University of Chicago Law School

Chicago Unbound

Coase-Sandor Working Paper Series in Law and Economics Coase-Sandor Institute for Law and Economics

1997

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

Randal C. Picker
dangelolawlib+randalpicker@gmail.com

Follow this and additional works at: https://chicagounbound.uchicago.edu/law_and_economics

Part of the [Law Commons](#)

Recommended Citation

Randal C. Picker, "Simple Games in a Complex World: A Generative Approach to the Adoption of Norms" (Coase-Sandor Institute for Law & Economics Working Paper No. 48, 1997).

This Working Paper is brought to you for free and open access by the Coase-Sandor Institute for Law and Economics at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

SIMPLE GAMES IN A COMPLEX WORLD: A GENERATIVE APPROACH TO THE ADOPTION OF NORMS

Randal C. Picker[†]

A key issue for law is defining its limits, understanding what the boundaries of law are and need to be. When must we regulate? When instead will behavior coalesce in an appropriate way without the intervention of law? These are Big Questions, as the recent explosion in the literature on norms— best evidenced by nearly 700 pages of new prose in the May, 1996 *University of Pennsylvania Law Review*— makes crystal clear. I don't purport to answer these questions here. Instead, I bring to bear a particular methodological approach to these situations, one that may serve as a substantial step beyond the formal tools often used by law professors in examining these questions. The hope is for a much richer feel for the possibilities and risks in these situations.

We should be clear upfront about what this paper will and will not generate. This is a paper about norm competition, about 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 social norms. I will generate a wide variety

[†] Professor of Law, The University of Chicago. I thank the Sarah Scaife Foundation and the Lynde & Harry Bradley Foundation for their generous research support; Dick Craswell, 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's *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 foldout that sets out snapshots of these dynamics. There is also a CD-ROM version of the paper 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 on my website at <http://www.law.uchicago.edu/Picker/Aworkingpapers:norms.html>.

of patterns of outcomes that might result in norm competition. What I will not generate, though, are the norms themselves. These will be taken as simply given and wholly outside the formal model. It is possible that we might create these norms in models of the sort described in this paper, but I confess to real skepticism. Understanding where we get competing norms will almost certainly require much painstaking investigation into particular institutions and situations.¹

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 is quite understandable: these games are tractable and provide a framework that is now well-understood. Nonetheless, simplicity is both vice and virtue. These models seem almost naked, stripped of a meaningful strategy space and shorn of the multiplicity of players that characterizes 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, as described in more detail below, there has been an explosion in computer modeling of interactive situations. Many labels are associated with this work—“complexity,” “artificial life,” “artificial societies,” “agent-based modeling” and “massively parallel microworlds” all come to mind. Yet the common elements of this work are an emphasis on dealing with substantial numbers of actors, specifying rules for their actions, and letting the system rip to see what happens. Results emerge and patterns form and the system organizes on its own. The idea of self-organization, which is central to this work, is hardly new. Hayek emphasized this idea,² and Adam Smith’s invisible hand is the defining metaphor of self-organization. Nor is it a new idea to use computers to test notions about basic social phenomena. Robert Axel-

¹ For recent examples, see Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code’s Search for Immanent Business Norms*, 144 U Pa L Rev 1765 (1996); J. Mark Ramseyer, *Products Liability through Private Ordering: Notes on a Japanese Experiment*, 144 U Pa L Rev 1823 (1996).

² See A.I. Ogus, *Law and Spontaneous Order: Hayek’s Contribution to Legal Theory*, 16 J L & Soc’y 393 (1989).

rod's classic *The Evolution of Cooperation*³ did exactly that more than a decade ago. Nonetheless, recent changes in computer modeling techniques make it possible to treat the computer as a laboratory to run experiments in self-organization, 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 rationality framework of the sort used by economists, but will offer no sense of how empirically important the phenomenon really is. A given result may be quite brittle, obtainable only if the key parameters are tuned just so, but lost if the values don't 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.

To preview the conclusions, when there are shared values about norms, under a broad set of assumptions, my model societies exhibit strong self-organization. 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 reached or that we cannot influence whether the good equilibrium obtains. 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 information available to neighbors in making decisions, and the rules they use to assess the information available to them. 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. Indeed, there is a funnel for each possible outcome; making the mouth of the funnel for the good equilibrium relatively larger— expanding its basin of attraction— emerges as an important way for the government to implement policy.

³ Robert Axelrod, *The Evolution of Cooperation* (Basic Books, 1984).

