind of intercourse for mutual convenience and advantage, the boundaries of justice still grow larger, in proportion to the largeness of men’s views, and the force of their mutual connexions.

Though prescient, the early classical scholars did not have the benefit of quantitative evidence or comparative experimental measures and game theory to support their intuitions and casual observations. Further, predicting or observing a relationship is not the same thing as positing mechanisms of origin and maintenance. In this chapter, we outline the broader theoretical significance of our empirical project for understanding the coevolution of the particular social norms and institutions that eventually paved the way for the emergence of large-scale societies and massive economic growth. We begin developing a theory of social norms and institutions by

drawing on converging lines of theoretical work arising from both evolutionary and economic frameworks. In developing this foundation, we consider evidence from sources as diverse as developmental psychology, social neuroscience, economic history, and anthropological ethnography. Building on this, we discuss how prosocial or group-beneficial norms arise and spread, and how this relates to the emergence of the more centralized and formal institutions that undergird the subsequent expansion of large-scale, complex societies.

THE EVOLUTION OF NORMS AND INSTITUTIONS

As empirical phenomena, anthropologists, sociologists, and others from across the social sciences have long noted the importance of social norms and institutions, often evoking norms or institutions to explain behavioral similarities within groups or to differentiate groups (Bendor and Swistak 2001; Bicchieri 2006; Sripada and Stich 2006). Despite broad usage and their seeming empirical importance, both the concept of social norms and its interrelationship with institutions have until recently lacked sufficient micro-level theoretical foundations to be taken seriously by researchers in the economic and evolutionary sciences.

In building a theoretical foundation, let's begin with a distinction between norms and institutions. *Norms* are mental representations stored in individual brains that got there through some form of learning, broadly defined (that is, they are not innate). Conceptually, depending on one's preference and disciplinary background, norms could be composed of a combination of preferences and beliefs, mental models (or scripts and schema) and motivations, or decision rules and expectations. In general, these all aim to include (1) what people believe others will do in some context and (2) what they think they and others *ought* to do, as well as (3) varying degrees of internalized motivations (including none) to meet those expectations and to see others meet those expectations.

Institutions are emergent phenomena that arise at the population or group level from the individuals' interactions, decisions, and learning. They are first and foremost self-reinforcing, dynamically stable equilibria that arise as individuals' norms converge and complement each other over time. We will use the term "formal institutions" to label those institutions that are reified and enforced, or reinforced, by written laws, police, and external sanctioning mechanisms. This allows us to distinguish cases in which individuals' norms (for example, their expectations and motivations) do not match the local institutions, or in which the local (informal) institutions do not match the formal prescriptions of extrinsically supplied formal institutions (based on written laws).

Without at least plausible answers to key theoretical questions regarding how social norms and institutions emerge, why individuals adopt norms that violate their narrow economic self-interest, what "adopt" means, how individuals' decisions interact with group patterns, how norms and institutions spread across groups, and how and why norms and institutions have changed over human history, neither social norms nor institutions can be fully incorporated into either economic or evolutionary frameworks. In recent decades, however, approaches arising from both evolutionary biology and economics are converging on a "ground-up" or "first principles" explanation of social norms and institutions. There are now plausible theoretical answers to the above questions, and in some cases there is an abundance of such proposals (see, for example, Chudek and Henrich 2010).

Learning Social Norms

The first step in approaching norms from an evolutionary perspective is to use the logic of natural selection, aided by formal evolutionary modeling, to consider what kinds of learning strategies or heuristics individuals—be they humans, guppies, or rats (Galef and Whiskin 2008b;

Lachlan, Crooks, and Laland 1998)—would evolve to adapt to uncertain, novel, or changing environments (Boyd and Richerson 1985; Schlag 1999; Wakano and Aoki 2006), including environments with social interactions (Guzmán, Rodríguez-Sickert, and Rowthorn 2007). These environments include social interactions in which information about the costs and benefits of alternative behaviors is costly, inherently uncertain, incomplete, or impossible to acquire. In such environments, these learning strategies, which include heuristics that integrate rules glossed as "copy the most successful" (prestige-biased transmission) and "copy the most common trait" (conformist transmission), can outcompete learning strategies that rely solely on the direct evaluation of perceived costs and benefits or innately specified repertoires (see, for example, Boyd and Richerson 1988; Henrich and Boyd 1998; Henrich and Gil-White 2001b; Schlag 1998; Wakano, Aoki, and Feldman 2004). This means that the direct evaluation of the costs and benefits of alternative actions is but one component in a suite of adaptive learning tools that permit individuals to calibrate to diverse informational environments. In this suite, direct cost-benefit evaluation is deployed and operates best in stable, well-structured (clear choices) environments with high informational content.

This line of evolutionary theorizing has led to many predictions about the kinds of learning rules or biases that individuals should use in calibrating their behavior to local environments. A substantial amount of evidence from both laboratory experiments and field observations supports these evolutionary predictions in humans—both in adults (Coultas 2004; Efferson, Lalive, and Fehr 2008; Henrich and Broesch 2011; Henrich and Gil-White 2001a; Henrich and Henrich 2007, ch. 2; Kameda and Nakanishi 2002; Kohler, VanBuskirk, and Rusavage-Barz 2004; McElreath et al. 2005; McElreath et al. 2008; Mesoudi 2009) and in children (Birch, Akmal, and Frampton 2010; Birch, Vauthier, and Bloom 2008; Corriveau, Fusaro, and Harris 2009; Corriveau and Harris 2009a, 2009b; Chudek et al. 2013; Jaswal and Neely 2006)—and in other species (Day et al. 2001; Galef and Laland 2005; Galef and Whiskin 2008a; Lachlan et al. 1998).

The second step in building an evolutionary theory of norms and institutions is to explore what happens when individuals equipped with these evolved learning strategies are placed in a formal (evolutionary) game theoretical model that permits different kinds of social interactions. The results of many such efforts indicate that stable equilibrium states at the population level often emerge when interacting individuals are deploying adaptive learning strategies. The result is a stable behavioral pattern for a given group. Moreover, such analyses often show that there are many different stable equilibrium states for a given situation. Which state emerges in a particular population is dependent on the initial conditions and details of particular shocks along the evolutionary path. In general, the existence and potential emergence of different stable states (that is, different group-level behavioral regularities) is amazingly robust across a variety of types of social interactions and different kinds of adaptive learning rules. The results begin to look like emerging social norms and institutions, at least in that these empirically grounded learning heuristics give rise to stable statistical regularities and behavioral prescriptions that vary across social groups.¹

Matters get even more interesting when one considers social interactions in which an individual has the possibility of exploiting others, such as in a public goods situation or other larger-scale cooperative dilemmas. One might expect free-riding strategies to always dominate, but evolutionary learning models show that if individuals are using empirically grounded adaptive learning heuristics (prestige-biased and/or conformist transmission), then prosocial or group-beneficial norms of cooperation, exchange, and respect for others' property can be maintained, even when interacting groups are large and the possibility of repeated interactions is low. There appear to be a number of ways to sustain such individually costly (and group-beneficial) equilibria in a manner that solves the well-known second-, third-, and *n*th-order free-rider problems. Two

important solutions for our purposes involve (1) permitting individuals to learn strategies that punish or reward non-prosocial behavior (for example, selfish behavior), along with strategies that cooperate or not (Boyd and Richerson 1992b; Henrich and Boyd 2001; Kendal, Feldman, and Aoki 2006); or (2) linking via reputation any norm-violating behavior (individually costly behavior) to other types of social interaction (Panchanathan and Boyd 2004). Other solutions exploit signaling opportunities (Gintis, Smith, and Bowles 2001; Boyd, Gintis, and Bowles 2010), or combinations of cultural learning biases (Henrich 2009a).

One problem with these “prosocial solutions” is that the same mechanisms can stabilize any equally costly behavior, *independent* of its benefit to the group. Such mechanisms can stabilize behaviors in a group that hurt both the individuals themselves and the average payoffs of the group. This robust aspect of these mathematical analyses is problematic if one is looking exclusively to solve the dilemma of large-scale cooperation or similar social dilemmas, since the cooperative solution is merely one among myriad others. For building a theory of norms, however, this oddity is a feature, not a bug, because it better reflects the world we actually observe.

In other words, these emergent phenomena are now looking even more like the norms and institutions recorded by social scientists: we have behavioral regularities stabilized by either direct punishment or indirect sanctioning (through reputational damage) that can be, but need not be, prosocial or group-beneficial. Anthropologists and sociologists have recorded a wide range of stable behavioral patterns that appear to be maladaptive (Diamond 2005; Edgerton 1992) but seem also to be enforced by concerns about reputational damage or informal punishment. Conspicuous examples include food taboos, Chinese foot-binding (Mackie 1996), supercision of adolescent males, cranial deformation in infants, ritual cannibalism of relatives that sustain prion diseases (Durham 1991), and female infibulation (Knight and Ensminger 1998). It is interesting that the same mechanisms that can theoretically account for the normative and self-reinforcing nature of these practices are the mechanisms that can explain cooperation in large groups, including management of property rights (Ostrom 1990) and economic exchanges among strangers in the absence of formal third-party enforcement (Ensminger 1992).

The Internalization of Norms

There is at least one major empirical aspect of norms that is missing so far from our “ground-up” development: the emotional, or motivational, aspect (Chudek and Henrich 2010; Sripada and Stich 2006). The desire to adhere to norms and to see them enforced appears to be internally motivated in some fashion. Once internalized, norms become ultimate ends, goals, or values in themselves. Economists would say these individuals have put the performance of the norm into their objective function (Greif 2006). By noting this, we emphasize that internalizing a norm does not make an individual a mindless “norm-executer.” People have plenty of competing internalized goals and motivations that demand choosing among alternative goals. Moreover, failure to comply with a norm on a particular occasion does not indicate a lack of any internalized motivations.

Why would natural selection build an organism that internalizes social norms as proximate motivations (Gintis 2004, 2007)? To approach this issue, the evolutionary analysis should focus on the costs of processing information and making errors (being sanctioned for violations or simply miscoordinating), the temptation of reaping immediate rewards, and the developmental circumstances of the adapting child. Natural selection may favor internalizing norms as ends in themselves if this saves on information-processing costs or reduces the associated errors. If norms are reliable and frequently sanctioned, an individual might be better off to “just do it” in

many situations, motivated by norm adherence as a goal in itself or by a pre-evaluation decision rule that sometimes skips cost-benefit evaluations. “Just doing it” would thereby save the costs of performing for each slightly different situation a cost-benefit analysis that would have to include considering the probability of being judged in violation of a norm and the resulting long-term reputational damage or punishment. Moreover, suppose that each time one runs such mental calculations, actors pay costs to acquire information and occasionally make processing errors. Internal motivation or pre-evaluation decision rules can save the costs of information search and help one avoid the errors in inherently noisy environments.

