

ARTICLE

Simple Games in a Complex World: A Generative Approach to the Adoption of Norms

Randal C. Picker[†]

Uncovering the boundaries of legal rules and defining their proper limits have traditionally vexed students of the law. When must we regulate? When will behavior coalesce in an appropriate way without the intervention of law? These are big questions, as the recent explosion of norms literature makes clear.¹ I do not try to answer these questions here. Instead, I introduce a particular approach to examining these questions, stepping beyond the formal tools that law professors often use. Applying this methodology could ultimately produce a much richer feel for the possibilities and risks in these situations.

[†] Professor of Law, The University of Chicago. I thank the Sarah Scaife Foundation and the Lynde & Harry Bradley Foundation for their generous research support; Seth Chandler, Dick Craswell, Bob Ellickson, Eric Posner, Mark Ramseyer, Matt Spitzer, and participants at workshops at the American Law and Economics Association Annual Meeting, Caltech, Chicago, Georgetown, and Stanford for comments; and Cass Sunstein for enthusiasm and helpful discussion. The title of this article obviously “borrows” from Richard A. Epstein, *Simple Rules for a Complex World* (Harvard 1995), and appropriate apologies (I hope) are hereby made.

A note about reading this paper. The computer simulations presented here are inherently dynamic, and the best way to grasp the dynamics is to see them. The published version of this paper includes a color insert that sets out snapshots of these dynamics. A CD-ROM version of the paper is also available from The University of Chicago Law Review. Requests should be directed in writing to Dawn M. Matthews, Business Manager, The University of Chicago Law Review, The University of Chicago Law School, 1111 East 60th Street, Chicago, Illinois 60637 or by phone at 773-702-9593. Finally, the simulations are also posted at <<http://www.law.uchicago.edu/Picker/aworkingpapers/norms.html>>.

¹ See, for example, *Symposium: Law, Economics, & Norms*, 144 U Pa L Rev 1643 (1996).

In looking at this traditional question of how and when must we regulate, I confine the focus of this article. This article examines norm competition: identifying the circumstances under which one norm will drive out a second norm or the conditions that will allow two norms to co-exist in a stable outcome. Put differently, this paper investigates the scope of collective action problems in the adoption of norms. I will generate a wide variety of outcomes that might result in norm competition; I will not, however, generate the norms themselves. It is possible that we might create these norms in models of the sort described in this paper, but I will not do so and will take particular norms as simply given and wholly outside the formal model. Understanding where we get competing norms will certainly require painstaking investigation into particular institutions and situations.²

This paper has three purposes. First, I want to step beyond simple game-theoretic formulations of norms to examine a larger, more realistic setting. Simple two-by-two games are a principal focus of analysis in game theory generally and in game theory and the law more particularly. This focus is quite understandable: these games are tractable and provide a familiar framework. Nonetheless, simplicity is both a vice and a virtue. These models seem almost naked, stripped of a meaningful strategy space and shorn of the multiplicity of players that typify real-life situations. The adoption of a particular norm is quintessentially a problem of more than small numbers, and we need to move beyond freestanding two-by-two games.

Second, I want to use the computer as a laboratory and run experiments in self-organization. Computer modeling of interactive situations has exploded. There are many labels associated with this work, including: "complexity," "artificial life," "artificial societies," "agent-based modeling," and "massively parallel microworlds." The common elements of this work, however, are straightforward: identify a situation with a substantial number of actors, specify rules for their actions, and let the system rip to see what happens. Results emerge, patterns form, and the system organizes on its own. Central to this work is the idea of self-organization, but this concept is hardly new. Friedrich von

² See, for example, Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U Pa L Rev 1765, 1768 (1996) (studying merchant practice and challenging the idea that courts should seek to discover and apply "immanent business norms"); J. Mark Ramseyer, *Products Liability Through Private Ordering: Notes on a Japanese Experiment*, 144 U Pa L Rev 1823 (1996) (studying the decision of many Japanese firms to subject themselves voluntarily to strict products liability at a time when Japan did not mandate it).

Hayek emphasized self-organization,³ and Adam Smith's invisible hand is its defining metaphor. Moreover, scholars have already used computers to test notions about basic social phenomena. Robert Axelrod's classic, *The Evolution of Cooperation*,⁴ did exactly that more than a decade ago. Nonetheless, recent changes in computer modeling techniques make it possible to test *in silico*, as the phrase goes, the circumstances under which a society will evolve on its own to a desired social outcome. These tests in societal self-organization are essential first steps before we can understand the possible domain for laws.

Finally, I try to contextualize my game theoretic results. Game theory—on its own and as applied to law—has generated little more than “possibility” results. A given model will show that a particular outcome is possible in the context of a coherent rational framework, but will not reveal the empirical importance of the phenomenon. A given result may be quite brittle, obtainable only if the key parameters are tuned just so, but lost if the values do not line up precisely. To understand whether the possibility result matters, we need to understand the full range of the relevant parameter space. The computer models described here allow us to test a wide range of possibilities directly and in so doing judge how robust our outcomes are.

My computer models lead to several conclusions. Under a broad set of assumptions, my model societies exhibit strong self-organization in the presence of shared values about norms. When norms are competing—when two norms are in play simultaneously—the individuals in the society successfully coalesce around the Pareto-superior norm. This is not to say that the good norm is invariably entrenched or that we cannot influence whether the good equilibrium is obtained. The set of starting conditions that leads to the superior norm—the basin of attraction for that norm—depends on the scope of connectedness among neighbors, the amount of information available to neighbors when making decisions, and the rules they use to assess that information. Each of these is a possible instrument for action as the government seeks to funnel a larger chunk of the possible initial conditions into the desired outcome. Think of each outcome as being associ-

³ Hayek identified the concept of spontaneous order with the market and with common law. See, for example, F.A. Hayek, *Studies in Philosophy, Politics and Economics* 96-105 (Chicago 1967). Hayek traced the belief in control planning to Cartesian rationalism, in reaction to which Adam Smith and other British philosophers of the eighteenth century “provided a comprehensive theory of the spontaneous order of the market.” Id at 98-99. See also A.I. Ogus, *Law and Spontaneous Order: Hayek's Contribution to Legal Theory*, 16 J L & Soc 393 (1989).

⁴ Robert Axelrod, *The Evolution of Cooperation* (Basic 1984).

ated with a funnel. Making the mouth of the funnel for the good equilibrium relatively larger—expanding its basin of attraction—then emerges as an important way for the government to implement policy.

Furthermore, the results suggest that we should be less sanguine about sequential norm competition, as occurs when a new norm competes with an old, entrenched norm. The old norm will likely continue to survive, notwithstanding that its useful life has expired.⁵ The seeding of norm clusters might jump-start the transition from old to new, causing a “norm cascade.”⁶ In this scenario, the government—or, for that matter, charities, for-profit organizations, or you and I—would encourage experimentation; and, if the conditions were right, the seeded norms would take root and spread, and society would successfully move from the old norm to the new norm. If the old norm really should survive, the experiment fails, and society loses very little.

Thus, to some extent, the collective action problem faced in norms and social meaning analysis has been overstated. I no longer lose sleep over this problem when we have simultaneous competition between two competing norms. Of course, in many social settings, the choice of norms may not be a binary on/off or yes/no choice. For example, in some tort and commercial settings, more than two norms may compete simultaneously, but my models will say nothing about this situation.⁷ And competition over time still remains a genuine problem, although to understand this problem, we need first to understand the circumstances that drive experimentation with new norms.

This article has five sections. Section I sets out the basic problem of norm competition and norm adoption and discusses the prior related literature. Section II lays out the basic frame-

⁵ Compare the problem of “excess inertia” in the network externalities literature. See Douglas G. Baird, Robert H. Gertner, and Randal C. Picker, *Game Theory and the Law* 209-11 (Harvard 1994) (illustrating excess inertia in the context of network externalities). See also Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va L Rev 757, 763-64 (1995) (arguing that because the value of a corporate contract term is related to the number of firms that adopt it, decentralized individual decisionmaking in the market may fail to produce a socially optimal equilibrium).

⁶ Cass R. Sunstein, *Social Norms and Social Roles*, 96 Colum L Rev 903, 909 (1996) (explaining that norm cascades occur when there are rapid shifts in norms, perhaps as a result of fragile social conditions). Compare Sushil Bikhchandani, David Hirshleifer, and Ivo Welch, *A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades*, 100 J Pol Econ 992, 994 (1992) (presenting a model and examples of individuals rapidly converging on one action despite imperfect information).

⁷ For an argument based on cultural evolution that suggests that commercial norms will not likely be optimal, see Jody S. Kraus, *Legal Design and the Evolution of Commercial Norms*, 26 J Legal Stud 337 (1997).

work of the analysis. Section III discusses the simulation results. Section IV ties these results to approaches that the law might take vis-à-vis norms. Section V concludes the paper, discusses its key limitations, and suggests directions for future work.

I. ORIENTATION: THE PROBLEM AND THE LITERATURE

The idea of norms is sufficiently well understood that I will introduce it only briefly before moving to consider the relevant literature. Consider three situations:

- You go to lunch with a business associate. It is Friday and the end of a long week. The waiter approaches your table and asks whether you would like to order a drink. You hesitate. You would like a drink, but at the same time you do not want your lunch partner to think ill of you for having a drink. Of course, she may be hoping that you will order a drink so that she can order one as well. What do you do? What does she do?
- During a speech you want to mention the substantial role played in your business by members of a particular racial group. Do you refer to these employees as “African-Americans”? “Blacks”? “People of Color”? You know of course that past terms for this racial group are no longer acceptable, notwithstanding continued use by organizations such as the NAACP and the United Negro College Fund. You do not want to be seen as following what might be seen as the new political orthodoxy, but at the same time you also do not want to offend these valued employees. What do you do?
- You are negotiating the terms of your employment with a new employer. You care about the parental leave policy, as you hope to have children soon. You are nonetheless reluctant to ask about this policy, because you fear that your new employer may doubt your commitment to the new job. What do you do?

These are situations in which the background context—whether described as a norm, a social meaning, or a social role—matters in an important way. The lunch presents a situation where neither person wants to move first. Other cases similar to this include asking first for a prenuptial agreement, which could be seen as a sign of doubts about the marriage, and moving to color-blind hiring unilaterally in a community dominated by discrimi-

natory norms.⁸ A social norm may exist that will resolve these situations in ways that benefit all interested parties. These norms could easily change over time or be subject to geographical or class variation.⁹

The second situation is more complex. It demonstrates clearly that norms can evolve and presents a clear example of a “norm entrepreneur.”¹⁰ Who used the term “African-American” before Jesse Jackson embraced it? Once Jackson did so, the norm shifted away from “Black,” and this created a complex range of possible social meanings reflected in the use of the phrase “African-American.” Initial use of the term could be seen as embracing Jesse Jackson personally or perhaps the broad set of social goals that he favored.

The third situation might be seen as just a problem in signaling theory,¹¹ but can also be understood as an issue embedded in a web of social roles and social norms. Mothers are expected to be quite involved with their children, fathers in the 90s increasingly so, and thus how one answers the question almost certainly depends on gender. Norms matter as well: if everyone routinely asks this question, it loses its signaling punch.

The issues raised by norms have generated a substantial literature. This literature establishes the context in which I will create my computer simulations. Accordingly, the following sections review the literature concerning norms and game theory; social norms, social meanings, and collective action problems; custom; social learning and social computation; agent-based computer simulations; and, finally, evolutionary and spatial games.

⁸ Eric A. Posner, *Law, Economics, and Inefficient Norms*, 144 U Pa L Rev 1697, 1730 (1996).

⁹ Indeed, the daily arbiter of American business life claims that the norm has shifted recently in favor of drinks at lunch, so enjoy. See Stephanie N. Mehta, *Lunch Hour Is Becoming Happy Hour Again*, Wall St J B1 (Sept 23, 1996).

¹⁰ See Sunstein, 96 Colum L Rev at 909 (cited in note 6) (defining “norm entrepreneurs” as people interested in changing social norms).

¹¹ See Philippe Aghion and Benjamin Hermalin, *Legal Restrictions on Private Contracts Can Increase Efficiency*, 6 J L Econ & Org 381 (1990) (arguing that restrictions on contracts that correct distortions when one party signals information to another can be valuable). See also Baird, Gertner, and Picker, *Game Theory and the Law* at 143 (cited in note 5) (suggesting that a law requiring employers to offer parental leave could solve the problem of potential employees not bargaining for parental leave because of the other inferences that the employer might draw about them before hiring them. In this case, the mandatory law could lead to a more efficient result by prohibiting signaling).

A. Norms and Game Theory

Simple two-by-two games provide much of the formal apparatus at work in the norm theory literature. For example, Sugden's masterful book *The Economics of Rights, Co-operation and Welfare*¹² establishes the two-by-two game as his central tool in studying the evolution of conventions and the rise of spontaneous order. Ellickson also sustains much of his theoretical analysis of norms in the two-by-two framework.¹³ As he makes clear up front, he views his inquiry into norms as seeking "to integrate three valuable—but overly narrow—visions of the social world: those of law and economics, sociology, and game theory."¹⁴ A survey of recent articles also reveals that the tools of game theory—and, in particular, free-standing two-by-two games—remain the formal vehicle of choice for understanding norms.¹⁵

B. Social Norms, Social Meanings, and Collective Action Problems

Recent theoretical work on social norms and social meaning advocates a substantial role for government in cultivating appropriate norms. Much of this analysis is rooted in the collective action problem faced by individuals in the adoption of a particular norm or social meaning.

For example, Cass Sunstein has recently mounted a vigorous defense of "norm management" by the government.¹⁶ He focuses on the context in which individuals make choices and identifies the important, if not pervasive, role that social norms play in de-

¹² Robert Sugden, *The Economics of Rights, Co-operation and Welfare* vii (Basil Blackwell 1986) (The book's object is "to show that if individuals pursue their own interests in a state of anarchy, order—in the form of conventions of behavior that it is in each individual's interest to follow—can arise spontaneously.").

¹³ Robert C. Ellickson, *Order without Law* 156-66 (Harvard 1991). Indeed, Ellickson devotes Chapter 9 of *Order without Law* entirely to game theory.

¹⁴ *Id.* at 1.

¹⁵ See, for example, Robert D. Cooter, *Decentralized Law for a Complex Economy: The Structural Approach to Adjudicating the New Law Merchant*, 144 U Pa L Rev 1643, 1657-77 (1996) (explaining how social norms arise and persist without state intervention); Posner, 144 U Pa L Rev at 1713-19 (cited in note 8) (examining the mechanisms by which preferences for various rules are transformed into a rule that governs behavior); Jason Scott Johnston, *The Statute of Frauds and Business Norms: A Testable Game-Theoretic Model*, 144 U Pa L Rev 1859, 1859 (1996) (analyzing the relationship between Article 2 of the Uniform Commercial Code and business norms). Compare Rajiv Sethi, *Evolutionary stability and social norms*, 29 J Econ Behav & Org 113 (1996) (examining the emergence of stable patterns of behavior that correspond to observed social norms such as vengeance and cooperation); Peter H. Huang and Ho-Mou Wu, *More Order without More Law: A Theory of Social Norms and Organizational Cultures*, 10 J L Econ & Org 390, 391 (1994) (modeling formally the control of corruption in principal-supervisor-agent relationships).

¹⁶ Sunstein, 96 Colum L Rev at 907 (cited in note 6).

fining choices. These norms will change the cost of a possible choice and thus will exert a substantial influence over observed choices.¹⁷ For example, Sunstein notes that social norms regarding smoking have changed over time.¹⁸ Behavior that once may have been admired is now seen as a sign of weakness and is a source of stigmatization. The social price of smoking has risen over time, and, like the direct cost of buying cigarettes or the health costs of cigarettes, should reduce the amount of smoking.¹⁹

Given the important way in which social norms influence the costs and benefits of particular choices, we need to understand how norms arise. Sunstein notes that individuals typically have little control over the content of a particular norm and almost no ability to push society from one norm to another.²⁰ This individual impotence raises the specter of a collective action problem where society cannot coalesce around a beneficial norm.²¹ Given this collective failure, direct government cultivation of particular norms—"norm management" in Sunstein's phrase²²—appears to be a plausible response.

Larry Lessig also has emphasized the collective action problem in his defense of social meaning regulation.²³ Again, social meanings are purely contextual. An action that might give offense in a cab in Hungary—putting on a seat belt, to pursue one of Lessig's examples²⁴—might go completely unnoticed in France. An individual is essentially powerless to alter the prevailing meaning associated with a particular act. She may take steps to mute that meaning, but her act has already spoken for her. In this framework, it is certainly imaginable that society will get stuck on a destructive convention.²⁵

C. Custom

Scholars also examine the same issues under the guise of custom, often in the context of evaluating reasonable behavior in

¹⁷ Id at 939-40.

¹⁸ Id at 905-06, 940.

¹⁹ Id at 950-51.

²⁰ Id at 911.

²¹ Id. See also the related discussion of coordination problems as a justification for legal intervention in Cass R. Sunstein, *After the Rights Revolution* 49-52 (Harvard 1990).

²² Sunstein, 96 Colum L Rev at 913 (cited in note 6).

²³ See Lawrence Lessig, *The Regulation of Social Meaning*, 62 U Chi L Rev 943, 991-1016 (1995). See also Lawrence Lessig, *Social Meaning and Social Norms*, 144 U Pa L Rev 2181, 2189 (1996) (asserting that norm analysis supplements traditional economics and should focus not just on describing behavior but also on the social meaning of norms).

²⁴ Lessig, 62 U Chi L Rev at 998.

²⁵ See also Dan M. Kahan, *Social Influence, Social Meaning and Deterrence*, 83 Va L Rev 349 (1997).

both commercial²⁶ and tort²⁷ settings. We used to say “customs” when we were talking about norms; now the norm, of course, is to say “norm.” Bad jokes aside, we might distinguish customs and norms based on the roles assigned by the legal system. Customs might be used to describe practices that have legal significance. They might be seen as norms that have been incorporated by reference into the law, statutory or otherwise. If so, then norms are assigned a role by the legal system only by negative inference. For example, Section 547(c)(2) of the Bankruptcy Code embraces a triple-ordinariness standard to insulate some pre-bankruptcy payments from avoidance as preferences. The standard looks to the practices between the parties but also to “ordinary business terms.” Embracing custom or current trade practices means, of course, that a court needs to figure out what those practices are, and that endeavor can be quite difficult.²⁸ In any event, nothing that I do in this paper distinguishes customs from norms.²⁹

D. Social Learning and Social Computation

This work examines the circumstances under which widely held information will be aggregated efficiently so that the right social outcome is reached. For example, imagine a new technology of uncertain quality. Individuals receive information about the technology through its use, but they receive a noisy signal. Aggregating these separate signals into an integrated framework is the work of social learning or social computation.³⁰

²⁶ See generally Bernstein, 144 U Pa L Rev 1765 (cited in note 2); Richard Craswell, *Do Trade Customs Exist?*, in Jody Kraus and Steven Walt, eds, *The Jurisprudential Foundations of Corporate and Commercial Law* (forthcoming Cambridge 1998); Kraus, 26 J Legal Stud 337 (cited in note 7).

²⁷ See generally Richard A. Epstein, *The Path to The T.J. Hooper: The Theory and History of Custom in the Law of Tort*, 21 J Legal Stud 1 (1992).