In contrast, the results suggest that we should be less sanguine about sequential norm competition, as occurs when a new norm arises to compete with an old, entrenched norm. There is good reason to think that the old norm will continue notwithstanding that its useful life has expired.⁴ Jumpstarting the transition—causing a norm cascade⁵—might be accomplished through the seeding of norm clusters. The government— or for, that matter, charities, for-profit enterprises or you and me— would encourage experimentation; if the conditions were right, the norms seeded would take root and spread, and we would successfully transition from the old norm to the new norm. If the old norm really should continue, the experiment fails, and very little is lost.

The consequence of all of this is that the collective action problem faced in norms and social meanings has, to some extent, been overstated. I no longer lose sleep over this problem when we have simultaneous norm competition between two competing norms. That said, while it may be realistic in many settings— especially social settings— to speak of two norms perhaps as a binary on/off or yes/no choice, in other contexts, perhaps in torts and commercial settings, more norms will compete simultaneously, and my models will say nothing about this.⁶ And competition over time still remains a genuine problem, though to be sure about that we need 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 framework of the analysis. Section III discusses the simulation results. Section IV ties these results to

⁴ 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* 208-13 (Harvard 1994); see also Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 Va L Rev 757 (1995).

⁵ Cass R. Sunstein, *Social Norms and Social Roles*, 96 Colum L Rev 903, 909 (1996); compare Sushil Bikhchandani, David Hirshleifer, and Ivo Welch, *A Theory of Fads, Custom, and Cultural Change as Informational Cascades*, 100 J Polit Econ 992 (1992).

⁶ For an argument based on cultural evolution that suggests that commercial norms won't be optimal, see Jody S. Krause, *Legal Design and the Evolution of Commercial Norms*, 26 J Legal Stud (forthcoming, 1997).

approaches that the law might take vis-à-vis norms. Section V concludes the paper, discusses key limitations of this paper, 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's Friday, 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 don't 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 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 don't want to be seen as following what might be seen as the new political orthodoxy, but at the same time, you also don't 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, as 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 prenuptial agreements, where asking first could be seen as a sign of doubts about

the impending marriage, and moving to colorblind hiring unilaterally in a community dominated by discrimination norms.⁷ A social norm may exist that will resolve these situations in ways that benefit all interested parties. This norm could easily change over time or be subject to geographical or class variation.⁸

The second situation is more complex. It demonstrates clearly that norms can evolve and presents a clear example of the idea of, to use Cass Sunstein's term, a norm entrepreneur.⁹ Who use 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 from 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 favors.

The third situation might be seen as just a problem in signaling theory,¹⁰ but can also be understood as embedded in a web of social roles and social norms. Mothers are expected to be quite involved with their children, Father in the '90s increasingly so, so 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 given rise to a substantial literature. In addition, I need to establish the context in which I will create my computer simulations. To do this, I will touch upon the literatures concerning norms and game theory; social norms, social meanings and collective action problems; custom; social learning and social computa-

⁷ See Eric A. Posner, Law, Economics and Inefficient Norms, , 144 U Pa L Rev 1697, 1730 (1996).

⁸ Indeed, the daily arbiter of American business life claims that this norm has shifted recently in favor of drinks at lunch, so enjoy. See Stephanie N. Mehta, Lunch Hour is Becoming Happy Hour Again, Wall Street Journal, September 23, 1996, pB1.

⁹ Cass R. Sunstein, *supra* note 5 at 909.

¹⁰ See Phillip Aghion and Benjamin Hermalin, Legal Restrictions on Private Contracts Can Increase Efficiency, 6 J L Econ & Org 381 (1990); see Baird et al, *supra* note 4, at 142-45.

tion; agent-based computer simulations; and, finally, evolutionary and spatial games.

Norms and Game Theory. Simple two-by-two games provide much of the formal apparatus at work in the norm theory literature. Robert Sugden's masterful book *The Economics of Rights, Co-operation and Welfare*¹¹ establishes the two-by-two game as his central tool as he studies the evolution of conventions and the rise of spontaneous order. Ellickson devotes an entire chapter to game theory, and sustains much of his theoretical analysis of norms in the two-by-two framework.¹² As he makes clear upfront, 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."¹³ A survey of recent articles makes clear that game theory, and, in particular, freestanding two-by-two games remains the formal vehicle of choice for understanding norms.¹⁴

Social Norms, Social Meanings and Collective Action Problems. Recent theoretical work on social norms and social meaning defends a substantial role for government. 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.¹⁵ He focuses on the context in which individuals make choices and identifies the important if not

¹¹ Robert Sugden, *The Economics of Rights, Co-operation and Welfare* (Basil Blackwell 1986).