In addition, internally motivated adherence to norms or pre-evaluation decision rules may be natural selection’s way of psychologically overcoming the pull of an immediate reward versus long-term benefits. If people overweight immediate rewards compared to future rewards or rewards amortized over years, as many researchers argue they do (Berns, Laibson, and Loewenstein 2007; Frederick, Loewenstein, and O’Donoghue 2002; Laibson 1997), then internal motivation might provide that extra push to forgo the short-term gains in favor of the long-term payoffs. Natural selection or adaptive learning could fix this problem by adjusting our temporal discounting, but since many nonhuman animals have the same problem, we might assume that such temporal discounting biases either provide other benefits that preserve them or are subject to biological constraints that make it too costly to build a better discounter.

Finally, when organisms live in rapidly changing or variable environments, allowing proximate motivations to be internalized by learning can help the organism make adaptive decisions. To understand this, consider the acquisition of social norms from the perspective of children. Since norms and institutions vary across groups but are generally locally stable, children can adapt to the local environment by first rapidly adopting and partially internalizing the local norms (pre-evaluation decision rule), thereby avoiding sanctions, and then begin evaluating the costs and benefits of norm violations once they have mastered the specifics of local norms, such as when the norm applies, how angry people get at violations, how much people monitor, and so on. The problem of errors in cost-benefit calculations is particularly acute for children since they lack the information possessed by adults to evaluate the consequences of violations, and many social rules protect individuals from environmental or social hazards (such as toxic substances, dangerous animals, or diseases).

Developmental work suggests that by adolescence, and often long before, children have much knowledge and have internalized many local norms. Children first acquire local norms in contextually specific circumstances and then rigidly apply them, before gradually, with observation and experience, adjusting the norms’ domain of applicability. Later, having mastered the knowledge of a norm but perhaps not yet having *fully* internalized it, children start figuring out how and when they can violate it for their own benefit (Fiske 1998; Harbaugh, Krause, and Liday 2002; Henrich and Henrich 2007, ch. 2; Lancy 1996; Sutter and Kocher 2007).

One further perspective on how norms are learned and internalized deserves attention. The creation of explicit rules, formal institutions, and laws, as we discuss later for more complex societies, helps to coordinate people’s beliefs about what norms are applicable to particular circumstances. Formalized rules also convey information about the incentives for adherence (sanctions); by influencing compliance, they may also influence the acquisition and internalization of learners trying to figure out what the local norms are (Cooter, Feldman, and Feldman 2008). Once internalized, social emotions (such as guilt, shame, and pride) can motivate norm adherence (Chudek and Henrich 2010; Chudek, Zhao, and Henrich, 2013).

The natural implication of this for our project is that the more highly developed these systems are, the better should be people’s predictions about the behavior of others, and thus the lower should be the variance in compliance; the result is a virtuous cycle (literally, in the case

of prosocial norms) of norm internalization and compliance. Iris Bohnet and Robert Cooter (2003, 19) summarize how this notion has been debated by legal scholars:

Our results bear on a disagreement among legal scholars about how law causes social change, such as the decrease in racial discrimination in the U.S. Some scholars argue that law has an “expressive function” that changes behavior, whereas other scholars deny that law has much influence on such phenomena as racial discrimination. Our research suggests that law changes society by changing beliefs more than preferences. According to this logic, laws imposing desegregation in the southern states may have changed behavior by changing beliefs about the willingness of others to integrate. Whereas beliefs changed relatively quickly, preferences probably changed relatively slowly.

Richard McAdams (2000, 1651) has also stressed the unique impact of law as a focal point for belief about how others will behave; as he puts it, “law works by what it says in addition to what it does.” The emphasis here is on coordination via focal points (and information about what others believe and prefer). In this case, even sanctionless laws can serve this function by signaling information. The notion that social norms change behavior by changing expectations regarding the behavior of others and evoking context-specific preferences is also supported by the literature on framing effects in economic experiments (Camerer 2003; Ross and Ward 1996), which we discuss in more detail later in the chapter.

Experimental Support for Learning and Internalization of Norms

This theoretical foundation for social norms permits us to unify a diverse range of empirical findings. Here we briefly summarize four supporting lines of experimental work. First, much work from developmental psychology shows that prosocial norms are transmitted culturally, at least in the laboratory: children acquire individually costly prosocial behaviors via observational learning and automatically adopt sanctioning tendencies toward violators. Second, prosocial behavior toward strangers, as measured in economic experiments in the United States and Europe, develops gradually over the first two or three decades of life, unlike motivations and behavior related to prosociality toward kin and in dyadic reciprocity situations. Third, behavior in experiments depends on population-specific contextual cues that aid individuals in mapping the games to their local norms. And finally, economic games measure internalized norms: subjects’ prosocial behavior (cooperation, fairness, and punishment) in experimental games in the West activates the same reward circuits as does receiving direct cash payments; being prosocial in games is internally rewarding in the brain, even when costly.

Norms are transmitted culturally. Substantial research with children in the 1960s and 1970s shows that context-specific, prosocial norms for altruistic behavior toward strangers can be acquired by observing others behaving altruistically in the same context. When exposed to either adult or peer models who donated more tokens (which could purchase toys) to poor children, subjects between five and eleven years of age also donated more tokens themselves. These effects were (1) not influenced by whether the child was alone when making the donation decision, and (2) not ephemeral, as they endured in re-tests months later that did not include observations of models. The altruistic effects extended to somewhat similar tasks, but did not generalize beyond this. When placed in the role of model, child subjects continued to donate higher amounts and also spontaneously scolded younger children who did not donate sufficiently; this scolding behavior had not been previously modeled in the experiment. Recent work shows that even toddlers

readily turn single observations of behavior into social rules and spontaneously sanction norm-violators (Rakoczy, Warneken, and Tomasello 2008; Rakoczy et al. 2010).

A comparison of methods reveals that exposing children to charitable models is the best method to induce charitable giving toward strangers, superior to rewards, exhortations, and direct instruction. Preaching and verbal instruction have little effect unless accompanied by costly donating actions by an adult or peer model (Henrich 2009a). Incentive schemes can induce prosocial behavior as long as the incentives persist, but they do not create internalized motivation. (In fact, they seem to inhibit internalization.) Natalie Henrich and Joseph Henrich (2007, ch. 2) detail these findings and provide references.

Economic games with children demonstrate that prosociality develops slowly over two or more decades. In experimental games similar to those used in this volume, research shows that Western children’s offers begin to approach adult levels by around age twelve; then there is a drop or plateau during adolescence, when they deviate only somewhat from the fairness norm. In the ultimatum game, in which the responder can punish a low offer, giving lower offers is rationally self-interested because adolescents’ willingness to punish (reject) has not risen sufficiently high to discourage unfairness, as it will in adulthood (Ensminger and Cook, chapter 18, this volume, available at: <http://www.russellsage.org/Ensminger>; see also Henrich and Henrich 2007, ch. 8). Experiments done over the developmental trajectory for both Americans and Europeans indicate that prosocial behavior does not reach its adult plateau until the mid or late twenties (Harbaugh et al. 2002; Sutter and Kocher 2007; Fehr, Bernhard, and Rockenbach 2008). Moreover, the behavior of university students in experimental games continues to change from the first to the fourth year (Carter and Irons 1991), and in some experiments, such as the dictator game, the differences between students and fully socialized adults is dramatic (Henrich and Henrich 2007; Carpenter, Burks, and Verhoogen 2005). Jean Ensminger and Kathleen Cook (chapter 18, this volume, available at: <http://www.russellsage.org/Ensminger>) demonstrate the same phenomenon for the rural United States.

Norms are context-specific. This approach to norms gives us a means to anticipate and theorize about how different contextual cues in laboratory experiments, which have no direct impact on the payoff structure of the game, influence game play. Subjects arrive at experiments equipped with norms, which include contextually specific beliefs (expectations of others’ behavior) and internalized motivations, and then face a novel situation. They have to figure out how to behave, in part, by figuring out which—if any—of their norms apply to the situation. Since most experimental games involve both money and anonymity, players from some societies with norms that apply to such contexts are influenced by both their norm-related beliefs (what they think others will do) and their internalized motivations, or pre-evaluation decision rules (such as what is “fair”).² These norm-effects come through when experimentalists set up the identical game (with the same payoff structure) but vary the language or some other aspect of the process. Lee Ross and Andrew Ward (1996), for example, used identical versions of a public goods game (a cooperative dilemma) and labeled one the “Wall Street” game and the other the “Community” game. They found that university students in the Wall Street game contributed less than those in the Community game (see also Hoffman et al. 1994; Pillutla and Chen 1999). In a different kind of variation, a variety of experiments show the positive effects on prosocial behavior of communicating before playing (Ostrom, Gardner, and Walker 1994). Communicating, even if not explicitly about the game, can help players anticipate the norm-driven behavior of other players, thus facilitating norm-coordination (Janssen et al. 2010).

Nahoko Hayashi and her colleagues (1999) show that simple framing differences strongly affect rates of cooperation in an otherwise identical two-person prisoner’s dilemma, and that these effects depend on whether the player is from Japan or the United States. This finding fits

with observed differences between the United States and Japan in non-experimental contexts and shows that the same contextual manipulations yield different effects in different populations. (For interactions between contextual cues and societal differences, also see Goerg and Walkowitz 2010).

We emphasize here that we do not think that norms are the only thing influencing play in experimental games, or anywhere else. Other aspects of the games—such as the material costs and benefits, the possibility of cultivating a reputation, and the prospect of repeated interaction—ought to influence game behavior in predictable ways independent of the norms. Additionally, some contextual effects that do not influence the actual payoff structure probably affect game play by influencing players' perceptions of the possibilities for reputation formation or repetition (see, for example, Haley and Fessler 2005). Of course, such effects do not stand as evidence against internalized motivations.

Finally, by living in a group with shared behavioral standards, individuals come to acquire these norms as internalized motivations. Work in neuroscience and neuroeconomics has recently contributed to this line of theorizing by showing that behaving in the manner demanded by norms—cooperating, contributing, or punishing in locally recommended or prescribed ways—activates the brain's rewards or reward anticipation circuits in the same manner as when a direct cash payment is received (Fehr and Camerer 2007). Complying with local norms "feels good" to the brain in the same way that receiving cash does. Cooperating and getting money (from the cooperation) feels better—and activates reward circuits more—than just getting the same amount of money without this association (Rilling et al. 2004). Being fair feels better—and activates more reward circuits—than receiving the same amount of money while being unfair (Tabibnia, Satpute, and Lieberman 2008). Punishing by really hurting norm-violators (physically or monetarily) also activates these reward circuits (Sanfey et al. 2003) more than punishing symbolically (de Quervain et al. 2004). Giving money to charity activates the same reward circuits as receiving money (Harbaugh, Mayr, and Burghart 2007), even when the actor's own actions do not cause the charitable giving. Activations of the brain's reward circuitry in these experiments generally predict behavioral outcomes (Fehr and Camerer 2007; Sanfey 2007).