²⁸ See Craswell, *Trade Customs* at 2 (cited in note 26) (arguing that custom is based more on individual case-by-case judgments than bright-line rules). See also *In the Matter of Tolona Pizza Products Corp*, 3 F3d 1029, 1033 (7th Cir 1993) (holding that “ordinary business terms” refers to practices in which similarly situated firms engage).

²⁹ A third and overlapping literature addresses conventions. For an introduction, see H. Peyton Young, *The Economics of Convention*, 10 J Econ Perspectives 105 (Spring 1996).

³⁰ See generally Glenn Ellison and Drew Fudenberg, *Rules of Thumb for Social Learning*, 101 J Pol Econ 612 (1993) (examining learning environment models where players consider the experience of their neighbors in deciding which technology to use); Glenn Ellison and Drew Fudenberg, *Word-of-Mouth Communication and Social Learning*, 110 Q J Econ 93, 93-97 (1995) (studying the way that word-of-mouth communication aggregates the information of individual agents and may lead to the socially efficient outcome). For work from the perspective of cultural evolution, see Kraus, 26 J Legal Stud 337 (cited in note 7).

E. Agent-Based Computer Simulations

The study of the social patterns that arise when individuals interact using simple decision rules dates back at least as far as Thomas Schelling's classic work, *Micromotives and Macrobehavior*.³¹ In his characteristically low-tech fashion using only a checkerboard and dimes and pennies, Schelling showed how even relatively mild associational preferences could give rise to substantial segregation.³² The coins were randomly placed on the squares of the checkerboard. They then were given simple preferences and rules for moving around the checkerboard. For example, each coin wanted at least one-third of its neighbors to be like it. If they were not, it would move to the nearest empty square that met its demands.³³ Schelling demonstrated that segregation could emerge quite naturally even if none of the participants had an affirmative taste for discrimination.³⁴

An emerging literature reports on the uses of computers to study self-organization in social systems. These developments spill over from the complexity and artificial life research in biology and the physical sciences, which maintains a similar emphasis on computer simulations of complex adaptive systems.³⁵ Mitchell Resnick's *Turtles, Termites, and Traffic Jams*³⁶ is a wonderful introduction to the possibilities in these large, decentralized models. But Resnick's focus is on epistemology, rather than on the detailed study of particular social phenomena. Joshua Epstein's and Robert Axtell's *Growing Artificial Societies*³⁷ represents the most sustained treatment to date in the social sciences. It describes the Sugarscape, an artificial society constructed with more than twenty thousand lines of computer

³¹ Thomas C. Schelling, *Micromotives and Macrobehavior* (Norton 1978).

³² Id at 147-55.

³³ Id at 148.

³⁴ Id at 153-54. For recent applications of this work, see generally Joshua M. Epstein and Robert Axtell, *Growing Artificial Societies* 165-71 (Brookings/MIT 1996); Paul Krugman, *The Self-Organizing Economy* 15-22 (Blackwell 1996); Mitchell Resnick, *Turtles, Termites, and Traffic Jams* 81-88 (MIT 1994).

³⁵ For representative work, see generally Stuart Kauffman, *At Home in the Universe: The Search for Laws of Self-Organization and Complexity* vii (Oxford 1995) (arguing that self-organization is the root source of order in the natural world); John H. Holland, *Hidden Order: How Adaptation Builds Complexity* xviii-xix (Addison-Wesley 1995) (providing an overview of "complex adaptive systems"); Per Bak, *How Nature Works* xi (Springer-Verlag 1996) (viewing nature as "self-organized criticality"—perpetually out of balance, but organized in a poised state, the critical state, where anything can happen within well defined statistical laws).

³⁶ Resnick, *Turtles* (cited in note 34).

³⁷ Epstein and Axtell, *Artificial Societies* (cited in note 34).

code.³⁸ Behavior in the Sugarscape mimics several key elements of behavior in society: agents are born, accumulate wealth, trade, reproduce, and die. The book offers a vision of social science as seeking to replicate—or to generate—particular macro patterns from well defined initial microspecifications.³⁹ This vision gives rise to a generative approach—hence the use of the term in the title of this article—to modeling.⁴⁰ The move from micro to macro itself is not new; leading macroeconomics theorists have spent the last two decades building rigorous micro foundations for macroeconomic phenomena.⁴¹ The use of explicit computer simulations with detailed specifications of decision rules as the means to fine-tune these micro-generated macro principles is, however, innovative.⁴²

Growing Artificial Societies provides a leading example of the possibilities of agent-based computer models, but also makes apparent the weaknesses of the best models to date. The book is bereft of institutional features: notwithstanding the desire to grow everything “from the bottom up,” the authors fail to generate the norms, contracts, laws, and organizations that constitute the basic institutional framework. Of course, the same criticism applies to this article: I model norm competition, but not the creation of the competing norms themselves. As noted above, I believe that modeling the creation of these institutions (or the norms) is at least an order of magnitude more complex than the problem tackled here.⁴³

F. Evolutionary and Spatial Games

Game theory is built up from a handful of key concepts. Perhaps most important is the notion of a Nash equilibrium. A Nash

³⁸ Id at 181.

³⁹ Id at 1.

⁴⁰ Id at 4.

⁴¹ See, for example, Robert E. Lucas, Jr., *Expectations and the Neutrality of Money*, 4 J Econ Theory 103, 121-22 (1972) (modeling why real output might sputter with monetary disturbances); Thomas J. Sargent, *Dynamic Macroeconomic Theory 1* (Harvard 1987) (describing general equilibrium models that have been built to help interpret time series observations of economic aggregates).

⁴² Though even here macro theorists have worked with automatons—agents acting under bounded rationality—to investigate macro phenomena. See, for example, Thomas J. Sargent, *Bounded Rationality in Macroeconomics* (Oxford 1993).

⁴³ See text accompanying note 2. For additional discussion of Sugarscape and for a general introduction to agent-based simulations, see John L. Casti, *Would-Be Worlds* 170-79 (Wiley 1997). And, for a recent effort to apply models of self-organization to economic phenomena, see Krugman, *The Self-Organizing Economy* at vi (cited in note 34) (showing, for example, how the principle of “order from instability,” which explains the growth of hurricanes and embryos, can also explain the formation of cities and business cycles).

equilibrium is a set of self-consistent strategy choices, in the sense that each player prefers no other strategy in response to the strategy of the other players. So, for example, in a two-player simultaneous-move game where each player has two choices, say, left or right, (left, left) forms a Nash equilibrium if player 1 would have no reason to deviate from left if player 2 were to play left, and the same holds for player 2 were player 1 to play left. Put this way, I hope to highlight a key problem with the Nash idea: it is far from obvious how the players actually effectuate a Nash equilibrium. It is one thing to say that player 1 would play left if she knew that player 2 would play left and that player 2 would play left if he knew that player 1 would play left; it is something else to say how the players choose when they do not know what the other player will do. Much of the recent work in game theory has examined the circumstances under which play by boundedly rational players using simple decision rules converges on Nash equilibria.⁴⁴ This is conventionally labeled as work in evolutionary game theory, notwithstanding the use of that phrase to describe an earlier, somewhat related literature typically associated with the work of John Maynard Smith.⁴⁵ The areas of overlap between this work and the current paper will be noted throughout the paper.

In addition to this literature, Luca Anderlini and Antonella Ianni use cellular automata to explore success in a pure coordination game, where the players are indifferent to two possible equilibria.⁴⁶ They use these models to examine the relationships between absorbing states—fixed equilibrium points given the decision rules used in the model—which may or may not be strategically optimal—and Nash equilibria, which are in some sense strategically optimal.⁴⁷ Finally, the scholarship that most clearly

⁴⁴ See George J. Mailath, *Economics and Evolutionary Game Theory* (manuscript on file with U Chi L Rev); Drew Fudenberg and David K. Levine, *Theory of Learning in Games* (manuscript on file with U Chi L Rev); Michihiro Kandori, George J. Mailath, and Rafael Rob, *Learning, Mutation, and Long Run Equilibria in Games*, 61 *Econometrica* 29 (1993); Glenn Ellison, *Learning, Local Interaction, and Coordination*, 61 *Econometrica* 1047 (1993); Siegfried K. Berninghaus and Ulrich Schwalbe, *Evolution, Interaction, and Nash Equilibria*, 29 *J Econ Behav & Org* 57 (1996).

⁴⁵ John Maynard Smith, *Evolution and the Theory of Games* (Cambridge 1982).

⁴⁶ Luca Anderlini and Antonella Ianni, *Path Dependence and Learning from Neighbors*, 13 *Games & Econ Behav* 141, 142 (1996) (studying evolutionary learning in a locally interactive system). See also Luca Anderlini and Antonella Ianni, *Learning on a Torus* (forthcoming Cambridge) (investigating the behavior of locally interactive learning systems for a finite population playing a coordination game).

⁴⁷ Anderlini and Ianni, 13 *Games & Econ Behav* at 142. See also Lawrence E. Blume, *The Statistical Mechanics of Strategic Interaction*, 5 *Games & Econ Behav* 387, 389 (1993) (presenting dynamic models of strategic interaction in a population of players whose direct interaction is local but indirect interaction is global).

parallels the approach taken here uses a spatial version of the prisoner's dilemma model to look at questions in evolution and biology.⁴⁸ I discuss this literature in Section II below.

II. THE BASIC SETUP

A. Physical Setup

The two-by-two interactions considered in this article will have the following form:

		Player 2	
		Left	Right
Player 1	Left	a,a	c,b
	Right	b,c	d,d

Payoffs: (Player 1, Player 2)

These are symmetric games, meaning that exchanging player 1 for player 2 (or vice versa) changes nothing. In that sense, only one type of player plays these games. This strategy space is a natural starting point, but also excludes some well-known games, including the Battle of the Sexes.⁴⁹

I will embed this game in a spatial framework, and to do so, I will lay out an $n \times n$ grid.⁵⁰ Focusing on just one block of nine players in a 10×10 version of this framework, we would have the following:

⁴⁸ See generally Martin A. Nowak and Robert M. May, *Evolutionary games and spatial chaos*, 359 *Nature* 826 (1992) (studying model in which players have no memory and either always defect or always cooperate); Martin A. Nowak and Robert M. May, *The Spatial Dilemmas of Evolution*, 3 *Intl J Bifurcation & Chaos* 35 (1993) (similar extension of evolutionary game theory to include spatial dimensions); Martin A. Nowak, Robert M. May, and Karl Sigmund, *The Arithmetics of Mutual Help*, *Scientific American* 76 (June 1995) (examining various cooperative strategies and reciprocal help in the biological world); Martin A. Nowak, Sebastian Bonhoeffer, and Robert M. May, *More Spatial Games*, 4 *Intl J Bifurcation & Chaos* 33 (1994) (extending earlier work).

⁴⁹ See Baird, Gertner, and Picker, *Game Theory and the Law* at 41-42 (cited in note 5). In the traditional "Battle of the Sexes," there is a conflict between two people who want to spend the evening together but have different preferences about whether to go to a fight or to the opera. Both would rather go together to their less preferred event than go alone to their preferred event. Neither is able to communicate with the other, however, so each must guess what the other will do. This game exemplifies coordination games where there are multiple Nash equilibria and conflicts over the equilibria. Both players want to coordinate their actions, but each player wants a different outcome. *Id.* at 302.

⁵⁰ For an introduction to the issues raised by embedded games, see *id.* at 191-95.

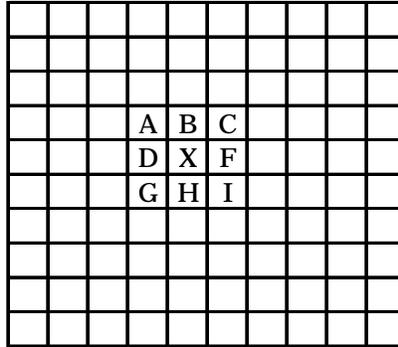


Diagram 1

Player X interacts with her immediate eight neighbors. She plays the free-standing two-by-two game with each neighbor, but she plays only one strategy per round. In other words, she will play either left or right, and that strategy will be the single play for each of the eight interactions. Player X's payoff is determined from the payoff function defined by the two-by-two game, given the plays of her neighbors.

For example, if our player played right, while all eight of her neighbors played left, she would receive a payoff of $8b$. If seven played left, while one played right, she would get a payoff of $7b + d$. This model is a natural extension of the two-by-two game to a somewhat more general framework. Note also that there are no boundaries here, notwithstanding the picture. Players at the top are treated as neighbors of the players at the bottom, at the left edge with those on the right edge. (Put differently, the layout is a doughnut, or a torus.) In the actual runs of the model, the grid is 101×101 , giving a total of 10,201 cells (and 10,201 players).

1. Payoff neighborhood.

I will discuss two kinds of neighborhoods in this article: payoff neighborhoods and information neighborhoods. A payoff neighborhood is the local area that directly impacts one player's payoffs. Once a player's strategy is chosen, the strategy choices made by the other players in the payoff neighborhood will determine the original player's payoffs and thus these strategies define the neighborhood's impact. In a basic two-by-two game, the payoff neighborhood for one player is simply the other player. In the game set forth in [Diagram 1](#), the payoff neighborhood is player X's eight neighbors.

While the notion of a payoff neighborhood is quite abstract, it would be a mistake to think that it does not track something

quite real. The payoff neighborhood is akin to a measure of community connectedness, or how linked we are to our neighbors. Of course, linkage could operate over many dimensions: linkage could be a sense of how much I care about my neighbor's welfare, or it could be much more instrumental, in the sense that your decisions help create the environment in which I operate. My payoff neighborhoods, however, are purely instrumental: my neighbor's strategy decisions set the environment, which, when coupled with my decision, create the consequences that flow to all of us.

Any number of payoff neighborhoods are possible, but I will work with the two standard neighborhoods from the cellular automata literature:

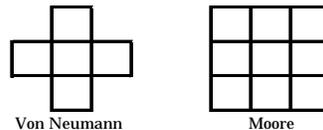


Diagram 2

The grouping on the left is the Von Neumann neighborhood, the grouping on the right is the Moore neighborhood. Focus on the cell at the center of each neighborhood. In the Von Neumann version, the payoff of each player is determined by her strategy choice and the choices of her neighbors to the immediate East, West, North, and South. The Moore neighborhood starts with the Von Neumann neighborhood and adds the four diagonal cells. The payoff of the center player is given by her decision and the decision of her eight neighbors. (Diagram 1 above thus represents a Moore neighborhood.) Of course, each player will be treated as centered at a payoff neighborhood, so the mosaic created is one of overlapping payoff neighborhoods.⁵¹

2. Information neighborhood.

An information neighborhood is the area in which a player observes results. One might think of it as the “vision” of the player. This information will form the basis for the player's strategy choice in the next round of the model. In many cases, the payoff neighborhood and the information neighborhood will

⁵¹ For a general introduction to the Von Neumann and Moore neighborhoods, see Richard J. Gaylord and Kazume Nishidate, *Modeling Nature* 4-7 (Springer-Verlag 1996). Also note, as Bob Ellickson did in a letter to me, that each individual in this model has a unique reference group, as the overlap is imperfect. These social groups are therefore different from kinship or professional groups where each person has the same “neighbors.”

be identical. But, as a general matter, it would be a mistake to assume that these neighborhoods need be coextensive. Information has a natural flow to it. Information also is a natural instrument for changing results in the models. Creating an ethic of welfare connectedness—what many would label a sense of community—is certainly extraordinarily difficult. Even creating instrumental connectedness is probably difficult, given the ability of individuals to isolate themselves from the consequences of the decisions of others. In contrast, information connectedness is much easier: it is relatively straightforward to provide information about others (though getting the recipients to listen is another hurdle).

To examine this divergence between payoff and information neighborhoods, I will vary the information neighborhoods. I will begin by looking at the coextensive cases and then move to cases in which the information neighborhood is larger than the payoff neighborhood. I will first embed a Von Neumann payoff neighborhood in a Moore information neighborhood. The player at the center will still have her payoffs determined by the choices made by her North, South, East, and West neighbors, but when she chooses her strategy for the next round, she will have seen the outcome of the prior round for her eight Moore neighbors. Next, I will embed the Moore neighborhood in a double Moore neighborhood, letting the player at the center see out a distance of two cells in all directions instead of just one cell. Finally, I will look at one version of a global information neighborhood, where players see all of the outcomes. Lest this be thought silly, information sources such as newspapers and stock markets may play exactly this kind of aggregation role.

B. Strategy Choice

Next, we need to specify some rules regarding strategies. In most of the models, initial strategies will be assigned at random. I will vary the distribution of initial left and right players and test how these initial conditions influence outcomes. As is conventional in this literature, I will use a color-coding scheme to track strategy changes between rounds. This scheme will be used in the color insert associated with this article and the [website](#) where videos of simulations are available.⁵² Other than in the initial round, a cell will be coded as blue if the player occupying it has played left in two consecutive rounds and will be coded as

⁵² The [website](http://www.law.uchicago.edu/Picker/aworkingpapers/norms.html) is at <http://www.law.uchicago.edu/Picker/aworkingpapers/norms.html>.

red if the player has played right in these same rounds. A player who switches from left to right is coded as yellow, and one who makes the reverse switch—from right to left—is coded as green. (In the initial random distribution, any cell playing left is coded as blue, playing right as red.) To review the color-coding for the strategy choices per round:

COLOR	STRATEGY
Blue	Two Rounds of Left
Red	Two Rounds of Right
Yellow	Switch from Left to Right
Green	Switch from Right to Left

Turn next to the substantive question of how we define strategy choice, which determines how the model will evolve from round to round, generation by generation. Contrast three frameworks for looking at decisionmaking and knowledge:

- *The Full Information/Full Rationality Model:* Players know the strategy space of the game, know the function giving rise to the payoffs, and have the ability to assess optimal strategy. This model tracks the usual assumption that the players have full knowledge of the rules of the game that they are playing and the ability to use that information in the best possible way.
- *The Biological Model:* Two types of players exist, one can only cooperate, one can only defect (that is, one plays left only, one plays right only). Thus, these “players” make no decisions at all; nature and instinct have programmed their “decisions.” The situation evolves by tying reproductive success in the next generation to payoffs in the current generation. A given cell is occupied in the next round by a player of the type that received the highest payoff among the nine cells centered on the cell in question. This model is a natural interpretation of the Darwinian model of adaptive success generating reproductive success.
- *The Limited Information/Incomplete Rationality Model:* Players lack full information regarding the strategy space or the way strategies interact to give rise to payoffs. For example, they might know that they are playing a two-by-two game, but they do not know which one. They may learn this information through the play of the game, though the approach to learning needs to be specified. Incomplete rationality may mean that the players’ abilities are bounded or

that rationality analysis is insufficient to guide behavior. Some heuristic rule is required for decisions, but this rule is not generated internally from the rationality assumptions. Examples might include selecting the strategy that does best among those observed by a player or playing a spatial version of tit-for-tat.⁵³

It is important to understand where these approaches converge and diverge. Consider a spatial version of the prisoner's dilemma. Given its overall prominence, I would be remiss in not addressing the prisoner's dilemma, though for reasons that will become clear, I think the spatial prisoner's dilemma is a dead-end for the law and economics crowd. Start with a free-standing version of the game that fits the general scheme described above:

		Player 2	
		Left	Right
Player 1	Left	a,a	c,b
	Right	b,c	d,d

Payoffs: (Player 1, Player 2)

Assume that $b > 1$. If we analyze this game, we should expect both players to play right. If player 1 expected player 2 to play right, player 1 would be indifferent between playing left or right because her payoffs are the same (0) under either strategy. But if player 1 expected player 2 to play left, she would clearly play right, as $b > 1$. This reasoning holds for both players, so we should expect (right, right) with a payoff of (0,0). Obviously, the players would both be better off with (left, left), as they each would receive a payoff of 1. This result replicates the essential feature of the prisoner's dilemma.