¹² Robert C. Ellickson, *Order without Law*, ch. 9 (Harvard, 1991).

¹³ *Id.* at p1.

¹⁴ See, e.g., Robert D. Cooter, Decentralized Law for a Complex Society: The Structural Approach to Adjudicating the New Law Merchant, 144 U Pa L Rev 1643, 1657-77 (1996); Posner, *supra* note 7, at 1713-19; Jason Scott Johnston, The Statute of Frauds and Business Norms: A Testable Game-Theoretic Model, 144 U Pa L Rev 1859 (1996); cf. Rajiv Sethi, Evolutionary stability and social norms, 29 J Econ Behav & Org 113 (1996); Peter H. Huang and Ho-Mou Wu, More Order without More Law: A Theory of Social Norms and Organizational Cultures, 10 JL Econ & Org 390 (1994).

¹⁵ Sunstein, *supra* note 5.

pervasive role played by social norms in defining that context. These norms will change the cost of a possible choice and thus will exert a substantial influence over choices seen. For example, as Sunstein notes, social norms on smoking have changed over time. 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 this, like the direct cost of buying cigarettes or the health costs of cigarettes, should reduce the amount of smoking. 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. This raises the specter of a collective action problem, that we will have no way for society to coalesce around a beneficial norm.¹⁶ In that framework, direct intervention by the government in norms–norm management in Sunstein’s phrase– appears to be a plausible response.

Larry Lessig has emphasized the collective action problem as well in his defense of the regulation of social meaning.¹⁷ 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 examples– might go completely unnoticed in France. An individual will be essentially powerless to alter the prevailing meaning associated with a particular act. She may take steps to mute that meaning, but all of this will be in a context where she has already spoken through her act. In this framework, it is certainly imaginable that society will get stuck on a destructive convention.¹⁸

¹⁶ See also the related discussion of coordination problems as a justification for legal intervention in Cass R. Sunstein, *After the Rights Revolution* 48-55 (Harvard 1990).

¹⁷ See Lawrence Lessig, *The Regulation of Social Meaning* 62 U Chi L Rev 943, 991-1016 (1995); Lawrence Lessig, *Social Meaning and Social Norms*, 144 U Pa L Rev 2181 (1996).

¹⁸ See also Dan M. Kahan, *Social Influence, Social Meaning and Deterrence*, 83 Va L Rev (forthcoming, 1997).

Custom. There is a second norms literature that examines these issues under the guise of custom, often in commercial settings¹⁹ or to help evaluate reasonable behavior in the law of torts.²⁰ We used to say custom 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. Custom might be used to describe practices that have legal significance. 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 practice are, and that can be a difficult problem.²¹ In any event, nothing that I do in this paper distinguishes as such between custom and norms.²²

Social Learning and Social Computation. This work is also related to the literature on social learning or social computation. This work examines the circumstances under which widely-held information will be aggregated efficiently so that the right social outcome is reached. Imagine, for example, that there is 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 is a work in social learning or social computation.²³

¹⁹ See Bernstein, *supra* note 1; see also Richard Craswell, Do Trade Customs Exist?, *The Jurisprudential Foundations of Corporate and Commercial Law* (forthcoming, Cambridge, 1997); Kraus, *supra* note 6.

²⁰ See 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).

²¹ See Craswell, *supra* note 19; see, e.g., *In re Tolona Pizza Products Corp.*, 3 F.3d 1029 (7th Cir. 1993).

²² A third related literature addresses conventions. For an introduction, see H. Peyton Young, The Economics of Convention, 10 *J Econ Perspectives* 102 (1996).

²³ See, e.g., Glenn Ellison and Drew Fudenberg, Rules of Thumb for Social Learning, 101 *J Polit Econ* 612 (1993); Glenn Ellison and Drew Fudenberg, Word-of-Mouth Communication and Social Learning, 110 *QJE* 93 (1995); for work from the perspective of cultural evolution, see Kraus, *supra* note 6.