Convergence with Economics

Evolutionary and economic approaches have now begun to converge on both a unified conception of learning and a theoretical foundation for social norms and institutions. By considering the impact of incomplete information and uncertainty on rational decision-making, combined with the simplicity of ecologically rational heuristics for dealing with complex situations (Brandstatter, Gigerenzer, and Hertwig 2006; Gigerenzer and Selten 2001), bounded rationality is providing an improved understanding of human social behavior paralleling that derived through evolutionary theory (Gintis 2007; Young 1998). Economists have shown that copying successful people and copying the plurality are—under particular conditions—quite rational strategies (Ellison and Fudenberg 1993, 1995; Schlag 1998, 1999; Spencer and Huston 1993), as well as fitness-maximizing. Such strategies are rational when information is costly to acquire or process, or when information about the costs and benefits of alternative behaviors or strategies is noisy (error-ridden)—that is, these strategies are rational in circumstances common to many real-life decisions. Economists have also explored bounded learning strategies based on personal trial and error learning (Fudenberg and Levine 1998).

Theoretical models that place individuals deploying these learning strategies in social interactions yield stable behavioral patterns that look like norms. This is not surprising in some cases, since the underlying learning heuristics used by evolutionary theorists and economists are similar, but in other cases this finding reaffirms the phenomenon that learning plus social interaction

robustly yields a wide variety of fairly stable outcomes (Weibull 1995; Young 1998). This multiplicity of stable outcomes is a feature of classical game theoretical models that assume perfect and free information and processing power, a finding enshrined in the folk theorem (Gintis 2000).

The emerging focus on bounded rationality and game theory has led to some of the best experimental work on learning in social interactions. The experiments confirm that, at least in the laboratory, learners do appear to be using learning heuristics like "copy the successful" and "copy the plurality," as well as following experienced-based learning rules (Apesteguia, Huck, and Oeschssler 2007; Camerer and Ho 2000; Pingle 1995; Pingle and Day 1996; Selten and Apesteguia 2005). Many of these findings converge with work in psychology, using quite different experimental tools, as well as with findings from field observations (see Henrich and Henrich 2007, ch. 2; Mesoudi 2009).

In returning to the inspirations of Adam Smith and others, economists are beginning to explore and theorize about the internalization of norms, or endogenous preference formation, and have specifically considered the effect of markets (Bowles 1998). Rather than waving off the question of where people's preferences come from, an increasing number of economists are examining the possibility that preferences emerge in part from interactions with the local institutional environment. As Avner Greif (2006, 37) puts it, "internalized norms are socially constructed behavioral standards that have been incorporated into one's superego (conscience), thereby influencing behavior by becoming part of one's preferences." Textbooks in microeconomics are also beginning to take the need to consider internalized motivations seriously (Bowles 2004). People's motivations or preferences partially calibrate in adapting to and performing in the local equilibrium, and these calibrations and preferences show durability (Alesina and Fuchs-Schündeln 2007; Francois and van Ypersele 2009; Nunn 2009). This means that history matters.

Scholars in the law and economics tradition have taken a particular interest in the internalization of social norms. Cooter (1996) has repeatedly emphasized that social norms are uniquely distinct from formal laws, in that they are socialized within individuals and thus become self-enforcing. Cooter does not dispute that we often have reputational reasons for abiding by social norms, or that individual members of societies may also impose external sanctions on norm-violators in the form of gossip and even ostracism. But as Cooter notes, to dwell on these aspects of social norms is to miss one of their unique features: we internalize social norms, including a sense of guilt (or other social emotion) should we violate a norm that we believe to be morally just (Cooter 1996, 152). Similarly, as we saw in our discussion of neuroeconomics, people also get "mental rewards" from seeing a norm-violator punished. Thus, social norms internally propel us (at least some of us) to engage in behavior that is not in our narrow economic self-interest. This also means that we may abide by norms even when no one is looking—for instance, in an anonymous economic experiment. Indeed, Cooter (2000, 1581) operationalizes a definition of the degree of internalization of a norm as the price one is prepared to pay to conform to it, which dovetails nicely with economic experiments that are designed to measure exactly that cost dimension of behavior.

Although economics does not generally concern itself with ultimate explanations of motivations or preferences, some efforts within economics have paralleled and even inspired some of the evolutionary explanations discussed here based on error management. If norm violations result in sanctions or miscoordination, individuals should develop internalized motivations (preferences) that allow them to avoid norm violations that will cost them in the long run (Frank 1988).

Multiple Mechanisms of Norm Stabilization

As noted earlier, there is a growing list of ways to stabilize costly norms and institutions, including group-beneficial ones. Earlier models solved the free-rider problem in larger-scale cooperative dilemmas by permitting the transmission (learning) of both cooperative and punishing

strategies. The higher-order free-rider problem created by the cost of punishing was solved by a meta-punishing strategy (punish all nonpunishers) or by learners using a combination of learning strategies under uncertainty, including some conformist learning (Axelrod 1986; Boyd and Richerson 1992a; Henrich and Boyd 2001; Kendal et al. 2006). This approach can have direct observers of violations do the punishing and spread reputational information about violations to create broader diffuse punishment, or it can allow those best positioned to punish to administer the sanctions. Thus, these approaches can work in the absence of reliable reputational information if direct observers of norm violation perform the sanctioning (Henrich and Henrich 2007). Herbert Gintis, Eric Alden Smith, and Samuel Bowles (2001) and Robert Boyd, Gintis, and Bowles (2010) provide models that explore how signaling might sustain costly punishment and how this might in turn sustain costly norms.

More recent models illustrate alternative routes to stable norms that do not involve costly punishment; these routes involve linking costly individual actions in larger-scale situations (for example, doing something costly to contribute to a group benefit) to a two-person interaction in a different social context. Karthic Panchanathan and Robert Boyd (2004), for example, show how costly norms can be stabilized by attaching players' reputations in a dyadic helping game to their reputations in a larger-scale, individually costly interaction. If an individual fails to "cooperate" (perform the costly action) in the larger interaction, he gets a "bad reputation," and other individuals can withdraw their help from him in the two-person game *without* getting a bad reputation themselves. Otherwise, individuals who refuse to help those with good reputations in the two-person helping game get a bad reputation and lose the help of others in the two-person game. There is no free-rider problem here because individuals "sanction" by withdrawing help, thus avoiding the costs of delivering help. Straightforward, narrow self-interest can drive sanctioning, so there is no costly punishment. Maciej Chudek and Joseph Henrich (2010) provide a more detailed discussion of various models that do not involve diffuse costly punishment.

These mechanisms provide different routes to stable social norms and institutions. As we argue in chapter 4, however, unlike models involving direct costly punishment, they require well-functioning, high-fidelity, reputational systems. Therefore, while such sanctioning regimes can explain some norms in some groups, including prosocial norms, they cannot support cooperation and fair exchange with those outside the reputational system—for example, in interactions in larger populations with strangers.

Cooter (1997) and Fehr and Gintis (2007) have argued that internalized norms make people more willing to directly punish norm-violators at a cost because they believe the normative behavior is the right thing to do. As we shall see in chapter 4, our findings suggest that this intuition applies only to norms sustained by diffuse costly punishment mechanisms and not to other kinds of sanctioning systems. We return to this debate in light of our findings and consider the circumstances in which the claim holds empirically. People internalize, in some sense, the local institutional forms.

THE SPREAD OF PROSOCIAL NORMS AND INSTITUTIONS ACROSS POPULATIONS

Once a combination of expectations, motivations, and beliefs converges in a group to create an institution, we have a somewhat sticky situation. When different societies, or groups, converge on different social norms, owing to the aforementioned path dependence or historical specificity of the process, is that the end? This problem is made even more poignant because, as we have already argued, models of norms indicate that many different norms can be stable and that most of these are not prosocial or group-beneficial. So what we have is a bunch of different

groups, each with different norms, only a few of which involve any prosociality among strangers. Selecting among these norms becomes the classic problem of *equilibrium selection*, an important challenge that emerges in both dynamic evolutionary approaches (Henrich 2006; Samuelson 1998) and those rooted in classical concepts of rationality (Harsanyi and Selten 1988).

At least three broad theoretical approaches confront the problem of equilibrium selection. We label these: (1) stochastic stability, (2) forward-looking decision-making, and (3) cultural group selection. The first approach is based on the stochasticity inherent in any real population (Young 1998). Different stable equilibria (institutions) are more or less susceptible to this stochasticity, meaning that in the long run some equilibria will be substantially more common than others because some institutions will be more likely to collapse and cause the group to evolve to a different institution. Over time, differences in the sizes of the basins of attraction among equilibria will gradually lead groups to aggregate at the equilibrium with larger basins of attraction. This force is most important in relatively small groups, as they are more susceptible to such stochasticity (Kendal et al. 2006).

The second theoretical approach, and perhaps the most intuitive, is that rational, forward-looking individuals recognize the long-term payoffs available at stable cooperative equilibria, assume that others are similarly sensible, and choose the prosocial state (see, for example, Harsanyi and Selton 1988). The main problem with considering this mechanism as a dominant force is that as one looks across time and space, the world and human history are full of non-prosocial, and even downright antisocial, norms and institutions that hurt the group as a whole (Edgerton 1992). Nevertheless, these three equilibrium selection mechanisms are not mutually exclusive, so this kind of mechanism is likely to be part of the story. Groups sometimes change norms quite consciously by meeting and reaching consensus, although actual cases suggest that they consciously adopt the norms of other more successful groups (Boyd 2001), making this a form of cultural group selection. Moreover, there is a tendency to focus on group decisions that yield beneficial outcomes as examples of foresight, but when they are placed in context with all the bad group-level decisions, luck rather than foresight seems the more probable cause.

The third mechanism, cultural group selection, results from competition among societies at different stable equilibria. Where population pressure and intergroup competition is absent, we see many examples of norms and institutions sustaining non-prosocial behavior. Where intergroup competition is strong, we see the spread of norm-bearers and/or practices and beliefs from groups stabilized at equilibria that favor success in intergroup competition, which includes institutions and norms that sustain large-scale cooperation, in-group harmony, and fairness among ephemeral interactants. In humans, competition between groups can take the form of warfare, demographic success, biased migration, or more subtle forms in which individuals learn decisions and strategies by observing and copying higher-payoff individuals—some of whom are from groups with higher-payoff norms and institutions (Fehr and Fischbacher 2003; Henrich 2004). This between-group learning can lead to a differential flow of decisions, strategies, and even preferences from higher- to lower-payoff groups (Boyd and Richerson 2002; Henrich and Boyd 2001), or it can lead to differential migration (Boyd and Richerson 1990), favoring the spread of the high-payoff norms (Boyd and Richerson 2009).