1. Full information/full rationality model.

Now look at the same prisoner's dilemma in the spatial context in each of our three frameworks. Start with the full information/full rationality model. Players know exactly what game they are playing and assess strategies in an individually rational way. As in the free-standing prisoner's dilemma, all players should defect. To see this result, walk through the possibilities one-by-one

⁵³ See Baird, Gertner, and Picker, *Game Theory and the Law* at 171-72 (cited in note 5). "Tit-for-Tat" is the strategy in a repeated game in which a player cooperates in the first period and, in all subsequent periods, defects if the other player defected in the immediately preceding period, but otherwise cooperates. Id at 316.

and assume again that $b > 1$. First, what should you do if you know that all eight of your neighbors are going to cooperate? If you cooperate you receive a payoff of 8, but if you defect you receive a payoff of $8b$, so clearly you should defect. Next, suppose that seven neighbors were going to cooperate and one was going to defect. You get 7 from cooperating, but $7b$ from defecting, so you defect. Now, approach the other extreme. Suppose that seven of your neighbors were going to defect and one intended to cooperate. You will get nothing from your defecting neighbors, regardless of whether you cooperate or defect, but you will get 1 from your cooperating neighbor if you cooperate and b from him if you defect. So you defect. Finally, suppose that all eight of your neighbors were going to defect. You would get 0 from cooperating, and 0 from defecting, so you are indifferent. Taking all of this analysis together, you never do worse by defecting and often do better, so you will defect. Everyone defects, just as in the free-standing model.

The spatial feature adds nothing in the full information/full rationality setting. The analysis does show that the central failure of the free-standing prisoner's dilemma—individually rational behavior is collectively foolish—carries over to the spatial setting, but we should have guessed that anyhow. So switch frameworks, and consider how the spatial version of the prisoner's dilemma fares if we use a biological model instead.

2. Biological model.

In the biological model, the spatial prisoner's dilemma is little more than a payoff function for the two strategies of cooperation and defection. Recall that no decisions are made in the biological model: actors simply have a preordained type. There is no obvious reason to think that this model will evolve in any particular way, and indeed, as Nowak, May, Bonhoeffer, and Sigmund show,⁵⁴ a rich variety of behavior emerges. In some cases, the model converges to the all-defection outcome that we saw in the full information/full rationality model. In other cases, the model cycles forever through a handful of states, all of which involve a mix of cooperation and defection. In yet other cases—and these are arguably the most interesting—the model appears to be almost chaotic. The model cycles only over a long time period, and the cycle will be all but undetectable to you or me.⁵⁵

⁵⁴ See note 48.

⁵⁵ To see an example of this, set $b = 1.65$, and start with an initial random distribution of 10 percent defectors and 90 percent cooperators. A movie showing the first hun-

This result is a dramatic change from the full information/full rationality model, which locks into the defection equilibrium immediately and is bereft of interesting behavior. In contrast, the biological framework provides an embarrassment of riches. Where does this result put us? Biology is a great subject, but we want models in which people make decisions and their behavior is not simply instinctive and preordained. To achieve this setting, switch to the limited information/limited rationality framework.

3. Limited information/limited rationality model.

In this framework, players know only that they are playing a two-by-two game and thus have a choice between two strategies. We now need to consider learning quite carefully. Suppose that our players learn nothing, and make decisions based on a simple observation. In each round, each player observes the payoffs that she receives and those received by her neighbors and the strategies that the neighbors play. Given this information, she adopts the strategy that did best—the strategy that resulted in the single highest payoff in the nine cells that she sees.

This strategy choice matches the biological framework exactly; indeed, that is the point of this exercise. The absence of learning means that our player never assesses whether defection is a dominant strategy. Instead, she uses a more basic decision rule that tracks exactly the reproductive success rule used in the biological model. By introducing limited information and limited rationality, we seem to have made the spatial prisoner's dilemma interesting for us.

But this outcome does not hold if we advance the model just a bit. It does not take much for our players to learn what game they are playing. A little experimentation and a little variation coupled with some simple calculations will let our players convert this limited information game to the full information version. Suppose, for example, that our test player cooperates in one round, while six of her neighbors cooperate and two defect. She will receive a payoff of 6, which she knows to be derived in the following way:⁵⁶

$$6p_{cc} + 2p_{cd} = 6$$

Obviously, an infinite number of pairs will satisfy this equation, including (1,0) and (0,3). She does not yet know much. Sup-

dred generations of this model is available at my [website](#) (cited in note 52).

⁵⁶ The symbols p_{cc} and p_{cd} represent the payoffs in the model when, respectively, both players cooperate, and player 1 cooperates and player 2 defects.

pose that in the next round, she again cooperates, while five of her neighbors cooperate and three defect. She now gets a payoff of 5, which she again knows to be generated in the following way:

$$5p_{cc} + 3p_{cd} = 5$$

Our player now has two equations with unknowns, and can solve them to learn that $p_{cc} = 1$ and $p_{cd} = 0$. Given this information, she can defect in the next round and learn the value of p_{dd} .⁵⁷ She now enjoys full information regarding the game, and can determine that defection is her dominant strategy.

The limited information model therefore tracks the interesting biological model when players do not learn, but quickly converts into the full information model when players begin to learn. To give a richer sense of this outcome, I have simulated the model one hundred times on the assumption that 25 percent of the players learn the game and defect. Moreover, I assumed that 25 percent of the remaining players defect initially, while 75 percent of this same group cooperate. After that, this predetermined group chooses the strategy that yields the highest payoff in the nine cells that they see. Under those parameter settings, eighty-five out of one hundred games ended in the all-defection equilibrium. In the other fifteen simulations, cooperation did poorly.

For me, this analysis yields the following results. We can generate complex, interesting behavior in the spatial prisoner's dilemma, and this behavior may have important implications in contexts in which individuals do not learn. Nonetheless, relatively simple learning converts the limited information framework to the full information framework, and only a modest fraction of the players need to master this learning for the interesting behavior to vanish. For most circumstances of interest to us, the spatial prisoner's dilemma should look a lot like the free-standing prisoner's dilemma. That result is worth noting, but it also means that playing the spatial prisoner's dilemma reveals nothing new.

This discussion has been an extended look down a dead-end. I started by emphasizing the need to be precise about the rationality and information assumptions used to generate strategies from generation to generation. The traditional assumptions of full information and full rationality make the spatial prisoner's dilemma an uninteresting vehicle for studying norm problems. Relatively simple learning converts a limited information/limited

⁵⁷ The symbol p_{dd} , of course, represents payoffs when both players defect.

rationality approach into the traditional rationality approach. We should therefore look elsewhere to model norm competition.

III. HOW NORM COMPETITION EVOLVES

Turn now to the traditional coordination game. I will focus on a particularly simple version of it:

		Player 2	
		Left	Right
Player 1	Left	1,1	0,0
	Right	0,0	b,b

Payoffs: (Player 1, Player 2)

As is generally known, we have little interesting to say about this model even when we use our full rationality assumptions. Dominance arguments will not solve this game. If Player 1 plays left, Player 2 wants to play left, and vice versa; if Player 1 plays right, Player 2 wants to play right, and vice versa. Neither player has a single best strategy to play. In the parlance of game theory, both (left, left) and (right, right) are Nash equilibria: neither player wants to switch strategy given the other player's strategy. Nonetheless, game theory provides no good way of choosing between these equilibria. We do not reach a determinate result, meaning that we cannot explain how the players would actually make decisions.⁵⁸ We must therefore use our limited information/limited rationality framework. Once we are in this framework, we can weaken our usual information assumptions without changing results in any material way.

Assume that players know only that they are playing a spatial two-by-two game. They know nothing about the payoffs and nothing about the strategy space. They are endowed with an initial strategy at random from all of the available strategies. They know only one strategy, and they play it. Round by round, players will observe the strategies played by their neighbors and thereby learn new strategies. Players will also observe the consequences of those strategies. Some players will learn the full payoff function, just as our players in the prior section learned that they were playing the spatial prisoner's dilemma. In this model,

⁵⁸ This overstates somewhat. Harsanyi and Selten have emphasized the idea of risk dominance to resolve these games. This idea focuses on the consequences of failing to coordinate and therefore the relative risk associated with each choice. See John C. Harsanyi and Reinhard Selten, *A General Theory of Equilibrium Selection in Games* 82-89 (MIT 1988).

the same linear learning will be fully informative about the payoff functions defined by the particular coordination game being played. But—and this is the key difference from the spatial prisoner’s dilemma—full knowledge of the strategies and the payoffs will not render this game uninteresting. In the spatial prisoner’s dilemma, learning converted the limited information framework into the full rationality framework and distanced us from the biological model. In contrast, here, because we do not have a determinate way of playing coordination games, we necessarily must reach outside the model for an operational decision rule. Learning does not convert the limited information framework into the full rationality model, and indeed, we remain quite close to the biological approach.

This discussion lays out the game and its setup. Next, we need to specify a choice rule for the players. For now, I will assume that each player uses the same rule. In [Section III.D](#), I will look at an alternative choice rule, but I will start with the rule that has received the most study in the literature: In the next round, the player will adopt the strategy that did the best, as measured by how her strategy performed and how her neighbors did in the previous round. Each player looks at the payoffs she obtained as well as those of her eight neighbors, figures out which payoff is the highest, and then adopts the strategy used by that player. As noted before, this scheme does not work in the first round—there are no prior payoffs to evaluate—so strategy choices will be assigned at random. While perfectly plausible, to me at least, this decision rule is created out of whole cloth. I do not justify it as emerging out of some other generally accepted framework.

A. Initial Examples

Before looking at the examples, note that all-left and all-right are both Nash equilibria and absorbing states (meaning that the model will not change from the state once it is reached). As will become clear, Nash equilibria and absorbing states are not interchangeable: we will have absorbing states in the models that are clearly not Nash equilibria. To see that all-left and all-right are Nash, if a given player thought that every other player was going to play right, she clearly would play right. She would get zero if she played left and a payoff of $8b$ from playing right. The same holds for left: if the player expected everyone else to play left, she would play left and get 8 rather than play right and get zero. Obviously, these two Nash equilibria in the spatial coordination game track those seen in the free-standing game.

All-left and all-right are also absorbing states—fixed points given the decision rules used. If everyone has played right, each player observes only how right has done, and therefore chooses right. If everyone plays left, each player observes only left outcomes, and chooses only left. As should be clear, the extent of initial variation in the number of players playing left or right will be important. To see this result, I will start with a few examples to give you a sense of the variety of behavior seen in this framework. Start with $b = 1.05$ and assume that the equal number of players initially play the left and right strategies. **Figures P1 to P6** on the color insert show six snapshots of the evolution of this model.⁵⁹

As is evident, this model converges to a mixed equilibrium, with large numbers of players adopting each strategy.⁶⁰ Of course, this mixed equilibrium is inefficient, as the social optimum is achieved when all players play right. Nonetheless, this result is not too surprising. The value of getting to the right equilibrium is low—1 versus 1.05—and the initial starting conditions do not tilt the tables significantly in favor of one of the equilibria.

Consider a second example. Bump b up to 1.25 and again assume that left and right are initially played in equal numbers. **Figures P7 to P12** on the color insert show six snapshots of the evolution of this model.⁶¹ All we have done is increase the value of coordinating on the second equilibrium, and now the model converges to the social optimum. Nonetheless, simply increasing b to 1.25 is not enough to assure convergence to the right equilibrium. Let 80 percent of the players start with the left strategy and 20 percent with the right, and consider the five snapshots of the model given on the color insert as **Figures P13 to P17**. Once again, the model fails to converge completely.⁶²

This outcome highlights the difference between absorbing states—fixed points—and Nash equilibria. The result in **Figure P17** is an absorbing state, but it is not a Nash equilibrium. The **diagram** below sets out the relevant chunk of the final result:

⁵⁹ To see this evolution directly, play the video at my [website](#) (cited in note 52).

⁶⁰ I mean “mixed” in the sense of having both strategies played at the same time by different players; I do not mean “mixed” in the sense of one player playing both strategies with some positive probability. For a discussion of mixed strategies played by single players, see Baird, Gertner, and Picker, *Game Theory and the Law* at 37-39 (cited in note 5).

⁶¹ See the [website](#) (cited in note 52).

⁶² *Id.*

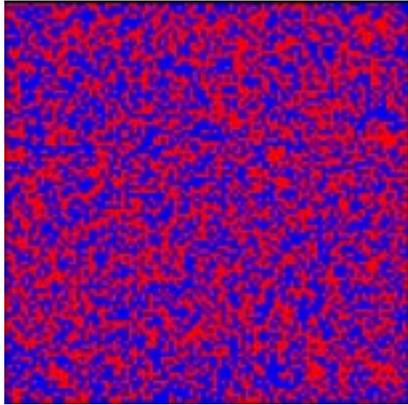


Figure P1

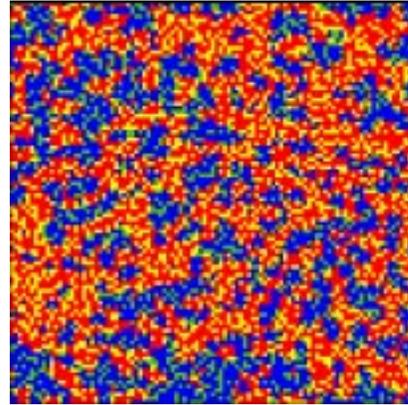


Figure P2

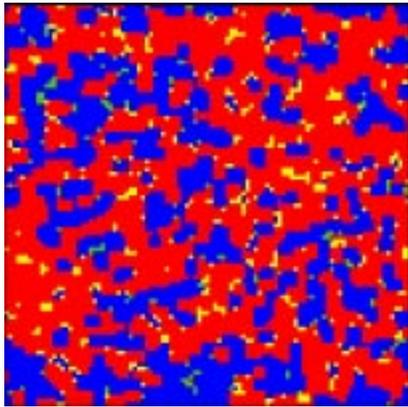


Figure P3

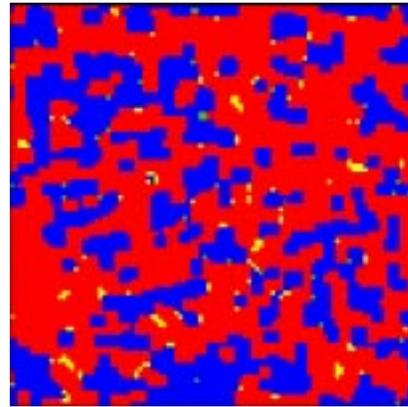


Figure P4

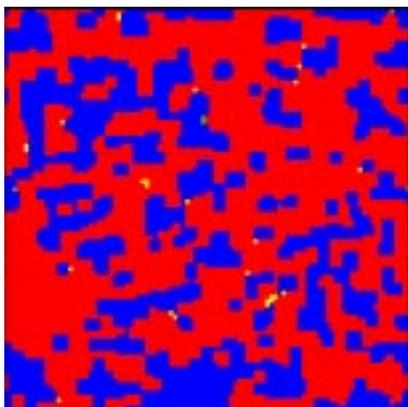


Figure P5

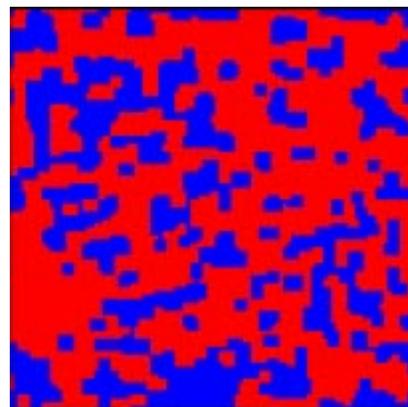


Figure P6

$b = 1.05$, Initial Setup: 50% Left, 50% Right

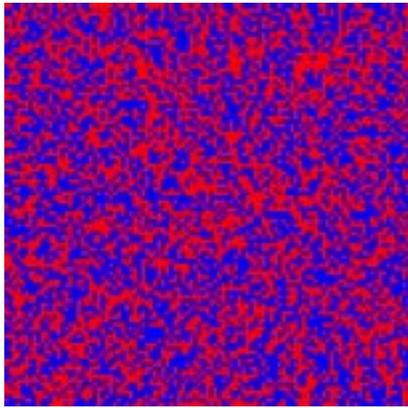


Figure P7

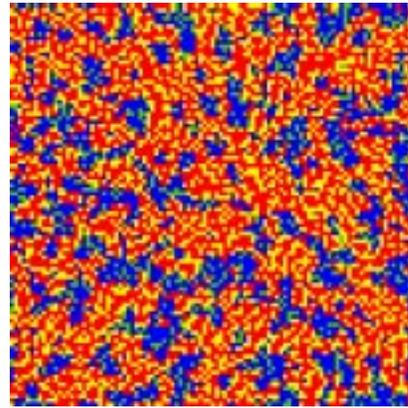


Figure P8

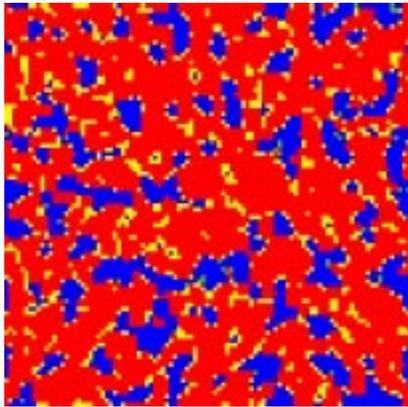


Figure P9

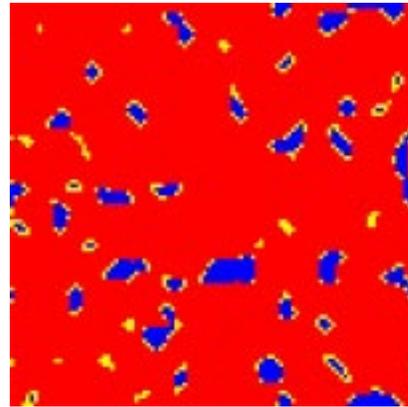


Figure P10

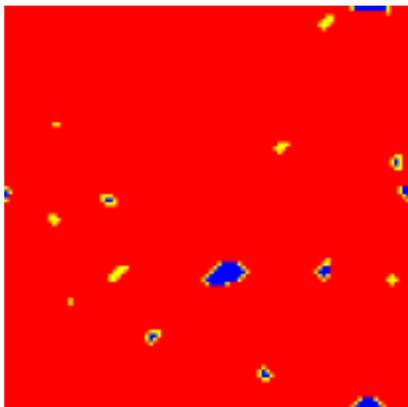


Figure P11



Figure P12

$b = 1.25$, Initial Setup: 50% Left, 50% Right

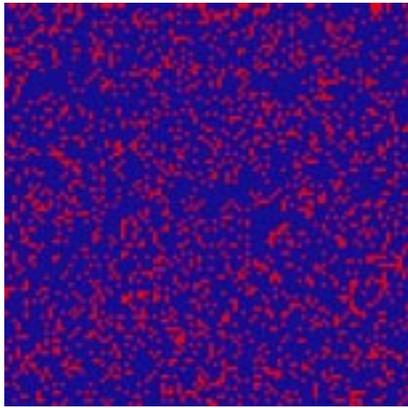


Figure P13

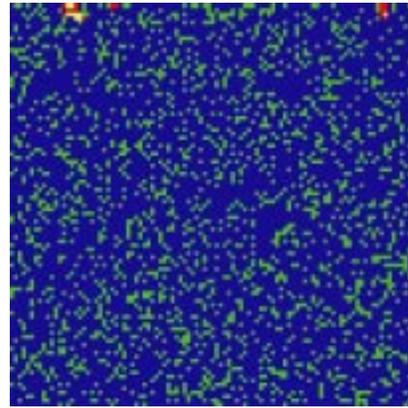


Figure P14

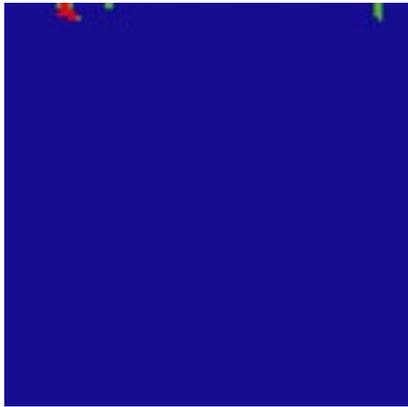


Figure P15

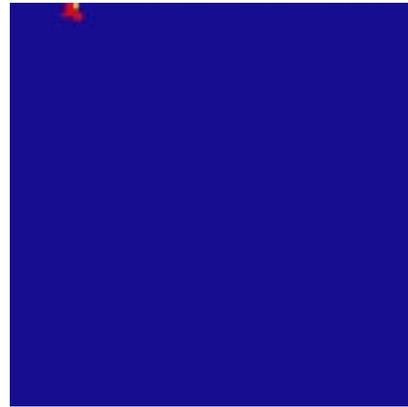


Figure P16

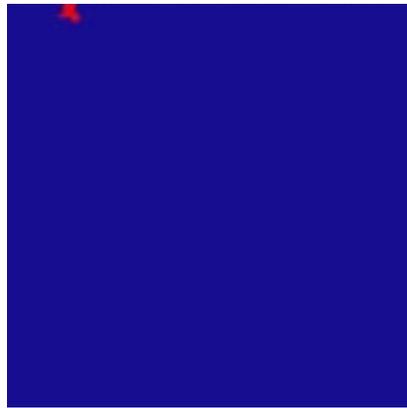


Figure P17

$b = 1.25$, Initial Setup: 80% Left, 20% Right

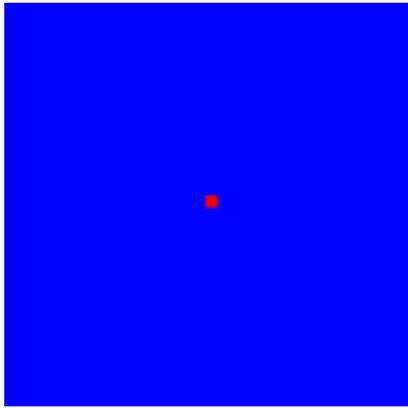


Figure P18

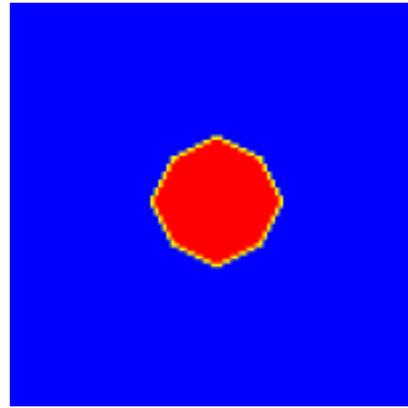


Figure P19

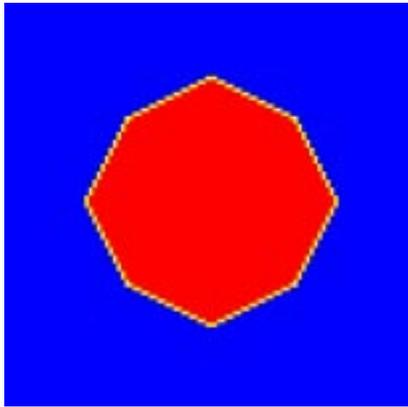


Figure P20

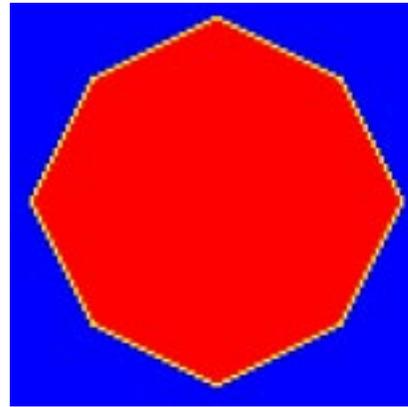


Figure P21

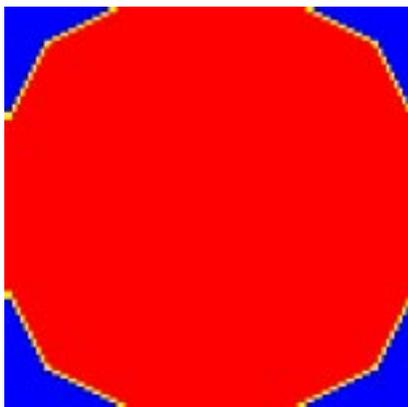


Figure P22



Figure P23

Norm Seeding: $b = 1.65$, Nine Cluster Start

B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	R	R	R	B	B	B
B	B	B	R	R	R	B	B	B
B	B	R	R	R	R	R	B	B
B	B	B	B	B	R	R	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B

Diagram 3

To understand why this equilibrium is not Nash, focus on the red (R) cell at the corner of the cluster of nine reds. Round after round in this model, this cell continues to play right and remains red because it observes the red at the center of the cluster of nine reds receiving a payoff of $8b$. This payoff, of course, is the best possible outcome in the model, and, on my assumed strategy for making decisions, any player who observes that outcome adopts right as her strategy. Nonetheless, this strategy cannot be a Nash equilibrium unless right is the best strategy for this player given all of the other strategies. Our corner red expects, however, three of her neighbors to play right, and five to play left. Given those strategies, she gets a payoff of $3b$ if she played right and a payoff of 5 from playing left, so she should switch strategies, so long as $b < 1.67$.

What should we make of this outcome? This particular mixed play absorbing state is sustainable only if we have “irrational” play round after round. Our corner red continues to play red because the center red is doing so well, even though the context in which the center red plays is quite different from corner red’s own setting. We might feel the need to abandon Nash approaches initially because they depend on highly refined introspection driven by high-level reasoning abilities. Nonetheless, we should be equally uncomfortable about results that depend on seemingly implausible play into perpetuity. Some middle ground may be more appropriate,⁶³ but, in any event, the approach taken in this paper builds or constructs an absorbing state based on a minimal set of key ideas. This generative approach may offer a more plausible account of how equilibria are actually achieved.⁶⁴

⁶³ See the suggested directions for future research at text following note 93.

⁶⁴ This parallels—indeed, overlaps with—the recent work in evolutionary game theory. See note 44.

A. The Emergence of a Phase Transition

The preceding examples give a flavor for the range of behavior that arises in the model. To get more systematic, I set $b = 1.65$ and ran sets of 100 simulations of the model for different initial densities. The results are set forth in Figure 1. To be clear on the meaning of the figure, I ran 100 simulations of the model with $b = 1.65$ for each of the initial densities shown along the x-axis (9900 simulations total). The three graphs in Figure 1 capture three possible results. All of the players could converge on playing right (“Red”); all could converge on left (“Blue”); and some could converge on left while others played right (“Mixed”).

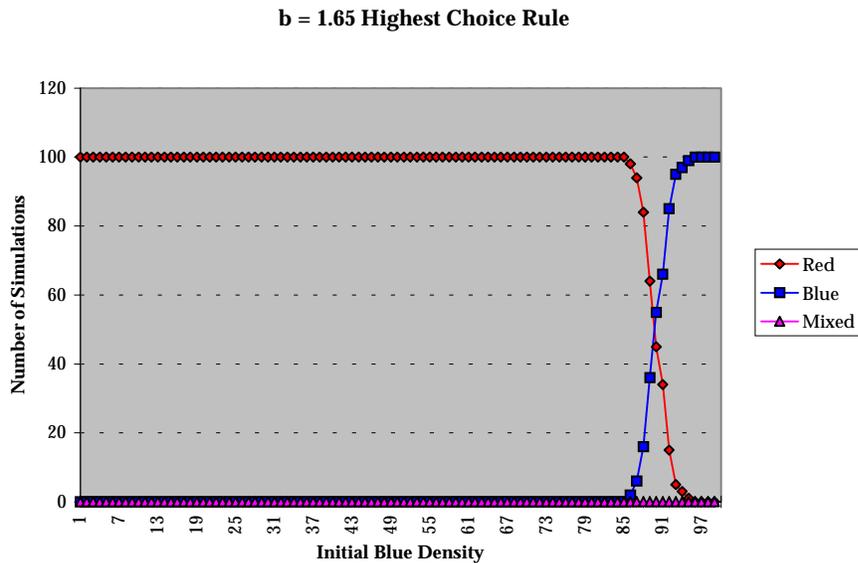


Figure 1

The graphs chart the number of times each possible outcome occurs in the 100 simulations for each initial value. If we start with 1 percent of the players playing left and 99 percent playing right, then in 100 times out of 100, the play of the game converged on the right-right (or all-red) equilibrium. In contrast, if we start with 99 percent of the players playing left and 1 percent playing right, then in 100 percent of the cases we converged on the inferior all-left equilibrium.

Neither of these results is surprising. The good (right) equilibrium, however, is particularly robust. Even if we start in tough conditions—say with 80 percent of the players playing left and 20 percent playing right—we still converge on the good equilibrium

in 100 percent of the cases. As we push the initial density of players playing the inferior choice ever higher, however, we run into problems. Some fraction of the simulations converge to the inferior equilibrium. By the time we reach just a bit more than 89 percent of the players playing left initially, the graphs cross: as many simulations converge on the bad equilibrium as converge on the good equilibrium. If we have more players choose initially the inferior strategy, more and more of the simulations converge on the inferior equilibrium. As we reach our maximal densities, the rout is complete, and all of our simulations converge on the inferior equilibrium.

The shape of these graphs is characteristic of a phase transition in physics or a model of punctuated equilibria in biology.⁶⁵ The system has two natural equilibria and shifts from one to the other occur over a narrow band. The combination of a standard two-by-two game and some neighborhood effects results in this phase transition.

What is going on here? Why is it that this system converges so well and why do we start to encounter trouble when 85 percent of the players initially choose the inferior play? To understand this phenomenon, start with a cluster of nine red cells surrounded by a sea of blue:

B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	R	R	R	B	B	B
B	B	B	R	R	R	B	B	B
B	B	B	R	R	R	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B

Diagram 4

Focus on decisionmaking by the players in the red cells. Each border red will observe the payoff of the center red, who will be getting 8b (that is, the center red interacts with eight players playing the strategy that she has played, so she is perfectly coordinated with her neighbors). The best blue cell that our border red could see has a payoff of 7. No border red will switch.

⁶⁵ See, for example, Stuart Kauffman, *At Home in the Universe* 57, 116-18 (Oxford 1995); Stephen Jay Gould and Niles Eldredge, *Punctuated Equilibria: The Tempo and Mode of Evolution Reconsidered*, 3 *Paleobiology* 115 (1977).

The center red, of course, has received the best payoff possible, and will see no reason to switch given the decision rule.

Now focus on the blue cells bordering the cluster of nine reds. There are four corner blues. A corner blue will see the best blue cell receiving a payoff of 8. The only red cell that it sees receives $3b$. Given our numbers, a corner blue will switch only if $b > 8/3$ or 2.66. Of the remaining blue cells, there are eight off-center border blues. Each will see a blue cell receiving 8, while the best red cell seen—the middle border red—will be getting $5b$. The off-center border blue will switch if $b > 8/5$ or 1.6. Finally, the analysis for the four center border blues tracks the analysis for the other border blues. Accordingly, if $b > 2.66$, all of the border blues will switch strategies; if $1.6 < b < 2.66$, all of the non-corner blues will switch, while if $b < 1.6$, no blue cells will change over.

Already, under this decision rule, we can have pockets of players playing blue and red simultaneously, if the benefits of coordinating on strategy right are not sufficiently great. This situation will give rise to clusters of blue cells and clusters of red cells and we will see a model in which multiple norms, conventions, or decentralized rules are in use at the same time.

Consider the intermediate case and roll over the cells to the next generation.

B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	Y	Y	Y	B	B	B
B	B	Y	R	R	R	Y	B	B
B	B	Y	R	R	R	Y	B	B
B	B	Y	R	R	R	Y	B	B
B	B	B	Y	Y	Y	B	B	B
B	B	B	B	B	B	B	B	B
B	B	B	B	B	B	B	B	B

Diagram 5

Nothing has changed for the red cells: the original cluster of nine play as before. Each yellow cell—recall that these are players who had been playing strategy left and who have now switched to strategy right—will see a red cell receiving a payoff of $8b$. Each yellow cell will therefore play strategy right again (and will turn red in the next round).

Now examine the blue cells as we did before: cell by cell. The former corner blues still see a blue cell playing with eight neigh-

boring blues, but will now see the best red cell playing with seven neighboring reds. The old corner blue will thus switch if $7b > 8$, or if $b > 1.14$. The new non-corner border blues will each see a red playing with five reds, and thus will switch if $b > 1.6$. The new corner blues will compare 8 with $4b$, and will switch if $b > 2$. The second iteration thus plays out in one of two fashions depending on whether b is less than or equal to 2.⁶⁶

Where $b > 1.6$, there is a powerful drive in the system to converge on the correct equilibrium. The model, however, does not necessarily get there. I started the cluster analysis with a block of nine red cells, but the relative scarcity of right strategy players may prevent the use of such a starting cluster. For example, suppose that the largest cluster of red cells is a square block of four cells, surrounded by a sea of blue. Each red cell will receive a payoff of $3b$, or 4.95 if $b = 1.65$. The best blue cell seen by each red cell will be the blue cell cater-corner to it. The blue cell will touch seven other blues and one red cell, and thus will receive a payoff of 7. Each red cell will switch strategies in the next round. This cluster is too small to support growth and it dies.⁶⁷

When $b = 1.65$, our cluster of nine red cells grows until it spreads throughout the entire domain, while our cluster of four red cells withers. A little bit more analysis makes clear that a two-by-three cluster of six red cells survives and grows when $b = 1.65$. Again, surround our cluster of six red cells with blue cells and consider the choices that each will make. Each red cell sees another red cell that received 8.25 (or $5b$) in the previous round. (In other words, each red cell sees at least one other red cell that has five red neighbors.) The payoff of 8.25 exceeds the highest payoff that a blue cell could enjoy—namely 8—and thus no red cell will switch in the next round. What will the neighboring blue cells do? Each of these neighboring blue cells sees a red cell receiving the 8.25 payoff. Again, they will see no blue doing better, so they will all switch. We now have a cluster of nine red cells, and the above analysis applies.⁶⁸ Finally, to complete the analysis, consider a cluster of five red cells. The best red payoff will be $4b$, or 6.6. Four of the five red cells will see a blue cell receiving a payoff of 7, and they will switch, then the fifth red cell will follow

⁶⁶ To see this pattern directly, play the video with a nine red cell starting point at my [website](#) (cited in note 52).

⁶⁷ To see this pattern directly, play the video with a four red cell starting point at my [website](#) (cited in note 52).

⁶⁸ To see this pattern directly, play the video with a six red cell starting point at my [website](#) (cited in note 52).

in the next round.⁶⁹ Six cells thus sustain growth while five cells do not when $b = 1.65$.

We can now make a bit more sense of [Figure 1](#). The charts represent exercises in applied probability. We know now that a cluster of six grows indefinitely until every player plays the socially preferred strategy. If we initially assign strategies at random, what is the probability that we will have one or more of the growth clusters? We could answer this question algebraically, but the chart itself provides the answer. For example, when we start with 90 percent of the players playing left and 10 percent playing right, then 55 times out of 100, we do not get at least one of our growth clusters. Absent a growth cluster, the small, scattered clusters of red die, and we converge to the inefficient blue equilibrium.⁷⁰

To put this in a different language—that of dynamic systems—we have mapped the basin of attraction for each of our point attractors. The basin of attraction for an attractor—or the catchment basin—is that chunk of the possible set of starting conditions that leads to that attractor.⁷¹ (The attractors in this example are the two Nash equilibria.) Think of a ball rolling over an undulating surface: at some point, the ball falls into a steep depression—this is a physical description, obviously, and not a statement of the ball's state of mind—and eventually comes to rest in that hole. That hole is an attractor and the starting points on the surface that result in the ball coming to rest in that particular hole form the basin of attraction for the hole. So, to return to the diagram, the basin of attraction for the red equilibrium includes initial distributions of 1 percent left to 85 percent left, while the corresponding basin of attraction for the blue equilibrium covers 96 percent left to 99 percent left. Initial distributions between 86 percent left to 95 percent left are as if our ball is running along an edge that separates the two basins of attraction and a slight nudge in one direction or the other pushes the ball into one equilibrium or the other.

⁶⁹ To see this pattern directly, play the video with a five red cell starting point at my [website](#) (cited in note 52). Note that there is not a unique way of laying out the five cells. This observation was true for our prior examples as well, but I have emphasized relative compactness of the cluster, and that gave a simple representation of the cluster.

⁷⁰ To see an example of this pattern, play the video at my [website](#) (cited in note 52). As both Dan Klerman and Seth Chandler noted in e-mails to me, this is one place where the size of the lattice matters. The bigger the lattice, the greater the chances of a growth cluster forming in the initial set of random plays.

⁷¹ For an introduction to these ideas, see I.M.T. Thompson and H.B. Stewart, *Non-linear Dynamics and Chaos* 9-12 (Wiley 1986).

Step back and now ask what we should make of this discussion. First, this game is path dependence writ large. The convergence of this system largely depends upon the initial starting conditions. The literature on path dependence grows by the day,⁷² and this model gives a crisp example of this phenomenon. Second, and perhaps of more interest given what we already know about path dependence, this setup converges on the superior equilibrium even in the face of tough starting conditions. If we just started with our free-standing coordination game, we could say very little about the likelihood that we would converge on the right equilibrium. Now, we should take some comfort that this system will get to where we want it to go.⁷³ In real situations, we might think of the initial choice of strategy as indeed random. This example says if these choices are essentially coin flips—a 50/50 chance—the model will always converge to the right norm. Even if the choice is substantially biased against the good strategy, we still converge on the best norm. (If we replace the coin with a single die, and play right only if six comes up, we still get to the good equilibrium.) And my intuition says that the bias should run in favor of the good strategy if players are choosing between both strategies at the same time. I have offered no other story of salience here other than the potential value that results from successful coordination. We should think that the extra value available would push us away from a 50/50 chance to odds favoring the good strategy.

This outcome is good news. We see a good chance of successful coordination on the right norm. We also see that it is easy to overstate the problems that define the coordination game and justify legal intervention. At least in this particular framework and on these values, I believe that it is highly unlikely that we would end up in the inferior equilibrium. For the inferior equilibrium to win over our players, we need either extremely bad luck or something that makes the inferior strategy especially salient.

Before leaving this particular example, we should look at one other variable of interest. The graph of the red outcomes in Figure 1 appears to be one of unremitting sameness until we get to particularly skewed distributions of initial choices at the end.