There is both laboratory and field evidence supporting cultural group selection. In the laboratory, Özgür Güreker, Bernd Irlenbusch, and Bettina Rockenbach (2006) permitted players to choose between one of two different "institutions." In the first institution, players could contribute money to a group project. All contributions were increased and divided equally among all players, regardless of their contributions. Previous experiments (Fehr and Fischbacher 2003) have established that when this interaction is repeated, average contributions to the public good drop to near-zero (a "noncooperative equilibrium"). The other "punishing" institution is very

similar, except now, after players have contributed, they can pay to punish (reduce the payoff) or reward other players. When this interaction is played repeatedly (Fehr and Gächter 2000), a substantial fraction of players punish low contributors, causing mean contributions to rise and to stabilize near full cooperation (a “cooperative equilibrium”) among samples of Western undergraduates (Herrmann, Thoni, and Gächter 2008). Both of these laboratory institutions were run concurrently for thirty interactions, and both initially and after each subsequent interaction (after seeing others’ payoffs), players could choose their institution for the next interaction.

The principal findings of Güerker and his colleagues (2006) can be summarized simply: initially most players picked the institution *without* sanctioning possibilities. But in response to being exploited by free-riders, cooperators in the nonpunishing institution began to reduce their contributions, and that began to drive total contributions toward zero. Meanwhile, punishers in the sanctioning institution started driving contributions up, despite the personal cost of punishing. After a few interactions, players from the nonsanctioning institution—presumably seeing the higher payoffs of those choosing the sanctioning institution—increasingly switched institutions. Despite the incoming flow of migrants from the nonsanctioning institution, the mean contributions in the sanctioning institution consistently increased or held stable near full cooperation. In fact, most incoming migrants, consistent with local norms in their new setting, increased their contributions during their first interaction in the sanctioning institution, and a majority administered some costly punishment.

What does this tell us about equilibrium selection? First, the student subjects’ expectations of others’ behavior did not permit them to foresee the final outcome and select the higher-payoff institution on the first interaction (Ertan, Page, and Putterman 2009). This occurred despite the simplicity of these experiments compared to the real world. Most players selected the lower-payoff institution, perhaps out of distaste for the possibility of being punished. Second, despite the stochasticity of human decisions, neither institution drifted into another equilibrium. There were only thirty interactions in this game, so one could hardly expect stochastic processes to begin selecting equilibria. But that is the point: observational learning and cultural group selection across institutions occur much faster than stochastic processes. What did happen is that once players from the lower-payoff institution observed the higher payoffs of the other institution, they wanted to adopt the practices of that institution, or the decisions and strategies of those other players. In this experiment, players could do that only by “migrating” to the other institution. These migrants, however, did not appear to be merely uninformed payoff-maximizers who needed to adjust their beliefs about others. A majority of migrants into the sanctioning institution not only cooperated in their new institution but also punished (cooperating but not punishing is the payoff-maximizing strategy). Formal evolutionary modeling of the influence of success on migration and the spread of group-beneficial equilibria converges with these experimental findings (Boyd and Richerson 2009).

Outside the laboratory, there are now many lines of empirical evidence to support cultural group selection, including data from archaeology, history, and ethnography. For example, using detailed quantitative ethnographic data, Scott Atran and his colleagues (2002) have shown how conservation-oriented ecological beliefs spread from locally prestigious Itza Maya to Ladinos in Guatemala, and how highland Q’eqchi’ Maya, with tightly bound cooperative institutions and commercially oriented economic production, are spreading at the expense of both Itza and Ladinos. In New Guinea, Joseph Soltis, Robert Boyd, and Peter Richerson (1995) have shown that even the slowest forms of cultural group selection (conquest) can occur on five-hundred- to one-thousand-year time scales. In Africa, using detailed ethnohistorical data, Raymond Kelly (1985) has demonstrated that differences in cultural practice regarding bride-price fueled the

Nuer expansion over the Dinka, and that different social institutions, underpinned by cultural beliefs about segmentary lineages, provided a decisive competitive advantage. Similarly, Marshall Sahlins (1961) has argued that cultural beliefs in segmentary lineages, which facilitated both the Nuer and Tiv expansions, have spread this social institution in different parts of Africa. At the global level, Jared Diamond (1997) has made a cultural group selection case for the European expansion after AD 1500, as well as for the Bantu, Chinese, and Austronesian expansions. Using archaeological data and cultural phylogenetics, anthropologists are increasingly arguing for the importance of cultural group selection in prehistory (Currie and Mace 2009; Flannery and Marcus 2000; Spencer and Redmond 2001), including competition among foragers (Bettinger and Baumhoff 1982; Young and Bettinger 1992).

Recent work suggests that religions and rituals that galvanize group solidarity and deepen communities’ mutual commitments can spread by cultural group selection (Henrich 2009a; Atran and Henrich 2010). For example, Ensminger (1997b) examines the spread of Islam in Africa. Islam is known to have spread particularly fast along the long-distance trade routes of Africa and South Asia. Conversion facilitated participation in trading networks (the group benefits). Although people may have originally been attracted to the economic benefits of conversion, there is no doubt that they also fully internalized the norms and belief systems of the religion, which worked to everyone’s advantage as they forged honest trading partnerships in which contracts were honored, thus vastly expanding the Islamic trade routes and simultaneously lowering transaction costs for all.

Up to this point, we have been largely concerned with the results of decentralized actions by individuals and the learning mechanisms they employ in acquiring norms and adapting to their social environment. In the historical record, of course, we also have many instances of important changes in social norms and institutions that were driven by bigger players possessing substantial economic or political power.

It is important to recognize a continuum between the most informal and the most formal of institutions—those with written laws, formal adjudication procedures, and so on—and to consider how they influence equilibrium selection. Societies vary dramatically in the strength and complexity of their centralized decision-making institutions. Such institutions can, by their very nature, influence other institutions and their associated norms. Some groups, like the Machiguenga of the Peruvian Amazon, lack any significant higher-level institutions. Many foragers and horticulturalists have at least a community-level meeting in which individuals seek to build consensus on proposed changes. If consensus is reached, changes can be made, but little compels compliance after the meeting, since achieving consensus does not mean internalizing motivations. We expect all the same factors to be at play at all levels of sociopolitical complexity. The existence of strong central decision-making institutions, however, backed by external sanctions and written laws, can raise the level of agreement, coordination, consensus, and enforcement achieved prior to and during experimentation with novel institutions.

In their writings about the Orma in Kenya, Ensminger and Knight (1997) and Knight and Ensminger (1998) provide examples of how bigger players (elites) can influence institutional shifts in a decentralized society. In these cases the relative bargaining power of the individuals leading the innovation in social norms can make a difference. A change in the norm of clan exogamy (rules mandating marriage outside of one’s clan) spread because of the relative bargaining power of those few individuals initiating the change. The authors argue that people in the society were less likely to sanction norm-violators of greater wealth and status, thus affording such people a greater ability to innovate social norms (either for their personal gain or to suit their values) without being sanctioned. People often did not wish to forgo future interactions (including marriage) with such people and would weigh such action differently than they would

if the violators were less pivotal in the social network. But other efforts by elites to change norms requiring female circumcision failed, owing to poor coordination. Similarly, Mackie's (1996) example of Chinese pledge societies that were involved in foot-binding eradication also involved elites as prime movers.

Moving along the continuum of sociopolitical complexity past typical hunter-gatherer and horticultural societies, we find societies with councils of elders, formal chieftains, and the politics of archaic and modern states. At this end of the spectrum, coordinated and centralized innovation of institutions may occur more regularly. It is here, we argue, that one is most likely to encounter coordinated manipulation both for strategic distributional ends and for the benefit of collective action in the interest of the common good, or success in intergroup competition (which are not always the same thing). However, in noting the importance of decision-making institutions, we do not wish to deny the fact that in many situations, even within the most centralized societies, it is the strength of local norms, not the force of formal institutions or laws, that governs behavior. In the developing world with which we are concerned, the sanctioning power attached to breaches of local social norms may be considerable, while the reach of state institutions is often limited and their legitimacy poorly internalized.

Economic history provides numerous examples of top-down "innovations" in institutional structures that have led to considerable economic prosperity (Greif 2006; North and Weingast 1989). Among ethnographic examples, Ensminger and Knight (1997) have examined the process by which sedentary elites in a herding society managed to engineer a gradual change in the property rights institution from one of common grazing to one involving more restrictive access that eased environmental degradation and favored sedentary elites (for additional examples of property rights change, see also Ensminger 1997a; Ostrom 1990).

Whether such institutional changes favored by elites or central political authorities should be understood as strategic foresight for individual or group gain, the imitation of more successful groups with similar practices—or merely a lucky guess that supplies variation to the engine of cultural group selection—depends on examining particular changes within a broader historical and multi-group context. It is easy to mistake lucky guesses or cross-group imitation for foresight, especially given the unintended consequences of so many well-motivated social engineering efforts (Henrich 2009b).

The converging evolutionary and economic approaches described thus far suggest that different social norms, whether they arise from rational decision-making or some evolved adaptive learning processes, will emerge in different places and contexts, leading members of different groups to calibrate their beliefs and internal motivations differently. If our experiments are measuring norms for dealing with strangers in monetary exchanges, our theorizing here leads us to expect three features. First, we expect variation across populations in both our measures of fairness and punishment for interactions with anonymous others (Henrich et al. 2004). This prediction stands in contrast to approaches that take students' behavior in experimental games to be a robust universal feature of our species and the product of cognitive adaptation favored by repeated interactions (Hoffman, McCabe, and Smith 1998; Nowak, Page, and Sigmund 2000). Second, this variation across populations ought to be bounded by theoretically possible, dynamically stable equilibria, and it should favor equilibria that are group-beneficial. Thus, although variation is expected, it is not that "anything goes" cross-culturally. We do not expect societies with normative expectations that favor, for example, giving one's entire windfall to an anonymous other. Societies tolerate much maladaptive behavior, especially in noncompetitive situations such as those associated with low population pressure, but societies subject to cultural group selection tend to possess prosocial norms that maintain harmony, extend exchange, and sustain large-scale collective action. Third, we expect to observe a relationship across popula-

tions between the social institutions (like markets and religion) and social norms for dealing with strangers, as measured in our games.

In the next section, we argue that a group's social norms influence how markets and other complex institutions operate, especially in large populations. Social groups with norms and institutions that facilitate trust, fairness, and cooperation in contexts involving strangers or in low-frequency interactions succeed in establishing institutions that achieve higher payoffs or are more competitive in intergroup competition. Similarly, larger groups that employ direct costly punishment for violations of such norms are more competitive. These differences in group payoffs or competitiveness favor the coevolution of the associated norms via one of the equilibrium selection mechanisms described here.

MARKETS, COMPLEX SOCIETIES, AND THE NORMS THAT MAKE THEM WORK

There is remarkably little consensus on two different, though inherently related, questions that researchers from diverse disciplines have confronted. The first is traditionally an anthropological question (Diamond 1997; Johnson and Earle 2000; Nolan and Lenski 2004): until about ten thousand years ago, our ancestors lived in relatively small, nomadic, or semi-sedentary, populations dependent on hunting and gathering for subsistence. In the ensuing millennia, sedentary agriculture arose in several places, then larger towns emerged, then cities. The scale and intensity of human cooperation and exchange expanded dramatically during what, in human evolutionary terms, was a relatively short time. How can a species adapted to living in relatively small foraging groups, often dependent primarily on kin relationships, expand the sphere of cooperation and exchange to such an extent in such a short time? And why did this seem to occur at different rates on different continents (Diamond 1997; Richerson and Boyd 1998, 1999; Hibbs and Olsson 2004)? The second question, one of critical humanitarian importance in the modern world and one that has long animated economic thinkers even before Adam Smith took up the challenge, asks: why are some societies rich and others poor?