⁷² See, for example, W. Brian Arthur, *Competing Technologies, Increasing Returns, and Lock-In By Historical Events*, 99 *Econ J* 116 (1989) (exploring “the dynamics of allocation under increasing returns in a context where increasing returns arise naturally: agents choosing between technologies competing for adoption”).

⁷³ Substantially more general results in a related model of coordination games regarding the success of convergence to the good equilibrium are obtained in Kandori, Mailath, and Rob, 61 *Econometrica* at 44 (cited in note 44). See also Ellison, 61 *Econometrica* at 1066-67 (cited in note 44).

Value after value ends up in the same place. But this graph omits the path of convergence to the social equilibrium. These paths actually look quite different, as [Figure 2](#) should make clear. [Figure 2](#) charts the average number of periods that the model used to converge to the red equilibrium. These numbers increase steadily, but slowly, for an extended period before reaching a region of a sharp increase in the time to convergence. Unsurprisingly, the initial distribution of strategy choices sets the ultimate rate of convergence to the good social equilibrium. Still, more comfort is found here. For values in the middle of the distribution, the time to convergence is relatively modest.⁷⁴

b = 1.65: Number of Rounds to Red Convergence

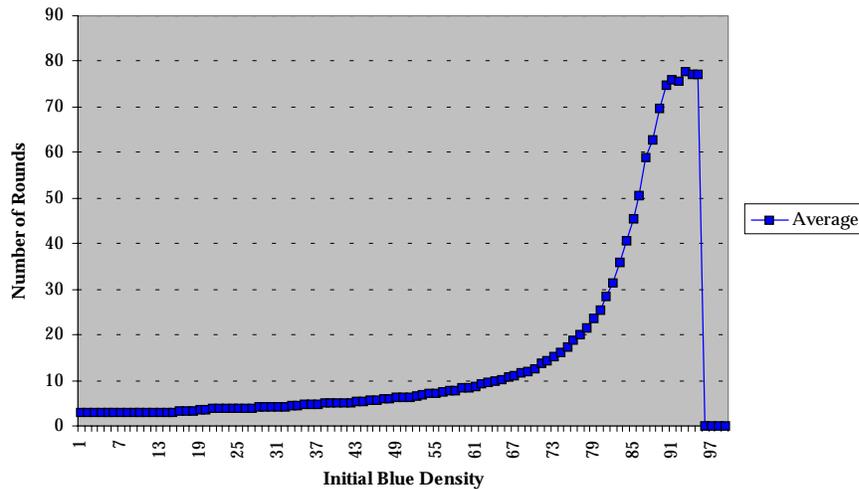


Figure 2

B. Two Phase Transitions and Mixed Play

So far we have looked at nearly ten thousand simulations of a ten thousand player model for a single value of $b = 1.65$. The next step is to understand how these results change as we alter the value of b . Start by noting that there is an inverse relation between the value of b and the minimal cluster required to assure convergence to the social optimum. Larger b values will

⁷⁴ The convergence times are also determined by the rules used to match players of the coordination game. The model used here fixes these matches in the initial $n \times n$ grid. Other matching rules might very well alter this pattern of convergence. Compare Ellison, 61 *Econometrica* at 1060-62 (cited in note 44).

support smaller initial growth clusters; smaller b values will require larger initial clusters. For example, a cluster of four red cells will survive and grow if $b > 2.66$. Each of the red cells receives a payoff of $3b$, which is greater than 8 if $b > 2.66$. These red cells will thus stay red. Adjacent blue cells will switch strategies, as they will see a red cell receiving more than 8 while no blue cell can do better than 8. We can still end up in the inefficient norm equilibrium, but that result would occur only if no cluster of four red cells formed in our initial distribution of strategies. As we increase b beyond 1.65, we keep pushing toward the right edge of the figures, until virtually all of the models converge to the good norm equilibrium.

So move b in the other direction and return to our discussion of the cluster of nine red cells. As noted above, when $b < 1.6$, the neighboring blue cells do not switch over. The blue cell most likely to switch sees at least one blue earning 8, while the best red cell seen will earn $5b$. When $b < 1.6$, the payoff for $5b$ is less than 8, and the blue cell stands pat. This situation raises the possibility of a mixed outcome, where some players play perpetual left, while other players play perpetual right. Indeed, were we to start with a single cluster of nine red cells and all other cells blue, the model would not move an inch. To see the latter point more clearly, I reran the model above with $b = 1.55$ and obtained the results in [Figure 3](#). The setup here is the same as before, save for the revised value of b , but the results change substantially. As predicted, we see a region in which the model converges to an outcome in which some players are playing left, while others play right.

$b = 1.55$ Highest Choice Rule

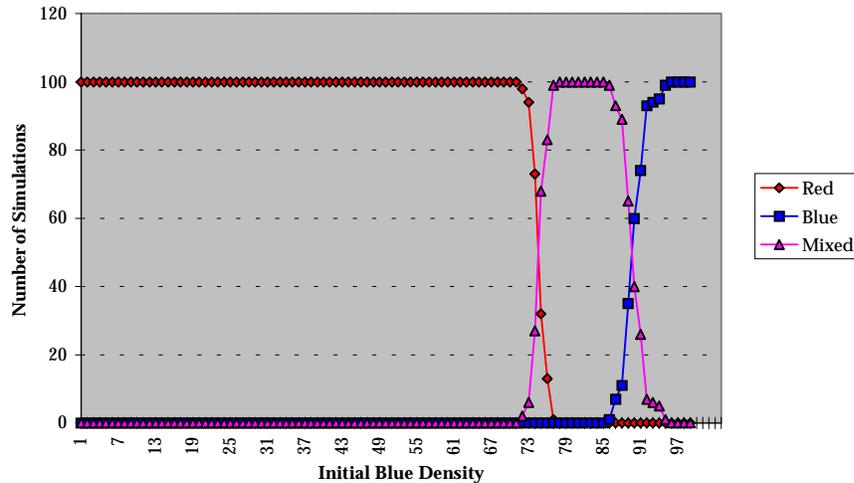


Figure 3

Figure 3 again graphs the number of all-red outcomes, all-blue outcomes and mixed outcomes against the initial blue densities. In broad terms, there are five distinct bands of behavior and two different phase transitions. For initial blue densities of 1 to 73, the model converges to the all-red equilibrium. The time to convergence is influenced by density—I do not set this chart out but it looks like the one in **Figure 2**—but we eventually get to the right steady state. We then reach our first phase transition: we move from the good all-red equilibrium to a mixed absorbing state, where the model locks on a mixture of left and right players. This transition occurs between initial blue density values of 74 and 76. For initial blue densities of 77 to 86, we converge to the mixed equilibrium. The second phase transition kicks in at an initial blue density of approximately 87 and is over by the time we get to 92. At that point, we move away from the mixed absorbing states to the bad all-blue equilibrium. Finally, for initial blue densities of 92 to 99, we converge to the blue equilibrium.⁷⁵

The existence of three different steady-state regions and two phase transitions is an important change from the prior analysis. Convergence on all-blue or all-red means that we eventually see

⁷⁵ I should note that at an initial density of 76, four of the one hundred simulations did not converge before reaching the then-applicable limit of two hundred and fifty generations.

society using only one norm. The good norm drives out the bad norm (or vice versa). We do see both norms in use out of equilibrium when $b = 1.65$, but only until we move to fixed, uniform play. In contrast, when we reduce b just slightly, we see a region in which we have two norms at work, in perpetuity. If our focus is on whether we will converge to the good norm, we should be reasonably confident that we will: so long as the initial density of players playing left is less than 74 percent, we converge to the good norm 100 percent of the time. Certainly this outcome represents a substantial drop from the results when $b = 1.65$, as we converged successfully in 100 percent of the cases up to an initial left density of 85 percent. But again, absent something that makes the left strategy particularly salient, we should expect to reach the good equilibrium.

Figures 4 through 8 set out runs of the same model with b set at, respectively, 1.35, 1.15, 1.14, 1.10, and 1.05. Remember that these values for b do not upset our rule that the all-right equilibrium is always better than the all-left equilibrium, but when $b = 1.05$, the differences are quite small. A quick glance at these figures reveals several facts. First, in each figure, we get three final state regions: all-red, all-blue, and mixed. Second, we see sharp phase transitions from region to region; change in the ultimate equilibrium is driven by a small fraction of the entire parameter space. Third, the location of these regions moves systematically as we reduce b . The right-most point where 100 percent of the simulations converges to the good norm decreases from 65 percent when $b = 1.35$ to 60 percent when $b = 1.15$; to 17 percent when $b = 1.14$; to 16 percent when $b = 1.10$; and then finally jumps up a bit to 17 percent when $b = 1.05$. Again, to return to the language of dynamic systems, the basins of attraction for the two point attractors change systematically as we reduce b . The basin of attraction of the good equilibrium shrinks steadily and then dramatically, while that for the bad equilibrium grows slowly over time.

b = 1.35 Highest Choice Rule

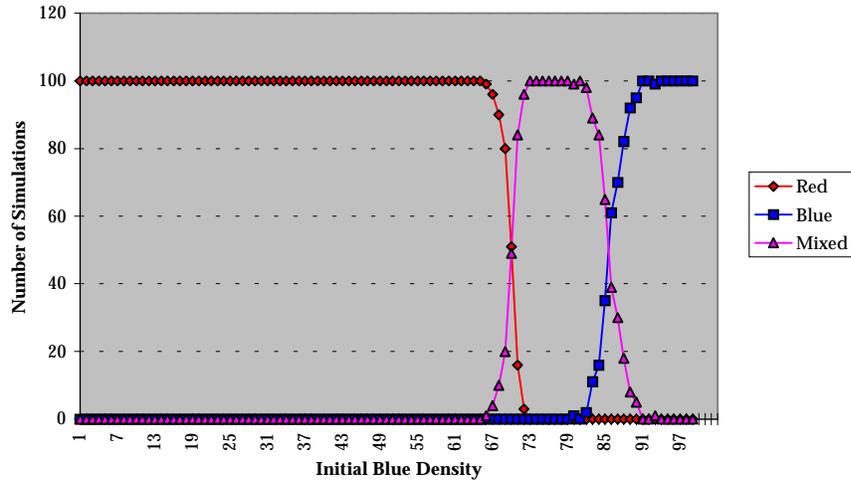


Figure 4

b = 1.15 Highest Choice Rule

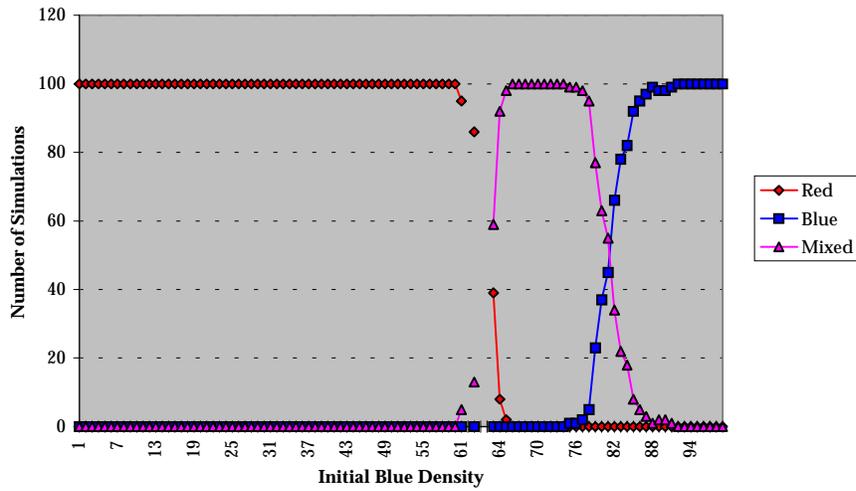


Figure 5

b = 1.14 Highest Choice Rule

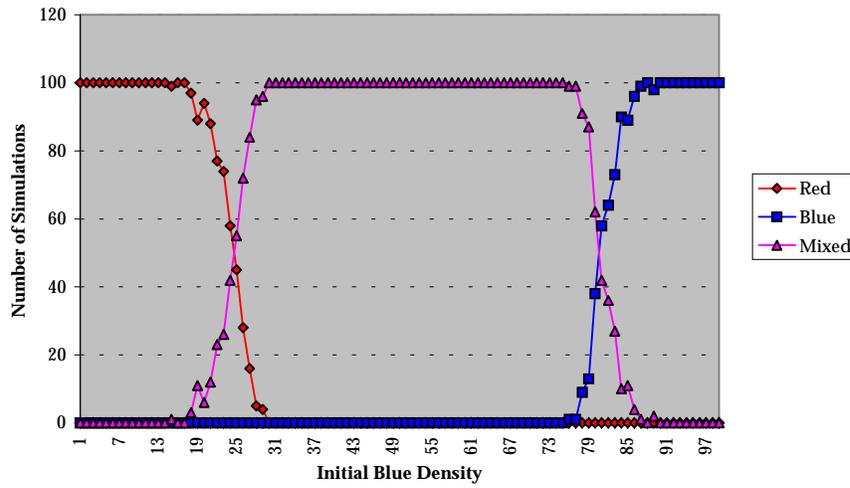


Figure 6

b = 1.10 Highest Choice Rule

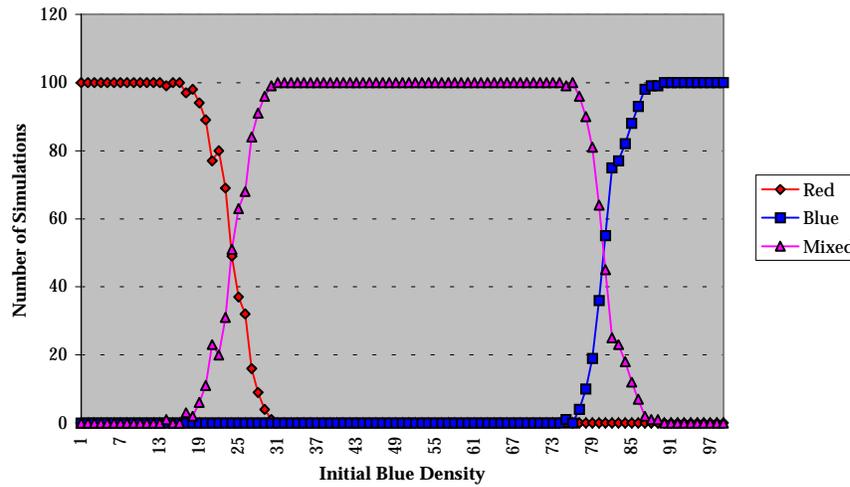


Figure 7

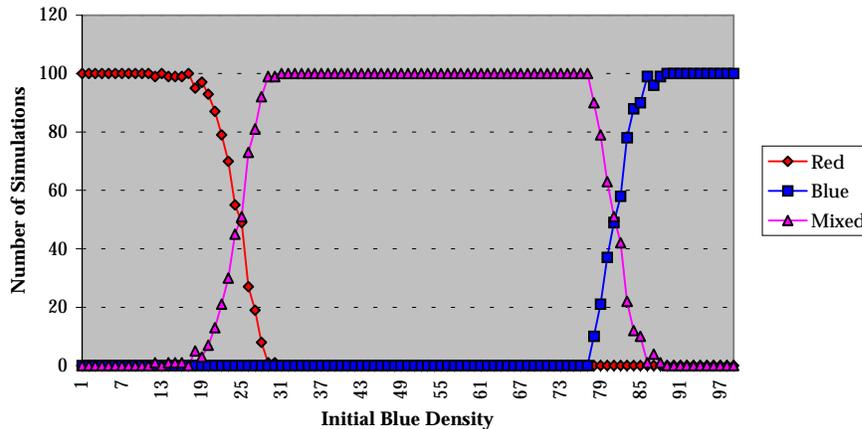
b = 1.05 Highest Choice Rule

Figure 8

As this analysis should make clear, we can remain fairly confident that the model will converge on the good norm so long as b is at least equal to 1.15. The chance that we will end up elsewhere, in one of the mixed play outcomes or the all-blue outcome, does rise as the value of b falls. But so long as not more than 60 percent of the players play left initially, we will converge on the good norm even if $b = 1.15$. This outcome again should be quite reassuring. If two norms are competing and one norm is better for everyone than the other by at least 15 percent, we will converge on the good equilibrium. Put differently, if the gains from getting to the right equilibrium are sufficiently large, we will get there. We could not get that result out of our free-standing coordination game, where we could only identify our two Nash equilibria and then punt. Moving to the good equilibrium when there is a substantial advantage to doing so instinctively seems right and it is comforting to see this result emerge in the model.

Note also the sharp break between 1.15 and 1.14. The probability of ending up in the good norm equilibrium in all cases plummets within this twilight zone of b values. The best bet here is that we will end up in a mixed play region. We will see both norms extant in the society, and perhaps in significant numbers. And this result holds as we move b toward 1. There are two natural questions. First, why the sharp change at this point? Second, are all the mixed outcomes identical, or are there meaningful differences for different initial densities, even if we know

that we end up in mixed play in 100 percent of the cases? Start with the second question and consider **Figure 9**:

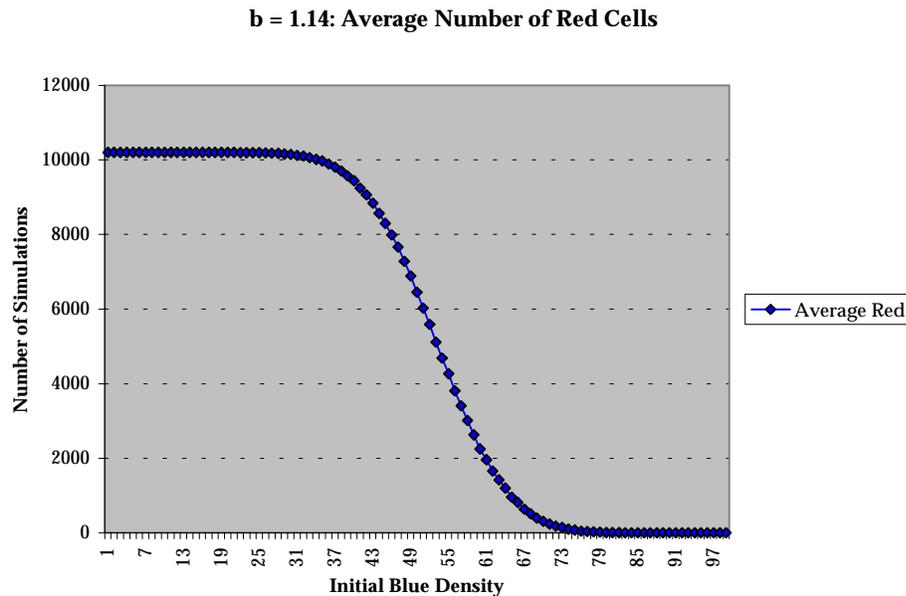


Figure 9

Figure 9 charts the average number of red cells in the eventual end state for the one hundred simulations for each initial density. Obviously, if all one hundred simulations converge to the all-red equilibrium, the average is the entire board of 10,201 cells. And, if each of the simulations converges to all-blue, there will be no red cells. The regions of interest are the two phase transitions and the mixed play region. As inspection of **Figure 9** makes clear, all mixed play outcomes are not created equal. Indeed, these outcomes systematically move from being mainly red, to being a fair mix of both blue and red, to being almost exclusively blue. Again, this outcome represents substantial and unsurprising path dependence.