Respected scholars from a variety of disciplines have diverse perspectives on these critical questions. Hypotheses concerning the role of climate, geography, and factor endowments generally have a long history as explanations of differential development, and they have recently been further elaborated in various ways by Jared Diamond (1997), Jeffrey Sachs (2001, 2003), Louis Putterman (2008), David Landes (1998) and Douglas Hibbs and Ola Olsson (2004). Explanations that emphasize seemingly immutable advantages or obstacles, such as geography, may seem to focus more on the fortuitousness (or not) of inherited circumstances than on that which is humanly engineered, such as the institutions of property rights that many believe undergird and incentivize durable economic growth (North 1981, 1990; North and Thomas 1973). The work of Stanley Engerman and Kenneth Sokoloff (1994) and Daron Acemoglu, Simon Johnson, and James Robinson (2001) adds further nuance by emphasizing the interaction effect between the original endowments and specific institutions.

In contrast to the focus on both factor endowments and institutions, scholars such as Gregory Clark (2007), Jean-Philippe Platteau (2000), Thomas Sowell (1998), Robert Putnam (Putnam, Leonardi, and Nanetti 1993), and Douglass North (2005), as well as modernization theorists of the 1950s and 1960s (for example, McClelland 1961), have stressed the role of cultural practices, beliefs, and values in the process of economic development and wealth generation. The theoretical framework for the evolution of social norms and institutions that we have laid out here is consistent with some of the arguments made in this literature, but we believe that better specification of the mechanisms and substantiation with

experimental and ethnographic evidence alter the specific substantive predictions and clarify this line of theorizing.

One way in which our data bear on the literature relating beliefs and values to economic growth has to do with religion. Religion plays a role in both our theory of the evolution of societal complexity and our data (Atran and Henrich 2010). Anthropologists have long noted a positive relationship between societal complexity and the presence of high moralizing gods. Religions in small-scale societies, especially foragers, often lack the omniscient moralizing gods of world religions (Johnson 2005; Roes 1995; Roes and Raymond 2003), who actively reward and punish proper behavior. For example, the idea that attaining a blissful afterlife could be contingent on proper moral behavior in this life is not found in traditional small-scale societies, and does not emerge in the historical record until 500 B.C. (McNeill 1991). Societies with high moralizing gods have spread dramatically, as have their beliefs, in the last two thousand years. Henrich and his colleagues argue that cultural group selection has favored certain religious beliefs and rituals (Atran and Henrich 2010; Henrich 2009a; Shariff, Norenzayan, and Henrich 2010). Our analyses in this volume indicate, independent of a wide variety of other factors like income, wealth, market integration, and settlement size, a positive relationship between practicing a world religion (Catholicism, Islam, or Protestantism—evangelical and non-evangelical) and prosocial behavior toward anonymous others.³ This is consistent with a variety of findings that have begun to delineate the relationship between faith, world religions, and prosociality (Norenzayan and Shariff 2008).

Another tradition close to what we propose here has focused on the consequences of institutional change as the driver of economic growth and development. In particular, Nathan Rosenberg and L. E. Birdzell (1986), North (1981, 1990), Greif (2006), and Acemoglu, Johnson, and Robinson (2002) have considered the evolution of institutional forms, including the development of property rights, as crucial for understanding differential global development. Acemoglu and his colleagues (2002) have argued that the reversal of economic fortunes in many regions after AD 1500 resulted from an interaction of European institutional forms with well-developed existing indigenous institutions. Where complex institutions already existed, imperialists exploited them to extract wealth. Where they did not exist, European institutions were imported, and these favored economic growth after the Industrial Revolution.

These diverse perspectives on the differences in the fortunes of nations are not mutually exclusive and can be linked by examining the first question—the evolution of societal complexity since the beginning of the Holocene epoch. Consider the origins of agriculture. Certain geographical regions had better combinations of natural endowments in the form of more easily domesticated grains and animals, as well as climates better suited to cultivation, population expansion, and long-term settlement (Diamond 1997). But as empirical data suggest, sustaining relatively stable, harmonious, large, sedentary settlements built around fixed resources (land, water, and pasture) requires at least some agreement on rudimentary property rights, access agreements, storage systems, and defensive collaborations before cereal agriculture can take off (North 1981). It appears likely that at least in some regions the technical know-how of domestication was poised for a takeoff long before the norms and institutions existed that incentivized the costs and risks of planting and nurturing a crop through to harvest. Working out cooperative property rights or defense capabilities may have taken far longer than culturally evolving the technical routines and know-how of stable productive agricultural packages (North 1981; Richerson and Boyd 2000; Richerson, Boyd, and Bettinger 2001). This is an example of the iterative, coevolutionary process that we are describing.

Merely learning to live together in large settlements with significant numbers of nonkin long enough to sow and reap a harvest may have required the development of norms of fairness that facilitated group organization and decision making. Initially, such authority systems may

have been nothing more than diffuse consensus among elders with the legitimacy to settle the disputes and diffuse the strains that arise from cohabitation and divisions of surplus. The development of more elaborate religious beliefs with professional specialists may have facilitated such submission to authority (Wright 2009) and made the potential or imagined consequences of violations more effective. Ethnographically, many groups, including some of the societies we have studied, lack such religious professionals and do not linger in large settlements, as disagreements and dispute cause dissolution. They also do not submit to authorities above the head of household (Henrich 2000; Johnson 2003; Johnson and Earle 2000).

The implementation of more intensive farming practices may have had to wait for the social and institutional systems to catch up (Richerson, Boyd, and Bettinger 2001). The problem is not dissimilar from what we observe in the modern world today. Clearly, there are many societies that are aware of, and sometimes in possession of, the technological capacities of the higher-performing economic systems that they wish to emulate, but have not yet acquired the institutional capacity and cultural capital to do so (Herrmann et al. 2008; Clark 2007).

Expanding the division of labor and trade, sustaining effective political and judicial decision-making, controlling corruption, and maintaining public safety and effective policing all require the evolution of a variety of norms and institutions for interacting prosocially to solve the inherent collective action problems that arise in interacting with strangers or low-frequency partners. As laid out earlier, coevolving norms and institutions provide a menu of potential solutions. Groups, or their norms, tend to spread because they allow larger populations to achieve higher payoffs in the competitive ecologies of other groups. Such norms can, and did, spread, probably by some version of the equilibrium selection mechanisms discussed here. But these processes are often slow, stochastic, and reversible, and they also often lead to the eventual collapse of specific societies (Diamond 2005).

Part of this process involves the emergence of the norms and institutions that permit mutually beneficial market exchanges among strangers or low-frequency partners. Certain forms of market exchange and their associated norms for dealing with strangers coevolve. Groups with norms that allow profitable exchange among ephemeral interactants achieve, on average, greater success than those societies lacking such norms. These market norms can spread as individuals, impressed by its success, seek to join the group and adopt its norms (just as they did in the experiment conducted by Gürer and his colleagues), or as members of other groups adopt the norms of the successful groups and shift the balance of their own group toward new norms and institutions. Successful groups may proliferate demographically, as did early Christians and Muslims (Levy 1957; Stark 1997), or the successful group may apply its excess wealth to military or commercial expansion, as has happened throughout history the world over. At the same time, markets and related institutions developed elsewhere can spread into groups already possessing sufficiently appropriate or applicable norms (even if suboptimal), though markets do not spread readily into groups lacking the appropriate norms without the operation of one of the equilibrium selection mechanisms discussed here. One implication of this reasoning is that more market-integrated societies ought to, *ceteris paribus*, have more prosocial norms for dealing with strangers, anonymous others, and low-frequency exchange partners.

Independent of markets, members of larger populations suffer from the daily need to interact with, cooperate with, not steal from, and exchange with other individuals well beyond their own circle of friends and family. Sustaining larger populations requires similar prosocial norms, though if the population is not too big, reputation can sustain the norms in the absence of direct costly punishment (Panchanathan and Boyd 2004). Thus, we predict a relationship between settlement size and punishment behavior in our experiments. Large populations require costly punishment to sustain prosocial norms, while small populations can use either reputation or

costly punishment. Using classical game theory, Avinash Dixit (2004, 76) has formalized this proposition and argues that it is systems of intermediate size that are at the greatest disadvantage, operating as it were between the advantages of face-to-face direct reputation-based systems and those at the other end of the institutional spectrum with strong mechanisms of third-party enforcement. This is consistent with ethnographic work in New Guinea indicating that communities lacking complex social institutions tend to break up as community size exceeds about three hundred (Forge 1972; Tuzin 2001). Our findings attest to this in the sense that diffuse costly punishment principally occurs in communities of more than twelve hundred people.

MOVING AHEAD

In this chapter, we have laid out a theoretical framework, together with experimental and empirical literature to support it, in which we attempt to explain the mechanisms by which certain kinds of prosocial norms can arise in the simplest of human societies. From there we describe the means by which we believe prosocial behavior can proliferate in a virtuous cycle of reinforcement as institutional structures become more complex. We recognize that multiple equilibria persist, and that many, if not most, are not group-beneficial. But as societies come into competition with one another, those where individuals and groups have drifted or designed their way into equilibria that stabilize higher levels of prosocial behavior have selective advantages over their neighbors. Quite commonly, aspects of successful systems are recognized and copied by neighbors, but never with exactly the same result, given path dependence. The level of economic exchange that is supported by complex societies today has its foundation in the earliest human societies, where some prosocial behavior is also maintained. But as institutions become more complex, they have the capacity to specify, internalize, and enforce higher levels of prosocial behavior that can allow societies to realize the economic benefits of overcoming diverse and substantial collective action challenges.

We believe that the data from this project are consistent with the mechanisms, described here, by which social norms and institutions evolved through human history. Our experiments were designed to test for the existence and strength of social norms for sustaining mutually beneficial exchange relationships, maintaining in-group harmony, and facilitating collective action among unrelated individuals. Our findings from the first phase of this project (Henrich et al. 2004; Henrich et al. 2005a, 2005b), now replicated and extended in this second phase with new experiments, sites, and samples, indicate a positive relationship between prosocial behavior and market integration, as well as a positive relationship between settlement size and the punishment of norm violations.

The sample of societies from which we draw the data for this project is virtually unique in that it runs the gamut from almost pure hunter-gatherers (absent most traces of modern development and material possessions) through numerous horticultural and nomadic herding societies (some equally remote from modern markets), to cash-cropping farmers, urban African workers, and small-town residents in rural America. As such, it offers a rare opportunity to address some core questions about the coevolution through time among social norms governing fair-minded behavior, institutional complexity, and level of market integration.