Now consider the first question: why the sharp transition between 1.15 and 1.14? The simple answer is that 7 times 1.14 is 7.98, which is less than 8, while 7 times 1.15 is 8.05, which is more than 8. To see why this difference matters, consider a cluster of nine blue cells surrounded by a sea of red (this is the flip of our prior example, where we started with a cluster of nine red cells):

R	R	R	R	R	R	R	R	R
R	R	R	R	R	R	R	R	R
R	R	R	R	R	R	R	R	R
R	R	R	B	B	B	R	R	R
R	R	R	B	B	B	R	R	R
R	R	R	B	B	B	R	R	R
R	R	R	R	R	R	R	R	R
R	R	R	R	R	R	R	R	R
R	R	R	R	R	R	R	R	R

Diagram 6

The red cells are all rock-solid here, so long as $b > 1$. Each red cell either will see a red receiving $8b$ or be such a red, and the best blue cell that could be observed can do no better than 8. No red cell will switch.

Next consider the blue cells, and start with a corner blue as that is the one most likely to switch. A corner blue will see the center blue receiving a payoff of 8 and the best red cell—the one cater-corner to it—receiving a payoff of $7b$. That blue will stay blue if $7b < 8$, or $b < 8/7$. The corner blue thus remains blue at $b = 1.14$ and switches at $b = 1.15$. In other words, a cluster of nine blue cells cannot be invaded if $b < 8/7$, and a mixed play outcome will result. In contrast, if $b > 8/7$, the four corner blues will switch to red, and the remaining blues will soon follow in subsequent rounds.

What should we make of this analysis? In one way, it should be quite comforting. If we have “reasonable” initial blue densities and sufficient benefit from the superior equilibrium, we converge to the good equilibrium. These circumstances may be present in many cases. For example, if two norms or standards are competing at the same time, with both trying to emerge as the accepted convention, we should anticipate that the middle initial densities will be most important. In contrast, if norms are competing over time—if one standard is the convention, and the situation evolves so that a new convention is socially preferable—society will be at either end of the initial densities. The low values in the parameter space will give us a sense of how successful a new standard will be in displacing a preexisting standard. In that case, we should be much less sanguine that our model will get to the right outcome.

D. Variations on the Model

Altering basic features of the model allows us to evaluate if and how these features drive the results of the model. In this section, I will discuss four variations of the original model: (1) a different decision rule; (2) a different payoff neighborhood; (3) different information assumptions; and (4) an alternative approach to parallel decision processing. Take these variations one by one.

1. Different decision rule.

Thus far, I have used a particular decision rule without offering any particular justification for the choice. I believe it a plausible rule, but not the only rule. It might be worth exploring the choice of decision rules systematically, but, to provide just one source of comparison, consider the following rule. Suppose that instead of choosing the strategy with the highest payoff from the nine observed results, our players choose the strategy with the highest *average* payoff. This rule imposes a much more severe calculation burden, but not one that we should deem too daunting.⁷⁶ How do the results change with the new decision rule?

Figures 10 to 14 set out the results when b is set to, respectively, 1.65, 1.35, 1.15, 1.10 and 1.05. A quick look reveals that the basic shape of the results is quite similar, but the precise break points do change in important ways. The one change in shape that we do see is mixed play regions at all values of b , even when $b = 1.65$. Again, if we believe that mixed play is important, the revised decision rule validates that possibility. The other interesting change is that we sustain 100 percent convergence to the good norm for lower values of b . For example, when $b = 1.10$, we get 100 percent convergence even with initial blue densities of 54 percent. Recall that under the highest decision rule, when b was 1.10, we lost 100 percent convergence when the initial blue density exceeded 16 percent. This difference is substantial, not only in absolute terms, but relative to where we think we are likely to be. If you believe that at least 50 percent of the players should adopt the good strategy, the highest average rule provides

⁷⁶ To see an example of this, return to [Diagram 4](#) and focus on the decisionmaking of a corner red. Under the highest payoff rule, the corner red bases its decision on the payoff of the center red, as that will be the highest observed payoff. Under the highest average payoff rule, the corner red is red, obviously, sees three other reds, and sees five blue cells. The blue cells receive payoffs of $5, 6, 7, 6$, and 5 for an average payoff of 5.8 . In contrast, the observed red cells have payoffs of $5b, 8b, 5b$, plus the cell itself obtaining a payoff of $3b$. The average red payoff is $21/4b$ or $5.25b$. The corner red will switch if $5.8 > 5.25b$, or $b < 1.105$, and will stay otherwise.

real evidence that we will converge on the good norm even when the good strategy provides relatively insignificant benefits (10 percent). We lose this result when b drops to 1.05—100 percent convergence to the good norm occurs last at an initial density of 43 percent—but this change is still a substantial improvement over the original decision rule. Again, altering the decision rule changes the relative size of the basins of attraction for our equilibria.

$b = 1.65$ Highest Average

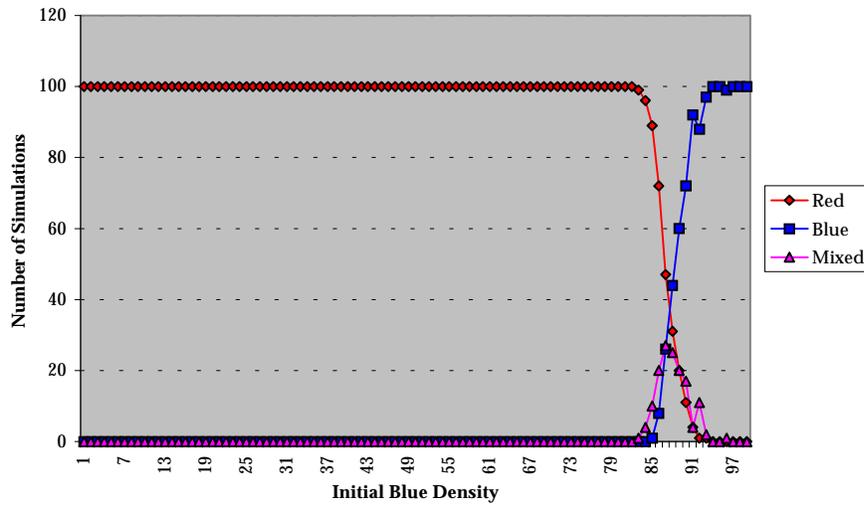


Figure 10

b = 1.35 Highest Average

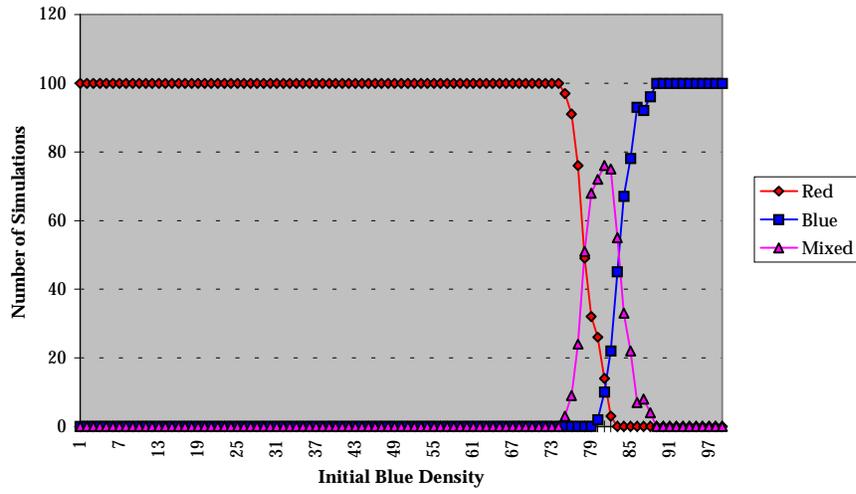


Figure 11

b = 1.15 Highest Average

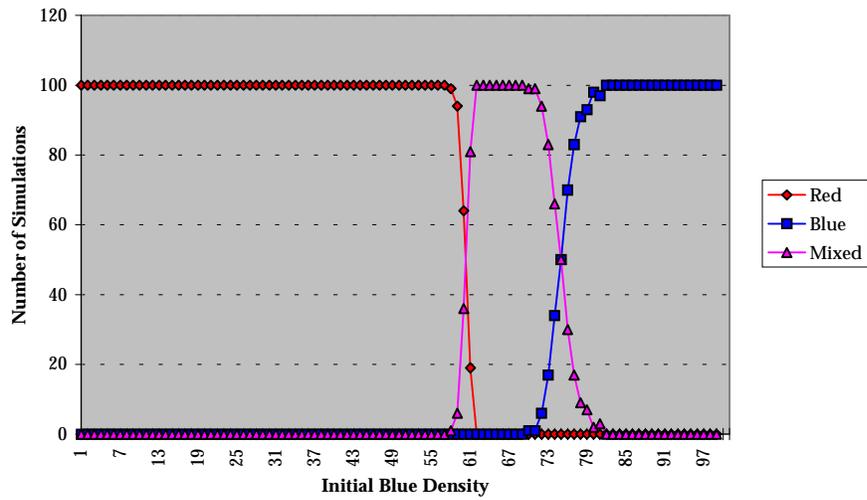


Figure 12

b = 1.10 Highest Average

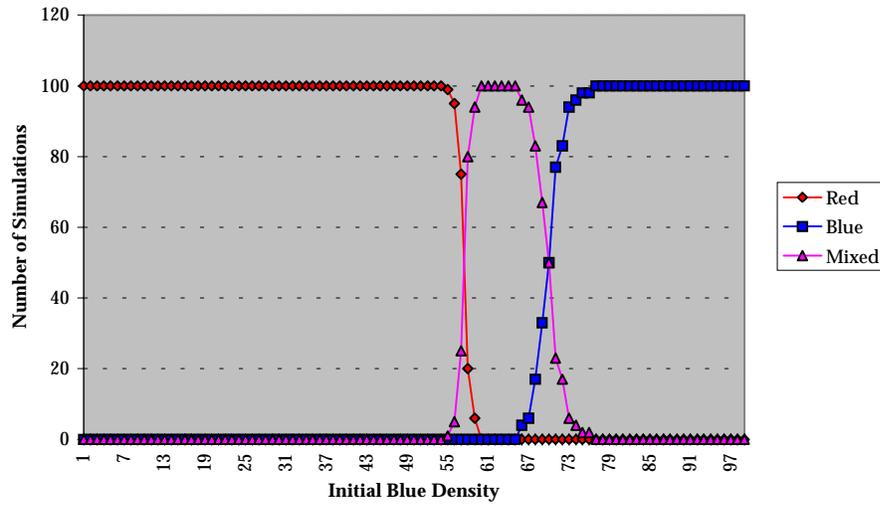


Figure 13

b = 1.05 Highest Average

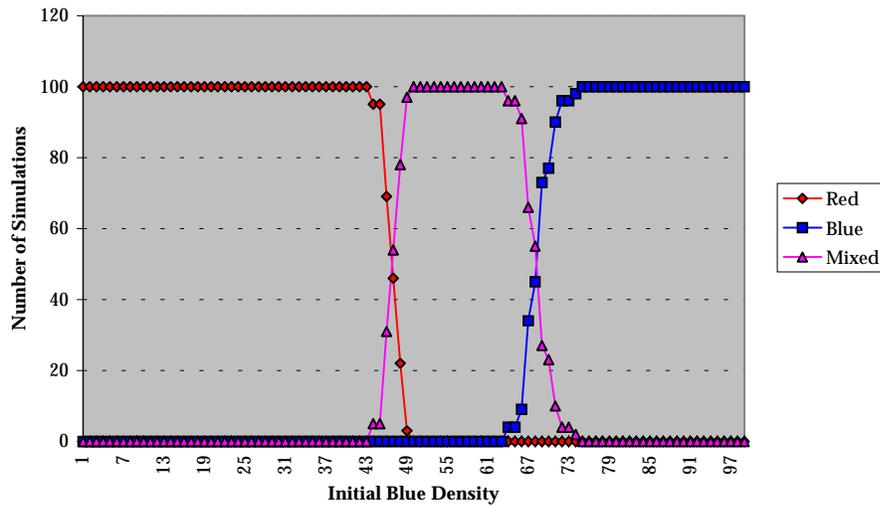


Figure 14

2. Different payoff neighborhood.

A second way to change the model is to switch payoff neighborhoods. We have focused on the Moore neighborhood, which likely drives particular features of the results. The significance of 1.6 and the breakpoint between 1.14 and 1.15 are both functions of having eight neighbors. As discussed before, the second prominent neighborhood used in the cellular automata literature is the Von Neumann neighborhood (see [Diagram 2](#)). This payoff neighborhood starts with a center cell and adds its North, South, East, and West neighbors. [Figures 15 to 18](#) set out results for the Von Neumann version of the model with b set at, respectively, 1.65, 1.35, 1.30, and 1.15:

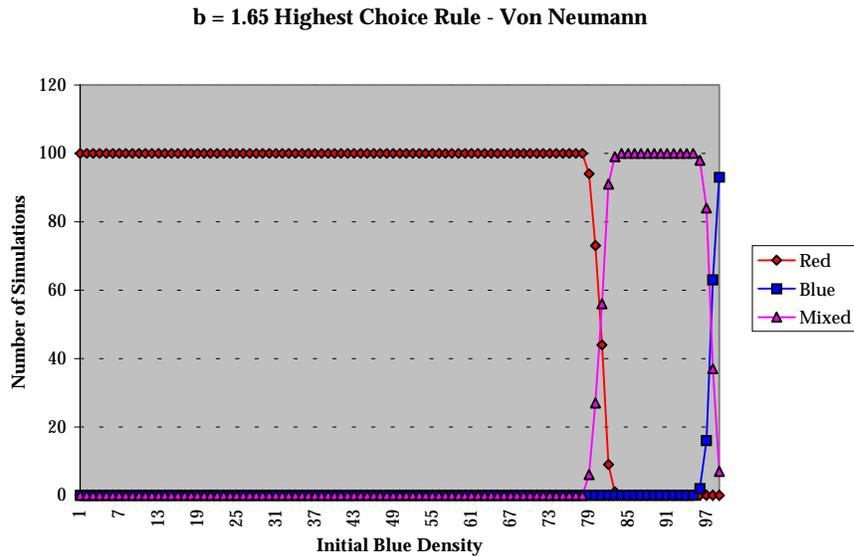


Figure 15

b = 1.35 Highest Choice Rule - Von Neumann

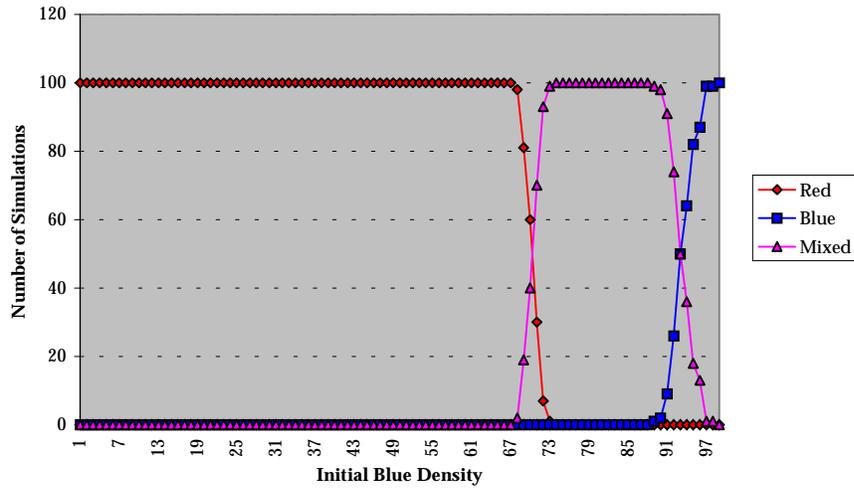


Figure 16

b = 1.30 Highest Choice Rule - Von Neumann

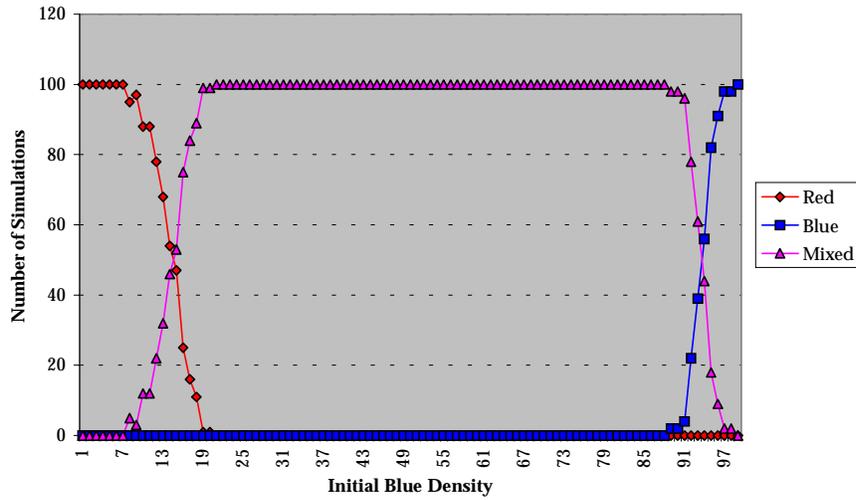


Figure 17

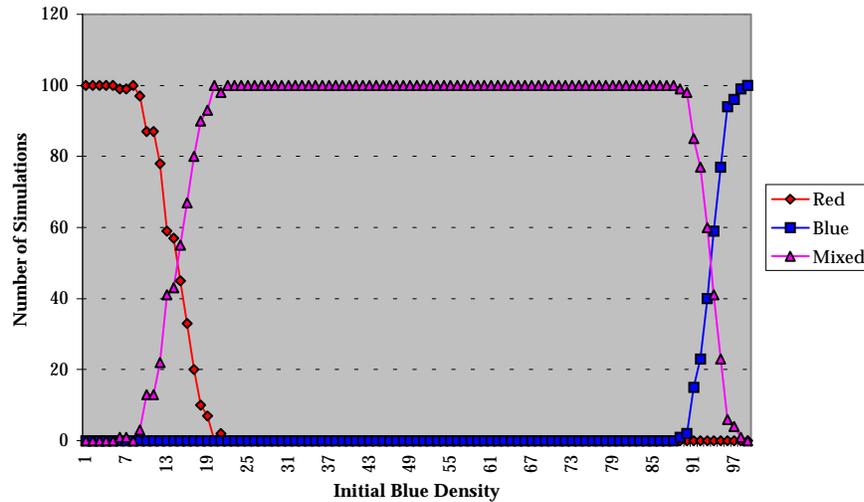
b = 1.15 Highest Choice Rule - Von Neumann

Figure 18

We do see important differences in the results. The mixed play region exists even when $b = 1.65$, and the last 100 percent convergence to the good norm drops from 85 percent in the original Moore version to 78 percent in the Von Neumann version. The results at $b = 1.35$ are almost identical for the two neighborhood versions. Note also that we get a break point at 1.33 rather than 1.14, so when $b = 1.30$, we find that a large mixed play region emerges. Again, this outcome is important, because it suggests that even with gains as large as 30 percent—with $b = 1.30$ the good norm is that much better—we cannot be confident that the model will converge on the good equilibrium.⁷⁷ Comparing the results of the Von Neumann and Moore neighborhoods therefore suggests that a bigger payoff neighborhood increases the chance that the model will converge on the good norm.⁷⁸

3. Different information assumptions.

To see a simple example of how a small change can alter the results of a model substantially, alter the information neighborhood. Recall that the payoff neighborhood is made up of the

⁷⁷ As to why $b = 1.33$ matters, see note 81.

⁷⁸ This result needs more analysis. Berninghaus and Schwalbe reach the opposite conclusion in a related model: “The smaller this reference group, the higher the probability that an efficient equilibrium will be reached.” See Berninghaus and Schwalbe, 29 *J Econ Behav & Org* at 79 (cited in note 44).

neighboring cells whose actions directly influence payoffs. The information neighborhood is the neighborhood that a given cell observes and provides information that players can use to formulate the choice of strategy in the next round. In all of the prior examples, the payoff neighborhood and the information neighborhood have been coextensive.

Now switch the assumptions. Set the payoff neighborhood as the Moore neighborhood and let the information neighborhood be the five-by-five cluster of twenty-five cells centered around the payoff neighborhood. Put differently, the payoff neighborhood starts with a single cell and extends out one cell in all directions. The information neighborhood starts with a center cell and extends out two cells in all directions. Payoffs are determined as before, but now each player chooses her next strategy based on the strategy yielding the single highest return in the twenty-five cells that the player observes. **Figures 19 through 21** reflect the runs with b set at, respectively, 1.65, 1.35, and 1.10:

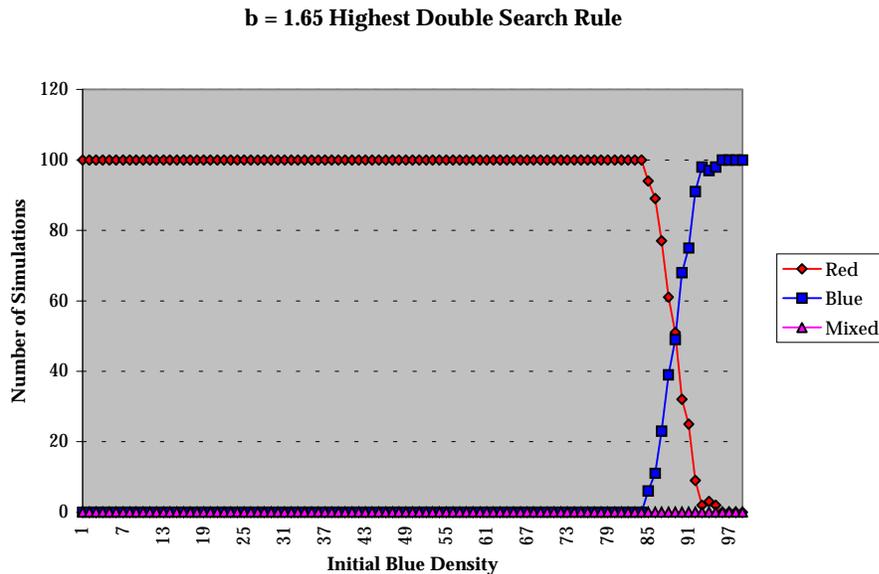


Figure 19

b = 1.35 Highest Double Search Rule

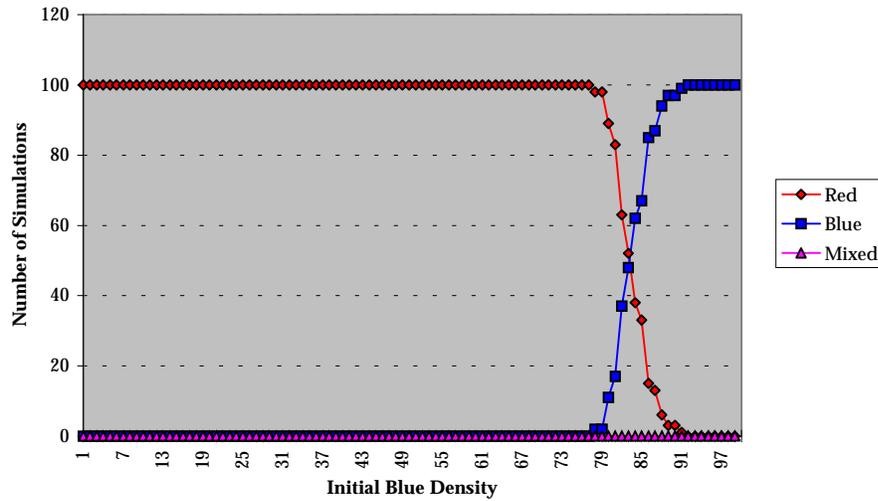


Figure 20

b = 1.10 Highest Double Search Rule

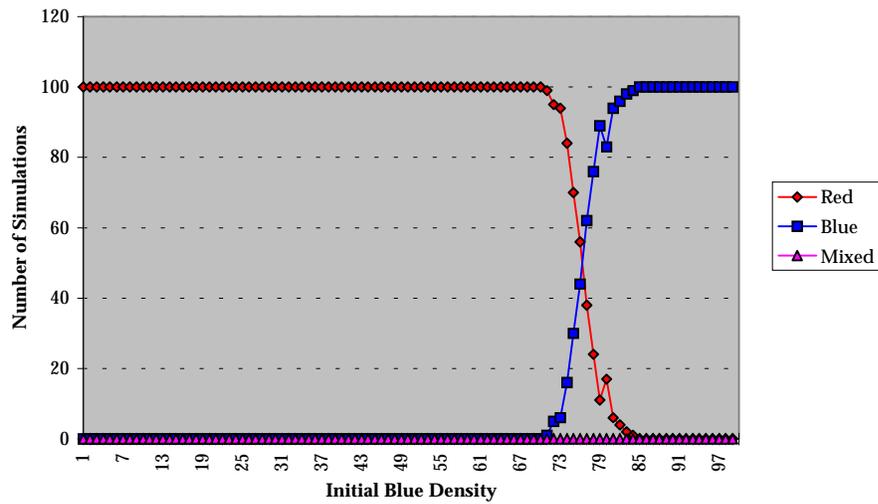


Figure 21

How does this model compare to our prior results? In the new results, when $b = 1.65$, there is little change, but when $b < 1.6$, the results change dramatically. We lose the mixed equi-

librium outcomes, and there is a single phase transition, moving from the superior equilibrium to the inferior equilibrium. The superior equilibrium is also reached much more often. In the original example, when $b = 1.10$, the region of 100 percent convergence on all-red ended at 16 percent, but with more information, that changes to 70 percent. Adding the additional layer of information enlarges the basin of attraction for the good equilibrium. A much larger set of initial conditions funnels into the good equilibrium.

Why the dramatic shift? Return to **Diagram 4**, where we had a cluster of nine red cells surrounded by blue cells. We noted before that the red cells would not change, so long as $b > 1$. Each red cell sees the center red receiving a payoff of $8b$, and no player can do better than that. Whether the neighboring blue cells turned over depended on how b compared to 1.6. If $b > 1.6$, the adjacent blue cells flipped over, but if $b < 1.6$, they all held firm. In the first case, the model would converge to the all-red equilibrium; but in the second, we reached a mixed outcome.

Now think about extending the information seen by the players. The cluster of nine red cells is as before: Each sees the center red receiving $8b$, and no one will switch. But the border blue cells now see the red cell in the center of the cluster of nine. When the blue cells could see out only one level, they saw only their immediate neighbors. Seeing out two levels brings the center red cell within view. The red cell does the best that anyone could do, and the blue cells shift. Note that this outcome holds regardless of how big b is, so long as $b > 1$.⁷⁹

We do not always reach the good equilibrium, nor is the size of the gain from getting to that equilibrium irrelevant. The graphs make clear that if we start with too many players playing the inferior strategy, we will converge on the inferior equilibrium. The size of the minimum cluster necessary to sustain growth is still determined by b . So long as we have at least one cluster of nine red cells, however, the model will converge to the right social norm whenever $b > 1$. Whether such a cluster exists is purely a question of probability and is independent of the size of b , but does depend on the size of the lattice itself.

There is one other point of interest. **Figure 22** sets out a comparison of the results for convergence to the good equilibrium as a function of the information available when $b = 1.65$:

⁷⁹ To see this pattern, play the video that starts with a single cluster of nine red cells and $b = 1.01$ at my [website](#) (cited in note 52).

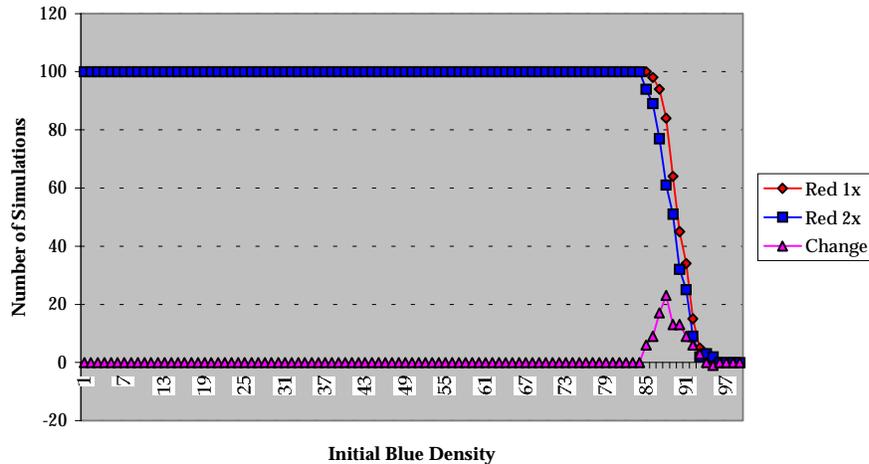
b = 1.65 Search Depth Comparison

Figure 22

The first two graphs—labeled “Red 1x” and “Red 2x”—just repeat the results from before when we set the information neighborhood equal to, respectively, the Moore neighborhood and the double Moore neighborhood. The third graph subtracts the number of times the model converged to the good equilibrium for the double Moore test from the single Moore test. Save for one density, we converge to the good equilibrium more frequently when we have less information. Believe it or not, when we give the players more information, we end up at the wrong equilibrium more often.

The double Moore results move to the inefficient blue equilibrium more often, sometimes substantially so (look, for example, at the spike in Figure 22 when initial blue density is 88). This result looks like a herd behavior in action,⁸⁰ but is less obviously tied to hidden information. Our players are not necessarily worse off with the additional information, as the double Moore model converges much more quickly (see Figure 23). This speed-

⁸⁰ Herd behavior arises when dispersed information is aggregated inappropriately. A chain of inferences arises that, while individually rational, leads the society—the herd—down the wrong path. In real life, you see this behavior frequently at traffic lights, where a few pedestrians trying to cross against the light lead an entire group into the street. See generally Abhijit V. Banerjee, *A Simple Model of Herd Behavior*, 107 Q J Econ 797, 798 (1992) (defining herd behavior as “everyone doing what everyone else is doing, even when their private information suggests doing something quite different”); Baird, Gertner, and Picker, *Game Theory and the Law* at 213-17 (cited in note 5).

ier convergence is a good thing, and we need to explicitly weigh faster outcomes against the possibility of more wrong outcomes.

b = 1.65 Search Depth Convergence Comparison

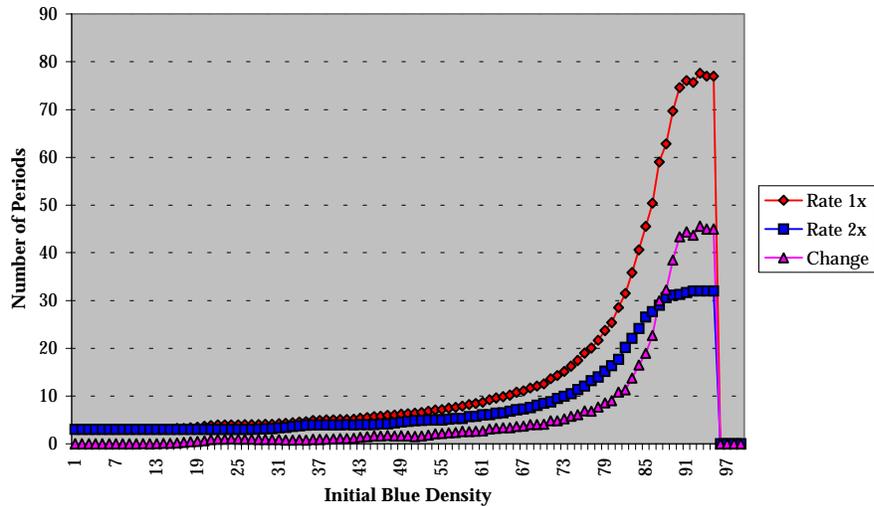


Figure 23

As a second example of how changes in information alter the results of our model, consider **Figures 24 to 26**. These runs use a Von Neumann payoff neighborhood and a Moore information neighborhood. Payoffs are therefore determined by the North, South, East, and West neighbors, but players receive information about strategy results from all eight of their immediate neighbors. The figures report results for $b = 1.65$, 1.35 , and 1.15 . Again, note how increasing information makes it more difficult for mixed play to be sustained. Recall (see **Figure 15**) that in the original Von Neumann run with $b = 1.65$, we had a substantial mixed play region, and we lost 100 percent red convergence at an initial blue density of 78 percent. The revised model (see **Figure 24**) has no mixed play region at all and converges successfully to the good equilibrium 100 percent of the time up to an initial left density of 89 percent.

b = 1.65 Highest Choice - Von N Payoff/Moore Info

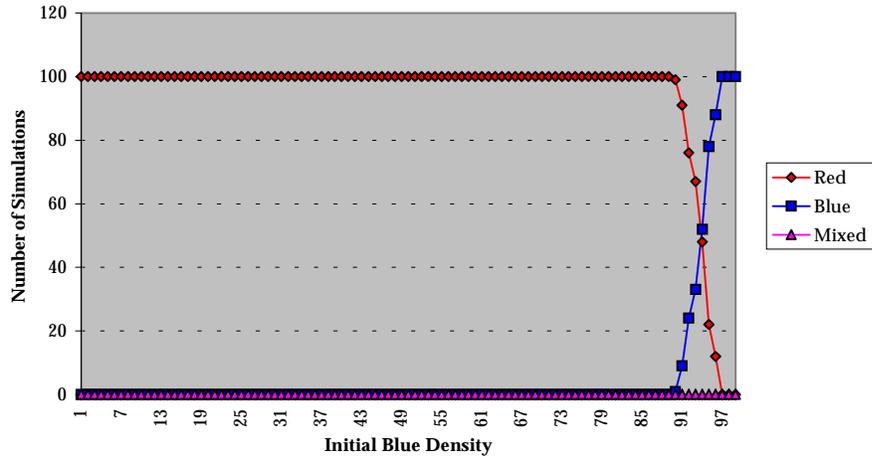


Figure 24

b = 1.35 Highest Choice - Von N Payoff/Moore Info

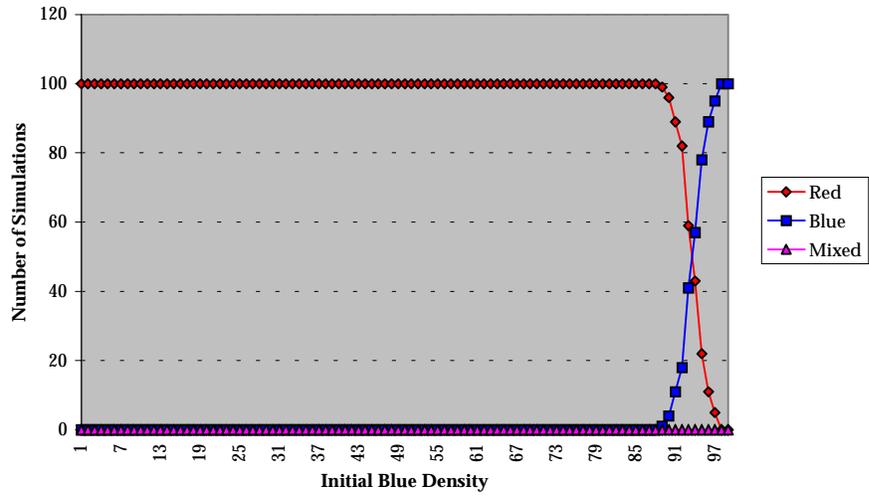


Figure 25

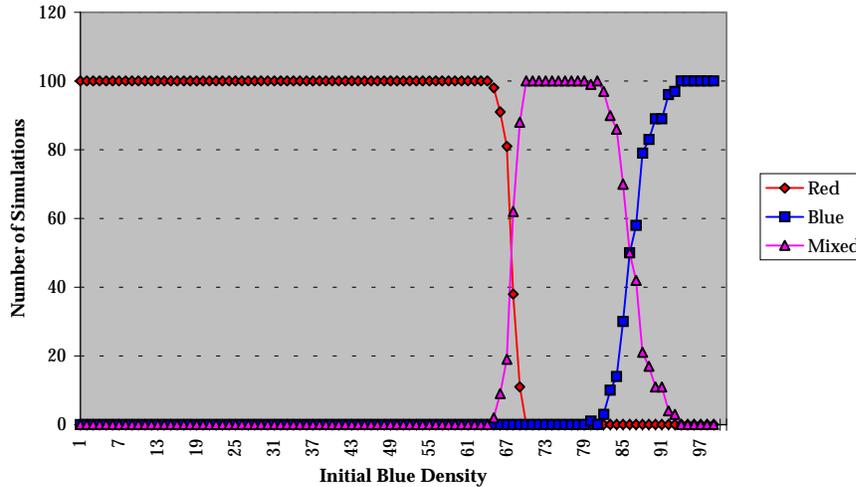
b = 1.15 Highest Choice - Von N Payoff/Moore Info

Figure 26

The same result holds at $b = 1.35$. Before (see Figure 16), we had mixed play and 100 percent red convergence through 67 percent. In the revised model (see Figure 25), no mixed play occurs and 100 percent red convergence runs through 88 percent. But these are not general results, as Figure 26 makes clear. With $b = 1.15$ and the Von Neumann/Moore fusion, there is a mixed play region, and we lose all-red convergence at 64 percent. This change in the basin of attraction for the good equilibrium is a sizable improvement over the purely Von Neumann run (see Figure 18), where all-red convergence was lost at 8 percent.⁸¹

⁸¹ If you have followed this analysis so far and care why the results change with the switch to the Moore neighborhood, the key point to note is that when $b > 1.33$, a red Von Neumann cluster will be a growth cluster when the model uses a Moore information neighborhood. When we start with five red cells in the form of a cross, none of the red cells will change. The four red cells forming the arms of the cross will observe the center red, who receives a payment of $4b$, the best possible payoff. None of the red cells will switch. Focus next on the four blue cells that complete the Moore neighborhood. Each of these cells now observes the center red. Note importantly that they would not have seen this red with Von Neumann vision, as the center red is on their diagonal and thus would be out of their sights. With the Moore information neighborhood, they too see the center red, and they switch over. This switch creates a cluster of nine red cells, surrounded by sixteen blue cells. The non-corner blue cells—all but four obviously—will each see one red cell receiving a payoff of $3b$, the center-edge red. The best blue cell that they observe will receive a payoff of 4, and the non-corner border blue cells will switch if $3b > 4$, or $b > 1.33$. This iteration drives the model round after round to the red equilibrium. To see the evolution of this model directly, play the video at my [website](#) (cited in note 52).

Finally, consider one last information comparison. Suppose that we let each player observe all of the other players' payoffs. Also assume that each player adopted the strategy with the single highest payoff that the player observed. As should be clear, each player will adopt the same strategy immediately, and the model converges to an all-blue or all-red absorbing state almost immediately. Figures 27 and 28 do just this for $b = 1.65$ and $b = 1.35$. The graph labeled "Global" charts the outcomes just described. These outcomes are compared with the outcomes from the double Moore information neighborhood. The results are quite close. The extended local information is almost equivalent to having information about the entire board, at least where each player decides based upon the single highest observed payoff.