NOTES

1. This approach also recognizes and explores the influences of “representational content” (the content of what is socially learned) on transmission as well as the impact and importance of inferential processes in understanding imitative learning (Henrich and McElreath 2003). For a debate on the evolutionary foundations of conformist transmission, see Nakahashi (2007), Guzmán et al. (2007), and Nakahashi, Wakano, and Henrich (2012).
2. See Ensminger (2004) for an example of a public goods experiment that did not use any specific framing but nevertheless was immediately labeled the “harambee game” by local participants, who were reminded of the local

institution of public goods provision. Behavior in the game tracked the locally accepted behavioral pattern: the wealthy made higher contributions to the public good, as local social norms dictated for the analogous institution. In contrast, in the dictator game, which triggered no particular local institutional reference, the wealthy did not contribute statistically more than others.

3. Notably, the relationship that we find does not support Weber (1958), who is the best-known proponent of the relationship between religion and commercial development. In our data, Protestants are not more prosocial than others.

REFERENCES

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91(5): 1369–1401.
- . 2002. “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution.” *Quarterly Journal of Economics* 117(4): 1231–94.
- Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. “Good-Bye Lenin (or Not?): The Effect of Communism on People’s Preferences.” *American Economic Review* 97(4): 1507–28.
- Apestequia, Jose, Steffen Huck, and Jorg Oeschler. 2007. “Imitation: Theory and Experimental Evidence.” *Journal of Economic Theory* 136(1): 217–235.
- Atran, Scott, and Joseph Henrich. 2010. “The Evolution of Religion: How Cognitive By-products, Adaptive Learning Heuristics, Ritual Displays, and Group Competition Generate Deep Commitments to Prosocial Religions.” *Biological Theory* 5(1): 1–13.
- Atran, Scott, Douglas L. Medin, Norbert Ross, Elizabeth Lynch, Valentina Vapnarsky, Edilberto Ucan Ek’, John D. Coley, Christopher Timura, and Michael Baran. 2002. “Folkecology, Cultural Epidemiology, and the Spirit of the Commons: A Garden Experiment in the Maya Lowlands, 1991–2001.” *Current Anthropology* 43(3): 421–50.
- Axelrod, Robert. 1986. “An Evolutionary Approach to Norms.” *American Political Science Review* 80(4): 1095–1111.
- Bendor, Jonathan B., and Piotr Swistak. 2001. “The Evolution of Norms.” *American Journal of Sociology* 107(2): 1493–1545.
- Berns, Gregory S., David Laibson, and George Loewenstein. 2007. “Intertemporal Choice: Toward an Integrative Framework.” *Trends in Cognitive Sciences* 11(11): 482–88.
- Bettinger, Robert L., and Martin A. Baumhoff. 1982. “The Numic Spread: Great Basin Cultures in Competition.” *American Antiquity* 47(3): 485–503.
- Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Birch, Susan A. J., Nazanin Akmal, and Kristen L. Frampton. 2010. “Two-Year-Olds Are Vigilant of Others’ Non-verbal Cues to Credibility.” *Developmental Science* 13(2): 363–69.
- Birch, Susan A. J., Sophie A. Vauthier, and Paul Bloom. 2008. “Three- and Four-Year-Olds Spontaneously Use Others’ Past Performance to Guide Their Learning.” *Cognition* 107(3): 1018–34.
- Bohnet, Iris, and Robert D. Cooter. 2003. “Expressive Law: Framing or Equilibrium Selection?” Faculty Research Working Papers Series. Cambridge, Mass.: Harvard University, John F. Kennedy School of Government.
- Bowles, Samuel. 1998. “Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions.” *Journal Economic Literature* 36(1): 75–111.
- . 2004. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton, N.J.: Princeton University Press.
- Boyd, David. 2001. “Life Without Pigs: Recent Subsistence Changes Among the Irakia Awa, Papua New Guinea.” *Human Ecology* 29(3): 259–81.
- Boyd, Robert, Herbert Gintis, and Samuel Bowles. 2010. “Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare.” *Science* 328(5978): 617–20.
- Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- . 1988. “An Evolutionary Model of Social Learning: The Effects of Spatial and Temporal Variation.” In *Social Learning: Psychological and Biological Perspectives*, ed. Thomas R. Zentall and Bennett G. Galef. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- . 1990. “Group Selection Among Alternative Evolutionarily Stable Strategies.” *Journal of Theoretical Biology* 145(3): 331–42.
- . 1992a. “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups.” *Ethology and Sociobiology* 13(3): 171–95.

- . 1992b. "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups." *Ethology and Sociobiology* 13(3): 171–95.
- . 2002. "Group Beneficial Norms Can Spread Rapidly in a Structured Population." *Journal of Theoretical Biology* 215(3): 287–96.
- . 2009. "Voting with Your Feet: Payoff Biased Migration and the Evolution of Group Beneficial Behavior." *Journal of Theoretical Biology* 257(2): 331–39.
- Brandstatter, Eduard, Gerd Gigerenzer, and Ralph Hertwig. 2006. "The Priority Heuristic: Making Choices Without Trade-offs." *Psychological Review* 113(2): 409–32.
- Camerer, Colin. 2003. *Behavior Game Theory: Experiments in Strategic Interaction*. Princeton, N.J.: Princeton University Press.
- Camerer, Colin, and Teck-Hua Ho. 2000. "Experience-Weighted Attraction Learning in Normal Form Games." *Econometrica* 67(4): 827–75.
- Carpenter, Jeffrey, Stephen Burks, and Eric Verhoogen. 2005. "Comparing Students to Workers: The Effects of Social Framing on Behavior in Distribution Games." In *Field Experiments in Economics*, ed. Jeffrey Carpenter, Glenn W. Harrison, and John A. List. Greenwich, Conn.: JAI Press.
- Carter, John R., and Michael D. Irons. 1991. "Are Economists Different, and If So, Why?" *Journal of Economic Perspectives* 5(Spring): 171–77.
- Chudek, Maciej, Patricia Brosseau, Susan Birch, and Joseph Henrich. 2013. "Culture-Gene Coevolutionary Theory and Children's Selective Social Learning." In *The Development of Social Cognition*, ed. Mahzarin R. Banaji and Susan A. Gelman. Oxford: Oxford University Press.
- Chudek, Maciej, and Joseph Henrich. 2010. "Culture-Gene Coevolution, Norm-Psychology, and the Emergence of Human Prosociality." *Trends in Cognitive Sciences* 15(5): 218–26.
- Chudek, Maciej, Wanying Zhao, and Joseph Henrich. 2013. "Culture-Gene Coevolution, Large-Scale Cooperation, and the Shaping of Human Social Psychology." In *Signaling, Commitment, and Emotion*, ed. Richard Joyce, Kim Sterelny, and Brett Calcott. Cambridge, Mass.: MIT Press.
- Clark, Gregory. 2007. *A Farewell to Alms: A Brief Economic History of the World*. The Princeton Economic History of the Western World Series. Princeton, N.J.: Princeton University Press.
- Cooter, Robert D. 1996. "Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant." *University of Pennsylvania Law Review* 144(5): 1643–96.
- . 1997. "Normative Failure Theory of Law." *Cornell Law Review* 82(5): 947–79.
- . 2000. "Do Good Laws Make Good Citizens? An Economic Analysis of Internalized Norms." *Virginia Law Review* 86(8): 1577–1601.
- Cooter, Robert D., Mical Feldman, and Yuval Feldman. 2008. "The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium." *Review of Law and Economics* 4(3): 890–911.
- Corriveau, Kathleen H., Maria Fusaro, and Paul L. Harris. 2009. "Going with the Flow: Preschoolers Prefer Nondissenters as Informants." *Psychological Science* 20(3): 372–77.
- Corriveau, Kathleen, and Paul L. Harris. 2009a. "Choosing Your Informant: Weighing Familiarity and Recent Accuracy." *Developmental Science* 12(3): 426–37.
- . 2009b. "Preschoolers Continue to Trust a More Accurate Informant One Week After Exposure to Accuracy Information." *Developmental Science* 12(1): 188–93.
- Coultas, Julie. 2004. "When in Rome . . . An Evolutionary Perspective on Conformity." *Group Processes and Intergroup Relations* 7(4): 317–31.
- Currie, Thomas E., and Ruth Mace. 2009. "Political Complexity Predicts the Spread of Ethnolinguistic Groups." *Proceedings of the National Academy of Sciences* 106(18): 7339–44.
- Day, Rachel L., Tom MacDonald, Culum Brown, Kevin N. Laland, and Simon M. Reader. 2001. "Interactions Between Shoal Size and Conformity in Guppy Social Foraging." *Animal Behaviour* 62(5): 917–25.
- de Quervain, Dominique J., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305(5688): 1254–58.
- Diamond, Jared M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W. W. Norton & Co.
- . 2005. *Collapse: How Societies Choose to Fail or Succeed*. New York: Viking.
- Dixit, Avinash K. 2004. *Lawlessness and Economics: Alternative Modes of Governance*. Gorman Lectures in Economics. Princeton, N.J.: Princeton University Press.
- Durham, William H. 1991. *Coevolution: Genes, Culture, and Human Diversity*. Stanford, Calif.: Stanford University Press.

- Edgerton, Robert B. 1992. *Sick Societies: Challenging the Myth of Primitive Harmony*. New York: Free Press.
- Efferson, Charles, Rafael Lalive, and Ernst Fehr. 2008. "The Coevolution of Cultural Groups and Ingroup Favoritism." *Science* 321(5897): 1844–49.
- Ellison, Glenn, and Drew Fudenberg. 1993. "Rules of Thumb for Social Learning." *Journal of Political Economy* 101(4): 612–43.
- . 1995. "Word-of-Mouth Communication and Social Learning." *Quarterly Journal of Economics* 110(1): 93–125.
- Engerman, Stanley L., and Kenneth L. Sokoloff. 1994. "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States." Historical Paper 66. Cambridge, Mass.: National Bureau of Economic Research.
- Ensminger, Jean. 1992. *Making a Market: The Institutional Transformation of an African Society*. Cambridge: Cambridge University Press.
- . 1997a. "Changing Property Rights: Reconciling Formal and Informal Rights to Land in Africa." In *The Frontiers of the New Institutional Economics*, ed. John N. Drobak and John V. C. Nye. New York: Academic Press.
- . 1997b. "Transaction Costs and Islam: Explaining Conversion in Africa." *Journal of Institutional and Theoretical Economics (Zeitschrift für die Gesamte Staatswissenschaft)* 153(1): 4–29.
- . 2004. "Market Integration and Fairness: Evidence from Ultimatum, Dictator, and Public Goods Experiments in East Africa." In *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, ed. Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis. New York: Oxford University Press.
- Ensminger, Jean, and Jack Knight. 1997. "Changing Social Norms: Common Property, Bridewealth, and Clan Exogamy." *Current Anthropology* 38(1): 1–24.
- Ertan, Arhan, Talbot Page, and Louis Putterman. Forthcoming. "Who to Punish? Individual Decisions and Majority Rules in Mitigating the Free Rider Problem." *European Economic Review* 53(5): 495–511.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach. 2008. "Egalitarianism in Young Children." *Nature* 454(7208): 1079–83.
- Fehr, Ernst, and Colin F. Camerer. 2007. "Social Neuroeconomics: The Neural Circuitry of Social Preferences." *Trends in Cognitive Sciences* 11(10): 419–27.
- Fehr, Ernst, and Urs Fischbacher. 2003. "The Nature of Human Altruism." *Nature* 425(6960): 785–91.
- Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90(4): 980–95.
- Fehr, Ernst, and Herbert Gintis. 2007. "Human Motivation and Social Cooperation: Experimental and Analytical Foundations." *Annual Review of Sociology* 33: 43–64.
- Fiske, Alan Page. 1998. "Learning a Culture the Way Informants Do: Observing, Imitating, and Participating." Available at: http://www.bec.ucla.edu/papers/learning_culture.htm (accessed October 2013).
- Flannery, Kent V., and Joyce Marcus. 2000. "Formative Mexican Chiefdoms and the Myth of the 'Mother Culture.'" *Journal of Anthropological Archaeology* 19(1): 1–37.
- Forge, Anthony. 1972. "Normative Factors in the Settlement Size of Neolithic Cultivators (New Guinea)." In *Man, Settlement, and Urbanisation*, ed. Peter J. Ucko, Ruth Tringham, and G. W. Dimbelby. London: Duckworth.
- Francois, Patrick, and Tanguy van Ypersele. 2009. "Doux Commerces: Does Market Competition Cause Trust?" CEPR Discussion Paper No. DP7368. London: Center for Economic Policy Research.
- Frank, Robert. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton & Co.
- Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* 40(2): 351–401.
- Fudenberg, Drew, and David Levine. 1998. *The Theory of Learning in Games*. Cambridge, Mass.: MIT Press.
- Galef, Bennett G., and Kevin N. Laland. 2005. "Social Learning in Animals: Empirical Studies and Theoretical Models." *Bioscience* 55(6): 489–99.
- Galef, Bennett G., and Elaine E. Whiskin. 2008a. "'Conformity' in Norway Rats?" *Animal Behaviour* 75(6): 2035–39.
- . 2008b. "Use of Social Information by Sodium- and Protein-Deficient Rats: Test of a Prediction (Boyd & Richerson 1988)." *Animal Behaviour* 75(2): 627–30.
- Gigerenzer, Gerd, and Reinhard Selten. 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, Mass.: MIT Press.
- Gintis, Herbert. 2000. *Game Theory Evolving*. Princeton, N.J.: Princeton University Press.
- . 2004. "The Genetic Side of Gene-Culture Coevolution: Internalization of Norms and Prosocial Emotions." *Journal of Economic Behavior and Organization* 53(1): 57–67.

- . 2007. "A Framework for the Unification of the Behavioral Sciences." *Behavioral and Brain Sciences* 30(1): 1–61.
- Gintis, Herbert, Eric Alden Smith, and Samuel Bowles. 2001. "Costly Signaling and Cooperation." *Journal of Theoretical Biology* 213(1): 103–19.
- Goerg, Sebastian J., and Gari Walkowitz. 2010. "On the Prevalence of Framing Effects Across Subject-Pools in a Two-Person Cooperation Game." *Journal of Economic Psychology* 31(6): 849–59.
- Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. New York: Cambridge University Press.
- Gürerk, Özgür, Bernd Irlenbusch, and Bettina Rockenbach. 2006. "The Competitive Advantage of Sanctioning Institutions." *Science* 312(5770): 108–11.
- Guzmán, Ricardo Andrés, Carlos Rodríguez-Sickert, and Robert Rowthorn. 2007. "When in Rome, Do as the Romans Do: The Coevolution of Altruistic Punishment, Conformist Learning, and Cooperation." *Evolution and Human Behavior* 28(2): 112–17.
- Haley, Kevin J., and Daniel M. T. Fessler. 2005. "Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game." *Evolution and Human Behavior* 26(3): 245–56.
- Harbaugh, William T., Kate Krause, and Steven G. Liday. 2002. "Bargaining by Children." Economics Working Paper 2002-4. Eugene: University of Oregon.
- Harbaugh, William T., Ulrich Mayr, and Daniel R. Burghart. 2007. "Neural Responses to Taxation and Voluntary Giving Reveal Motives for Charitable Donations." *Science* 316(5831): 1622–25.
- Harsanyi, John C., and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.
- Hayashi, Nahoko, Elinor Ostrom, James Walker, and Toshio Yamagishi. 1999. "Reciprocity, Trust, and the Sense of Control." *Rationality and Society* 11(1): 27–46.
- Henrich, Joseph. 2000. "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga." *American Economic Review* 90(4): 973–80.
- . 2004. "Cultural Group Selection, Coevolutionary Processes, and Large-Scale Cooperation." *Journal of Economic Behavior and Organization* 53(1): 3–35.
- . 2006. "Cooperation, Punishment, and the Evolution of Human Institutions." *Science* 312(5770): 60–61.
- . 2009a. "The Evolution of Costly Displays, Cooperation, and Religion: Credibility Enhancing Displays and Their Implications for Cultural Evolution." *Evolution and Human Behavior* 30(4): 244–60.
- . 2009b. "The Evolution of Innovation-Enhancing Institutions." In *Innovation in Cultural Systems: Contributions in Evolution Anthropology*, ed. Stephen J. Shennan and Michael J. O'Brien. Cambridge, Mass.: MIT Press.
- Henrich, Joseph, and Robert Boyd. 1998. "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences." *Evolution and Human Behavior* 19(4): 215–42.
- . 2001. "Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas." *Journal of Theoretical Biology* 208(1): 79–89.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton, and David Tracer. 2005a. "'Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in Fifteen Small-Scale Societies." *Behavioral and Brain Sciences* 28(6): 795–815.
- . 2005b. "Models of Decision-Making and the Coevolution of Social Preferences." *Behavioral and Brain Sciences* 28(6): 838–55.
- Henrich, Joseph, and James Broesch. 2011. "On the Nature of Cultural Transmission Networks: Evidence from Fijian Villages for Adaptive Learning Biases." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1567): 1139–48.
- Henrich, Joseph, and Francisco Gil-White. 2001a. "The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission." *Evolution and Human Behavior* 22(3): 165–96.
- . 2001b. "The Evolution of Prestige—Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission." *Evolution and Human Behavior* 22(3): 165–96.
- Henrich, Joseph, and Richard McElreath. 2003. "The Evolution of Cultural Evolution." *Evolutionary Anthropology* 12(3): 123–35.
- Henrich, Natalie, and Joseph Henrich. 2007. *Why Humans Cooperate: A Cultural and Evolutionary Explanation*. Oxford: Oxford University Press.
- Herrmann, Benedikt, Christian Thoni, and Simon Gächter. 2008. "Antisocial Punishment Across Societies." *Science* 319(5868): 1362–67.
- Hibbs, Douglas A., and Ola Olsson. 2004. "Geography, Biogeography, and Why Some Countries Are Rich and Others Are Poor." *Proceedings of the National Academy of Sciences of the United States of America* 101(10): 3715–20.
- Hirschman, Albert. 1982. "Rival Interpretations of Market Society—Civilizing, Destructive, or Feeble." *Journal of Economic Literature* 20(4): 1463–84.
- Hoffman, Elizabeth, Kevin McCabe, Keith Shachat, and Vernon Smith. 1994. "Preferences, Property Rights, and Anonymity in Bargaining Games." *Game and Economic Behavior* 7(3): 346–80.
- Hoffman, Elizabeth, Kevin McCabe, and Vernon Smith. 1998. "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology." *Economic Inquiry* 36(3): 335–52.
- Hume, David. 2006. *An Inquiry Concerning the Principles of Morals*. New York: Cosimo Classics. (Originally published in 1751.)
- Janssen, Marco A., Robert Holahan, Allen Lee, and Elinor Ostrom. 2010. "Lab Experiments for the Study of Social-Ecological Systems." *Science* 328(5978): 613–17.
- Jaswal, Vikram K., and Leslie A. Neely. 2006. "Adults Don't Always Know Best: Preschoolers Use Past Reliability over Age When Learning New Words." *Psychological Science* 17(9): 757–58.
- Johnson, Allen. 2003. *Families of the Forest: Matsigenka Indians of the Peruvian Amazon*. Berkeley: University of California Press.
- Johnson, Allen, and Timothy Earle. 2000. *The Evolution of Human Societies*, 2nd ed. Stanford, Calif.: Stanford University Press.
- Johnson, Dominic D. P. 2005. "God's Punishment and Public Goods: A Test of the Supernatural Punishment Hypothesis in 186 World Cultures." *Human Nature: An Interdisciplinary Biosocial Perspective* 16(4): 410–46.
- Kameda, Tatsuya, and Daisuke Nakanishi. 2002. "Cost-Benefit Analysis of Social/Cultural Learning in a Non-stationary Uncertain Environment: An Evolutionary Simulation and an Experiment with Human Subjects." *Evolution and Human Behavior* 23(5): 373–93.
- Kelly, Raymond C. 1985. *The Nuer Conquest*. Ann Arbor: University of Michigan Press.
- Kendal, Jeremy, Marcus W. Feldman, and Kenichi Aoki. 2006. "Cultural Coevolution of Norm Adoption and Enforcement When Punishers Are Rewarded or Non-punishers Are Punished." *Theoretical Population Biology* 70(1): 10–25.
- Knight, Jack, and Jean Ensminger. 1998. "Conflict over Changing Social Norms: Bargaining, Ideology, and Enforcement." In *The New Institutionalism in Sociology*, ed. Mary C. Brinton and Victor Nee. New York: Russell Sage Foundation.
- Kohler, Timothy A., Stephanie VanBuskirk, and Samantha Ruscavage-Barz. 2004. "Vessels and Villages: Evidence for Conformist Transmission in Early Village Aggregations on the Pajarito Plateau, New Mexico." *Journal of Anthropological Archaeology* 23(1): 100–118.
- Lachlan, Robert F., Lucy Crooks, and Kevin N. Laland. 1998. "Who Follows Whom? Shoaling Preferences and Social Learning of Foraging Information in Guppies." *Animal Behaviour* 56(1): 181–90.
- Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112(2): 443–77.
- Lancy, David. 1996. *Playing on the Mother-Ground: Cultural Routines for Children's Development*. London: The Guilford Press.
- Landes, David S. 1998. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: W. W. Norton & Co.
- Levy, Reuben. 1957. *The Social Structure of Islam; Being the Second Edition of "The Sociology of Islam"*. Cambridge: Cambridge University Press.
- Mackie, Gerry. 1996. "Ending Footbinding and Infibulation: A Convention Account." *American Sociological Review* 61(6): 999–1017.
- McAdams, Richard H. 2000. "A Focal Point Theory of Expressive Law." *Virginia Law Review* 86(8): 1649–1729.
- McClelland, David C. 1961. *The Achieving Society*. New York: Van Nostrand.
- McElreath, Richard, Adrian V. Bell, Charles Efferson, Mark Lubell, Peter J. Richerson, and Timothy Waring. 2008. "Beyond Existence and Aiming Outside the Laboratory: Estimating Frequency-Dependent and Pay-off-Biased Social Learning Strategies." *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1509): 3515–28.