$b = 1.65$ Global Information Comparison

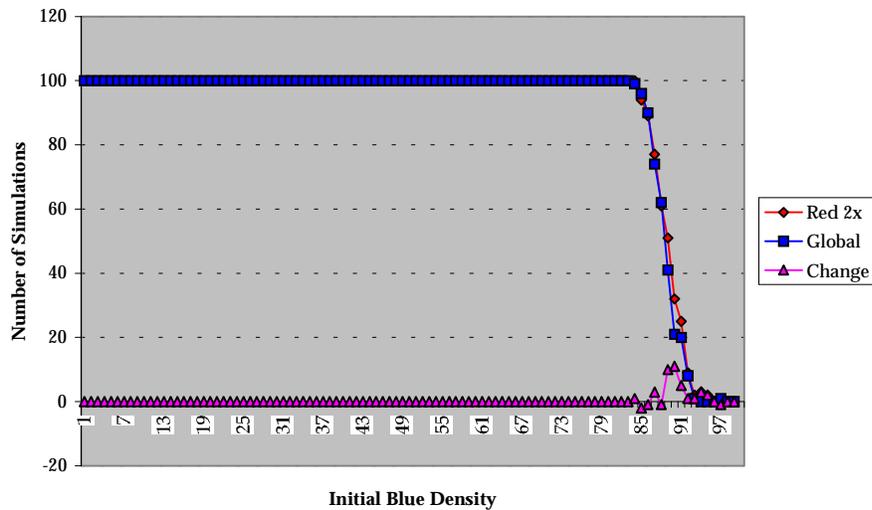


Figure 27

b = 1.35 Global Information Comparison

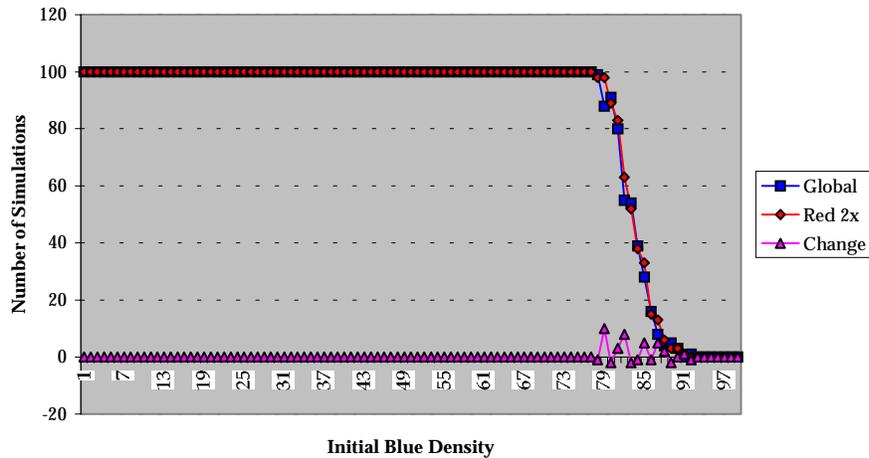


Figure 28

4. Parallel decision processing: SIMD v MIMD.

What appears to be the name of an obscure case, turns out to be an important issue in approaching parallel programming. “SIMD” is an abbreviation for single-instruction, multiple-data; “MIMD” is multiple-instruction, multiple-data. Putting technical niceties to one side, the key issue is whether all of the players remain in perfect sync. If you think of the players as executing a computer program—as they actually do in the computer simulations—does each player execute the same instruction at exactly the same time? In a SIMD scheme they do; in a MIMD scheme, they do not. (In reality, the setup is more complicated than this explanation, as we are using a single processor machine to simulate a multiple processor scheme, but this really is just a computational point.)

What turns on whether the players act exactly at the same time? Turn to [Figure 29](#) and compare it to [Figure 4](#):

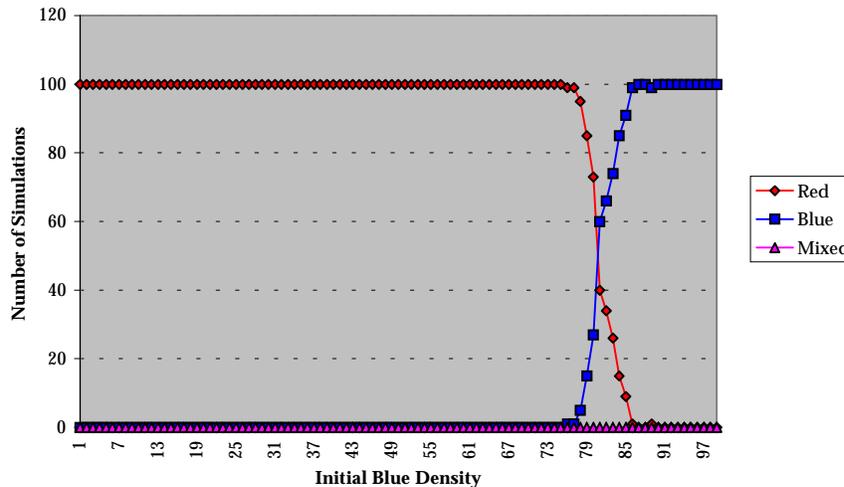
b = 1.35 Highest Choice Rule MIMD Version

Figure 29

Both models set $b = 1.35$; both use the same highest choice decision rule; and both use the Moore payoff and information neighborhoods. What differs are the results. The MIMD version (see Figure 29) has a single phase transition, no mixed play region, and 100 percent convergence on the good social norm at an initial density for the inferior strategy at 75 percent. In contrast, our original SIMD version (see Figure 4) has two phase transitions, a mixed play region, and the final 100 percent convergence on the good social norm at an initial inferior density of 65 percent.

These findings represent mixed to good results. We actually do better on reaching the good norm equilibrium and should be even more confident in our prior forecast that if $b = 1.35$, we will converge on the good equilibrium. That result, however, comes at the expense of losing the mixed play region.⁸²

⁸² The natural thing to do is to run more versions of the MIMD model to see whether the mixed play region reappears at lower values of b . Unfortunately, the single run presented in the paper took twenty-four days of computer time. And, there is substantial reason to doubt in any event the overall stability of the mixed play region. A large level of mutations in the players would tend to disrupt the boundaries that define the mixed play outcomes. Compare Kandori, Mailath, and Rob, 61 *Econometrica* 29 (cited in note 44).

IV. SPECULATIONS ON LAW: DECENTRALIZED RULEMAKING, NORMS AND SOCIAL MEANING

I want to be cautious about inferring too much about the relative roles of government and the private sectors in these models. I have said nothing about sources of market failure or of inadequacies in private ordering, nor have I addressed the range of infirmities associated with government action. With that caveat issued, I will nonetheless offer the following few speculative thoughts.

A. Echoing the Free-Standing Coordination Game

The basic coordination game has been the model of choice for illustrating the problem of coalescing around a Pareto-superior norm. I had very little to say about the original free-standing coordination game, but, for better or worse, quite a bit to say about the spatial version of this game. The extended analysis of the basic spatial game and possible variations reveals that the results do depend on particular settings of the model. The Von Neumann variation suggested that the basic problem of the coordination game might persist, as we faced mixed play outcomes with b as high as 1.33. To fail to converge on the good norm with this much at stake is disappointing. Still, the more general message has to be quite positive, at least for the case of simultaneous norm competition. The basic Moore model did quite well, and the revised versions did even better. Moving to the highest average decision rule and adding information typically led to better outcomes. Whether we think that the government should intervene depends on quite a lot—see the caveat above—but the model does suggest that the possible loss of value from inadequate coordination is naturally self-limiting.⁸³ This outcome makes me much less concerned about the problem seen in the free-standing coordination game.

B. The Importance of Phase Transitions for Policymakers

I find it striking how small ranges matter for the outcomes in these models. The vast majority of the parameter space is completely irrelevant to whether the model gets to the right steady state. I think, on the whole, that the small ranges should comfort us. If a policymaker were to try to switch this system

⁸³ Again, see the prior results to this effect in Kandori, Mailath, and Rob, 61 *Econometrica* at 44 (cited in note 44); Ellison, 61 *Econometrica* at 1066-67 (cited in note 44).

from one equilibrium to the other, the narrow band effect would work in our favor. If we were near the phase transition, a policymaker who mistakenly sought to switch from the red equilibrium to the blue equilibrium, and who took blue density as the variable on which to work, might do real damage. Such proximity, however, would arise from exceedingly bad luck—unless we have a story as to why these systems should gravitate toward the point of a phase transition.⁸⁴ Instead, we would expect to be at some distance from the transition, and one would hope that our policymaker would get quite discouraged before getting to the transition point. After all, our policymaker could push blue densities from 50 to 60 to 70 to 80 and accomplish nothing. Few policymakers would persevere for so long in the face of such apparent failure.⁸⁵

In contrast, the narrow band effect should work to our benefit if our policymaker sought to move us from the bad blue equilibrium to the good red equilibrium. In this case, a policymaker might correctly want to try to trigger a norm cascade, which looks a lot like swooshing down the phase transition to the good equilibrium. Here, little effort would be rewarded quickly, and the reward might be vastly disproportionate to the effort expended.

When might this analysis be relevant? Imagine norms competing over time. A norm or standard becomes entrenched at one time, and appropriately so: it represents the socially efficient outcome. We converge on the red equilibrium and everyone plays red. Things change. A new norm or standard is now superior to the old standard, but everyone is still playing the old standard. (Recorded video technology, with VCR tapes representing the old, locked-in standard, and DVD the new superior technology, may represent this situation today.)

Indeed, the law itself may help entrench a particular standard and thereby make it more difficult for our players to move to a new, superior equilibrium. For example, if custom is a good defense against a charge of negligence, there will be little reason for a new custom to evolve. Indeed, the status attached to the pre-existing custom further entrenches it against newcomers.

The models proper will not let us get from the old standard to the new standard, and even if a small number of folks start to

⁸⁴ We do get stories in the chaos literature of models that move to the edge of chaotic behavior, see Kauffman, *At Home in the Universe* at 26-29 (cited in note 65), but that is a far cry from where we are now.

⁸⁵ Of course, my guess is that politicians at either end of the spectrum would say that their worthy opponents routinely display such bullheadedness in promoting policies!

experiment with the new option, they may be too few and too dispersed to move the system to the new, better equilibrium. To track our charts, it is as if we have 98 percent or 99 percent playing the old, inferior standard, and only 1 percent or 2 percent playing the new standard. As the charts demonstrate, even in the face of a substantial improvement—remember that when $b = 1.65$, it represents 65 percent more value compared to 1—we might end up in the wrong equilibrium. Our policymaker now may be able to push us down the phase transition to the new equilibrium, where only a small shift in densities will be required.

C. Seeding Norm Clusters

If we take the model literally, the government may use a more direct route to effect the transition: seed norm or standard clusters. Given a cluster of the right size—for example, when $b = 1.65$ start with six red players clustered together in a sea of 10,195 blue players—the model will converge to the appropriate social equilibrium, even if the absolute number of players of the strategy in issue is almost zero. Look at the development of the model as seen in the six snapshots of its evolution in the color insert at **Figures P18 to P23**.⁸⁶

Gerry Mackie provides a striking example of the power of seeding norm clusters in an account of the end of footbinding in China.⁸⁷ Mackie argues that footbinding should be understood as a Schelling convention at work in the marriage market. China appears to have been locked into this convention for centuries, notwithstanding recognition of the harmful consequences of the practice. The practice, however, vanished in a generation. Mackie cites data showing, for example, that in Tingsien, 99 percent of the women were footbound in 1889, 94 percent in 1899, and virtually none in 1919. This dramatic shift is easily understood as a rapid shift from an inferior to a superior equilibrium, a norm cascade as we have described it.

What accounts for the change? Local missionaries in China established the first antifootbinding society in 1874. Families pledged that they would not footbind their daughters and that they would not let their sons marry the footbound. This local convention created sufficient density to make it self-sustaining—this is our norm cluster—and these clusters grew until the old

⁸⁶ To see this evolution directly, play the video at my website (cited in note 52).

⁸⁷ Gerry Mackie, *Ending Footbinding and Infibulation: A Convention Account*, 61 *Am Sociological Rev* 999, 1015 (1996).

convention was overrun. This is a dramatic example of the power of seeding norm clusters, but it also emphasizes that the government need not play a unique role in creating these clusters. Any number of groups can play this role: the government certainly can, but so can charities and for-profit entities.⁸⁸

This approach is probably more important in cases in which the established, but appropriate, norm or meaning changes, and we need to navigate from the formerly appropriate norm to the new norm. Norm seeding is a low-risk strategy. If the government seeds an inefficient cluster, it will die, and little will be lost. If the new norm is superior to the old norm, however, the artificially created norm cluster will thrive and spread. This analysis suggests that the government should embrace test policies or norms or take steps to foster social meanings in particular local contexts as a way of testing whether a superior approach can take root and spread.

D. Instruments for Policymakers

Finally, the analysis of the variations on the model suggests possible instruments to facilitate adoption of good norms. These are all directed at expanding the basin of attraction for the superior equilibrium, so that the initial starting conditions are less likely to influence the ultimate outcome of the system. First, larger payoff neighborhoods appear to do a better job of converging on norms that represent modest improvements (say, $1.15 < b < 1.33$).⁸⁹ As discussed before,⁹⁰ creating instrumental connectedness among individuals within a neighborhood may be difficult, but this model suggests that success yields substantial returns. Second, more refined decisionmaking—averages rather than a single, highest value—also supported success for lower values of b . Third, additional information usually increases the likelihood of success for the good norm. A strategy of providing information from beyond the boundaries of the payoff neighborhood generally increases the chance of converging on the good norm.

⁸⁸ Anti-dueling societies played a similar role in the decline of dueling in Europe, see V.G. Kiernan, *The Duel in European History* 216-17 (Oxford 1988), cited in Posner, 144 U Pa L Rev at 1740 n 108 (cited in note 8); and in the Old South, see Jack Kenny Williams, *Dueling in the Old South: Vignettes of Social History* 64 (College Station 1980), cited in Warren F. Schwartz, Keith Baxter, and David Ryan, *The Duel: Can These Gentlemen Be Acting Efficiently?*, 13 J Legal Stud 321, 329 n 27 (1984).

⁸⁹ But see Berninghaus and Schwalbe, 29 J Econ Behav & Org 57 (cited in note 44).

⁹⁰ See [Section II.A.2](#).

CONCLUSION: LIMITS AND FUTURE DIRECTIONS

The recent interest in norms and law almost certainly dates from Bob Ellickson's important work on Shasta County.⁹¹ This work started as an inquiry into the Coasean Irrelevance Proposition and emerged as a separate Ellicksonian Irrelevance Proposition. The bastardized version of the Coase result states that the allocation of property rights is irrelevant, as parties will recontract efficiently. Law plays very little role, at least in a world of low transactions costs. Now Coase himself would probably contend that the point of *The Problem of Social Cost* was precisely that transaction costs were important, and that the role for law needed to be understood in a particularized context of pre-existing transaction costs. In contrast, Ellickson's investigation of cattle in Shasta County led him to conclude that law was irrelevant when robust local norms could evolve.⁹² Neighbors did not know the law, did not use it to their advantage or disadvantage, and resolved their disputes in full sunlight—no bargaining in the shadow of the law for these folks. Law simply did not matter in a community with well developed norms.

More recent work takes as a given that norms, social meanings, and social roles matter enormously, and turns to the role that government might play in shaping these essential features of society.⁹³ The central concern of this work is that too often society will end up with weak norms, or, even worse, norms that are affirmatively harmful. Individuals who recognize the problem will be trapped and will lack a mechanism to move the collective to the superior norm.

This is certainly possible, but the computer experiments described here suggest local interactions will often lead to convergence on the superior norm. The benefits obtained by clusters of individuals who successfully embrace the better norm will often lead that norm to be propagated throughout the entire population of players. Again, this is not to say the government is irrelevant. The simulations identify at least three policy instruments of interest—the scope of local connectedness (my payoff neighborhoods); the information available to the players (the information neighborhoods); and the manner in which individuals process available information (the decision rule)—plus a strategy of seeding norm clusters so as to perturb an existing equilibrium to test whether a superior equilibrium will take root and spread.

⁹¹ See Ellickson, *Order without Law* §§ 1-6 (cited in note 13).

⁹² *Id.*

⁹³ See Section I.B.

The current paper suffers from at least five weaknesses, all of which suggest directions for future work:

- The players do not contest the choice of norm. How many potential norms benefit everyone identically? In contrast, how often will the norm chosen have substantial consequences for individual players (or player types) even if all benefit from a single norm? The traditional Battle of the Sexes is exactly this situation: the players want to coordinate on a single choice but each cares about the particular choice made.⁹⁴ Migration will also become an important feature of these models, as players may reduce the conflict over norms through separation by type.
- Competition between norms has been restricted to only two norms at a time. As suggested before, this model may capture many social contexts accurately, but it almost certainly does not track all commercial dealings.⁹⁵
- My models are awfully static for dynamic models. There is no baseline of change, either through replacement of players or spontaneous mutation or experimentation. The zone of mixed play may not be as sustainable when baseline change is introduced.
- The spatial game setup itself tends to obscure the way that information is transmitted in these models. There are no holes in the lattice, meaning that there is a player at every spot. Connectedness may matter in important ways. Full connectedness maximizes the chances that a good strategy will propagate throughout the entire lattice. Separation of clusters would introduce physical barriers to the spread of a good strategy, and might create situations in which both strategies would thrive in physically distinct locations. Whether a good strategy would percolate throughout the entire lattice will undoubtedly depend on how connected our lattice is. At the same time—and this works in the other direction—the players in the spatial game do not move about. Movement serves as another way in which information is transmitted—in which strategies are spread. A richer model

⁹⁴ For a discussion of the “Battle of the Sexes,” see note 49.

⁹⁵ See text accompanying note 7. Moving to a continuous strategy space has been shown to have substantial consequences for the results in spatial games. See Bernardo A. Huberman and Natalie S. Glance, *Evolutionary games and computer simulations*, 90 Proc Natl Acad Sci 7716, 7717-18 (1993) (showing that the results of digital simulations regarding territoriality and cooperation in the prisoner’s dilemma differ greatly when time is continuous rather than discrete).

would incorporate holes in the lattice (and thus explicitly introduce something akin to networks and the extent of connectedness) and allow movement by the players. This would allow a more refined analysis of information and strategy transmission.

- Finally, there is no cost of switching strategies in these games. In real life, switching strategies is costly (have you switched from a Macintosh to a Wintel machine lately?). Friction of this sort would certainly slow down the move to new strategies. Whether it would alter the outcomes in other important ways awaits more work.

As this suggests, this article is just a first step down a path of uncertain length. The silicon models set forth here make it possible to test the importance of a variety of factors for the evolution of norms. Given the pervasive role that norms play in shaping the contexts in which we act, it is hardly surprising that norms have become the focus of so much attention. It may be possible for the law to play a role in shaping norms and thereby alter the backgrounds against which much activity takes place. Whether we think legal intervention in norms is appropriate should turn on whether we believe that private ordering will result in inefficient or positively harmful norms. The message of this article, where individuals have a shared interest in the norms, is that large-scale intervention is unjustified, as individual decisions turn out to aggregate nicely and to coalesce around the appropriate norms. Small-scale intervention—norm perturbation—may be appropriate, and a strategy of norm seeding may be an effective way for the government to test at little cost whether a norm improvement should be and can be effectuated.

[Click here to return to text.](#)

[Click here to return to text.](#)