- McElreath, Richard, Mark Lubel, Peter Richerson, Timothy Waring, William Baum, Edward Edsten, Charles Efferson, and Brian Paciotti. 2005. "Applying Evolutionary Models to the Laboratory Study of Social Learning." *Evolution and Human Behavior* 26(6): 483–508.
- McNeill, William Hardy. 1991. *The Rise of the West: A History of the Human Community*. Chicago: University of Chicago Press.
- Mesoudi, Alex. 2009. "How Cultural Evolutionary Theory Can Inform Social Psychology and Vice Versa." *Psychological Review* 116(4): 929–52.
- Montesquieu, Charles de Secondat. 1900. *The Spirit of Laws*, trans. Thomas Nugent. New York: P. F. Collier & Son. (Originally published in 1749.)
- Nakahashi, Wataru. 2007. "The Evolution of Conformist Transmission in Social Learning When the Environment Changes Periodically." *Theoretical Population Biology* 72(1): 52–66.
- Nakahashi, Wataru, Joe Yuichiro Wakano, and Joseph Henrich. 2012. "Adaptive Social Learning Strategies in Temporally and Spatially Varying Environments." *Human Nature* 23(4): 386–418.
- Nolan, Patrick, and Gerhard Lenski. 2004. *Human Societies: An Introduction to Macrosociology*, 9th ed. Boulder, Colo.: Paradigm Publishers.
- Norenzayan, Ara, and Azim F. Shariff. 2008. "The Origin and Evolution of Religious Prosociality." *Science* 322(5898): 58–62.
- North, Douglass Cecil. 1981. *Structure and Change in Economic History*. New York: W. W. Norton & Co.
- . 1990. *Institutions, Institutional Change, and Economic Performance*. New York: Cambridge University Press.
- . 2005. *Understanding the Process of Economic Change*. Princeton, N.J.: Princeton University Press.
- North, Douglass C., and Robert Paul Thomas. 1973. *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press.
- North, Douglas C., and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49(4): 803–32.
- Nowak, Martin A., Karen M. Page, and Karl Sigmund. 2000. "Fairness Versus Reason in the Ultimatum Game." *Science* 289(5485): 1773–75.
- Nunn, Nathan. 2009. "The Importance of History for Economic Development." *Annual Review of Economics* 1(1): 65–92.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Ostrom, Elinor, Roy Gardner, and James Walker. 1994. *Rules, Games, and Common-Pool Resource Problems*. Ann Arbor: University of Michigan Press.
- Panchanathan, Karthi, and Robert Boyd. 2004. "Indirect Reciprocity Can Stabilize Cooperation Without the Second-Order Free Rider Problem." *Nature* 432(7016): 499–502.
- Pillutla, Madan M., and Xiao-Ping Chen. 1999. "Social Norms and Cooperation in Social Dilemmas: The Effects of Context and Feedback." *Organizational Behavior and Human Decision Processes* 78(2): 81–103.
- Pingle, Mark. 1995. "Imitation vs. Rationality: An Experimental Perspective on Decision-making." *Journal of Socio-Economics* 24(2): 281–315.
- Pingle, Mark, and Richard H. Day. 1996. "Modes of Economizing Behavior: Experimental Evidence." *Journal of Economic Behavior and Organization* 29(2): 191–209.
- Platteau, Jean-Philippe. 2000. *Institutions, Social Norms, and Economic Development: Fundamentals of Development Economics*, vol. 1. Amsterdam: Harwood Academic Publishers.
- Putnam, Robert D., with Robert Leonardi and Raffaella Nanetti. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, N.J.: Princeton University Press.
- Putterman, Louis. 2008. "Agriculture, Diffusion, and Development: Ripple Effects of the Neolithic Revolution." *Economica* 75(300): 729–48.
- Rakoczy, Hannes, Katharina Hamann, Felix Warneken, and Michael Tomasello. 2010. "Bigger Knows Better: Young Children Selectively Learn Rule Games from Adults Rather Than from Peers." *British Journal of Developmental Psychology* 28(4): 785–98.
- Rakoczy, Hannes, Felix Warneken, and Michael Tomasello. 2008. "The Sources of Normativity: Young Children's Awareness of the Normative Structure of Games." *Developmental Psychology* 44(3): 875–81.

- Richerson, Peter, and Robert Boyd. 1998. "The Evolution of Ultrasociality." In *Indoctrinability, Ideology, and Warfare*, ed. Irenaus Eibl-Eibesfeldt and Frank Kemp Salter. New York: Berghahn Books.
- . 1999. "Complex Societies: The Evolutionary Dynamics of a Crude Superorganism." *Human Nature* 10(1): 253–89.
- Richerson, Peter J., Robert Boyd, and Robert L. Bettinger. 2001. "Was Agriculture Impossible During the Pleistocene but Mandatory During the Holocene? A Climate Change Hypothesis." *American Antiquity* 66(3): 387–411.
- Rilling, James K., Alan G. Sanfey, Leigh E. Nystrom, Jonathan D. Cohen, David A. Gutman, Thorsten R. Zeh, Giuseppe Pagnoni, Gregory S. Berns, and Clinton D. Kilts. 2004. "Imaging the Social Brain with fMRI and Interactive Games." *International Journal of Neuropsychopharmacology* 7(supplement 1): S477–78.
- Roes, Frans L. 1995. "The Size of Societies, Stratification, and Belief in High Gods Supportive of Human Morality." *Politics and the Life Sciences* 14(1): 73–77.
- Roes, Frans L., and Michel Raymond. 2003. "Belief in Moralizing Gods." *Evolution and Human Behavior* 24(2): 126–35.
- Rosenberg, Nathan, and L. E. Birdzell Jr. 1986. *How the West Grew Rich: The Economic Transformation of the Industrial World*. New York: Basic Books.
- Ross, Lee, and Andrew Ward. 1996. "Naive Realism: Implications for Social Conflict and Misunderstanding." In *Values and Knowledge*, ed. Edward S. Reed, Elliot Turiel, and Terrance Brown. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Sachs, Jeffrey D. 2001. "Tropical Underdevelopment." Working Paper 8119. Cambridge, Mass.: National Bureau of Economic Research.
- . 2003. "Institutions Don't Rule: Direct Effects of Geography on per Capita Income." Working Paper 9490. Cambridge, Mass.: National Bureau of Economic Research.
- Sahlins, Marshall. 1961. "The Segmentary Lineage: An Organization of Predatory Expansion." *American Anthropologist* 63(2): 322–45.
- Samuelson, Larry. 1998. "Evolutionary Games and Equilibrium Selection." Cambridge, Mass.: MIT Press.
- Sanfey, Alan G. 2007. "Social Decision-Making: Insights from Game Theory and Neuroscience." *Science* 318(5850): 598–602.
- Sanfey, Alan G., James K. Rilling, Jessica A. Aronson, Leigh E. Nystrom, and Jonathan D. Cohen. 2003. "The Neural Basis of Economic Decision-Making in the Ultimatum Game." *Science* 300(5626): 1755–58.
- Schlag, Karl H. 1998. "Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-Armed Bandits." *Journal of Economic Theory* 78(1): 130–56.
- . 1999. "Which One Should I Imitate?" *Journal of Mathematical Economics* 31(4): 493–527.
- Selten, Reinhard, and Jose Apesteguia. 2005. "Experimentally Observed Imitation and Cooperation in Price Competition on the Circle." *Games and Economic Behavior* 51(1): 171–92.
- Shariff, Azim, Ara Norenzayan, and Joseph Henrich. 2010. "The Birth of High Gods: How the Cultural Evolution of Supernatural Policing Agents Influenced the Emergence of Complex, Cooperative Human Societies, Paving the Way for Civilization." In *Evolution, Culture, and the Human Mind*, ed. Mark Schaller, Ara Norenzayan, Steve Heine, Toshi Yamagishi, and Tatsuya Kameda. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Smith, Adam. 2000. *The Theory of Moral Sentiments*. New York: Prometheus Books. (Originally published in 1759.)
- Soltis, Joseph, Robert Boyd, and Peter J. Richerson. 1995. "Can Group-Functional Behaviors Evolve by Cultural Group Selection? An Empirical Test." *Current Anthropology* 36(3): 473–94.
- Sowell, Thomas. 1998. *Conquests and Cultures: An International History*. New York: Basic Books.
- Spencer, Charles, and Elsa Redmond. 2001. "Multilevel Selection and Political Evolution in the Valley of Oaxaca." *Journal of Anthropological Archaeology* 20(1): 195–229.
- Spencer, Roger W., and John H. Huston. 1993. "Rational Forecasts: On Confirming Ambiguity as the Mother of Conformity." *Journal of Economic Psychology* 14(4): 697–709.
- Sripada, Chandra, and Stephen Stich. 2006. "A Framework for the Psychology of Norms." In *The Innate Mind: Culture and Cognition*, ed. Peter Carruthers and Stephen Stich. New York: Oxford University Press.
- Stark, Rodney. 1997. *The Rise of Christianity: How the Obscure, Marginal Jesus Movement Became the Dominant Religious Force in the Western World in a Few Centuries*. New York: HarperCollins.

- Sutter, Matthias, and Martin Kocher. 2007. "Trust and Trustworthiness Across Different Age Groups." *Games and Economic Behavior* 59(2): 364–82.
- Tabibnia, Golnaz, Ajay B. Satpute, and Matthew D. Lieberman. 2008. "The Sunny Side of Fairness: Preference for Fairness Activates Reward Circuitry (and Disregarding Unfairness Activates Self-control Circuitry)." *Psychological Science* 19(4): 339–47.
- Tuzin, Donald. 2001. *Social Complexity in the Making: A Case Study Among the Arapesh of New Guinea*. London: Routledge.
- Wakano, Joe Yuichiro, and Kenichi Aoki. 2006. "A Mixed Strategy Model for the Emergence and Intensification of Social Learning in a Periodically Changing Natural Environment." *Theoretical Population Biology* 70(4): 486–97.
- Wakano, Joe Yuichiro, Kenichi Aoki, and Marcus W. Feldman. 2004. "Evolution of Social Learning: A Mathematical Analysis." *Theoretical Population Biology* 66(3): 249–58.
- Weber, Max. 1958. *The Protestant Ethic and the Spirit of Capitalism*. New York: Scribner.
- Weibull, Jorgen W. 1995. *Evolutionary Game Theory*. Cambridge, Mass.: MIT Press.
- Wright, Robert. 2009. *The Evolution of God*. Boston: Little, Brown and Co.
- Young, David, and Robert L. Bettinger. 1992. "The Numic Spread: A Computer Simulation." *American Antiquity* 57(1): 85–99.
- Young, H. Peyton. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, N.J.: Princeton University Press.