Agent-Based Computer Simulations. The study of the social patterns that arise when individuals interact using very simple decision rules date back at least as far as Thomas Schelling's classic work in this area.²⁴ In his characteristically low-tech fashion— using only a checkerboard and markers of two colors— Schelling showed how even relatively mild associational preferences could give rise to substantial segregation. Agents were given simple preferences and rules for moving around the checkerboard. Schelling demonstrated that segregation could emerge quite naturally even if none of the participants had an affirmative taste for discrimination.²⁵

More recently, there is an emerging literature that uses computers to study self-organization in social systems. This is a spillover from the complexity and artificial life literatures in biology and the physical sciences, where again the emphasis is on the computer simulation of complex adaptive systems.²⁶ Mitchell Resnick's *Turtles, Termites, and Traffic Jams*²⁷ is a wonderful introduction to the possibilities in these large, decentralized models, but its focus is epistemological rather than the detailed study of particular social phenomena. Epstein and Axtell's *Growing Artificial Societies*²⁸ represents the most sustained treatment to date in the social sciences. It describes the Sugarscape, an artificial society constructed with more than 20,000 lines of computer code. Behavior on 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. This gives rise to a generative approach— hence the use of the term in

²⁴ Thomas C. Schelling, *Micromotives and Macrobehavior* (Norton, 1978).

²⁵ For recent implementations of this work, see Joshua M. Epstein and Robert Axtell, *Growing Artificial Societies* 165-71 (Brookings Institution Press/MIT Press, 1996); Paul Krugman, *The Self-Organizing Economy* 15-22 (Blackwell Publishers 1996); and Mitchell Resnick, *Turtles, Termites, and Traffic Jams* 81-88 (MIT Press 1994).

²⁶ For representative work, see Stuart Kauffman, *At Home in the Universe* (Oxford, 1995); John H. Holland, *Hidden Order* (Addison-Wesley 1995); Per Bak, *How Nature Works* (Springer-Verlag 1996).

²⁷ Resnick, *supra* note 25.

²⁸ Epstein and Axtell, *supra* note 25.

the title of this article— to modeling.²⁹ In mission, moving from micro to macro isn't new; the work of the leading macroeconomics theorists of the 1980s was to build rigorous micro foundations for macroeconomic phenomena.³⁰ What is newer is the use of explicit computer simulations with detailed specifications of decision rules as the means to accomplish these ends.³¹

The book provides a leading example of the possibilities of agent-based computer models, but also makes apparent the weaknesses of these models to date. The book is bereft of institutional features. That is institutions play very little explicit role, and notwithstanding the desire to grow everything “from the bottom up,” none of the basic institutional framework— norms, contracts, laws and organizations— is generated. (Of course, do note that the same criticism applies to this paper: 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) to be at least an order of magnitude more complex than the problem tackled here.)^{32, 33}

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 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

²⁹ See Epstein and Axtell, *supra* note 25, at 177.

³⁰ See, e.g., the classic Robert E. Lucas, Jr., *Expectations and the Neutrality of Money* 4 J Econ Theory 103 (1972); Thomas J. Sargent, *Dynamic Macroeconomic Theory* (Harvard, 1987).

³¹ Though even here macrotheorists have worked with boundedly-rational agents— automata— to investigate macro phenomena. See Thomas J. Sargent, *Bounded Rationality in Macroeconomics* (Oxford 1993).

³² For additional discussion of Sugarscape and for a general introduction to agent-based simulations, see John L. Casti, *would-be worlds* (John Wiley & Sons 1997).

³³ And, for a recent effort to apply these principles in economics pure, see Krugman, *supra* note 25.

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 this highlights 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 don't 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.³⁴ 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.³⁵ The areas of overlap between this work and the current paper will be noted throughout the paper.

In addition to this literature, Anderlini and Ianni³⁶ use cellular automata to explore success in a pure coordination game, that is, where the players are indifferent as to which of two possible equilibria obtains. 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.³⁷ And, finally,

³⁴ See George J. Mailath, *Economics and Evolutionary Game Theory* (manuscript, November 11, 1996); Drew Fudenberg and David K. Levine, *Theory of Learning in Games* (manuscript, May 8, 1997); 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).

³⁵ John Maynard Smith, *Evolution and the Theory of Games* (Cambridge Univ. Press, 1982).

³⁶ Luca Anderlini and Antonella Ianni, *Path Dependence and Learning from Neighbors*, 13 *Game & Econ Behav* 141 (1994); Luca Anderlini and Antonella Ianni, *Learning on a Torus, The Dynamics of Norms* (Forthcoming, Cambridge).

³⁷ See also Lawrence E. Blume, *The Statistical Mechanics of Strategic Interaction*, 5 *Games & Econ Behav* 387 (1993).

the most developed related literature use a spatial version of the prisoner's dilemma model to look at questions in evolution and biology.³⁸ This literature is discussed in more detail below.

II. The Basic Setup

The two-by-two interactions considered 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, these are games with only one type of player. This is a natural starting point, but also excludes some well-known games, including the Battle of the Sexes.

I will embed this game in a spatial framework, and to do so, I will lay out an $n \times n$ grid.³⁹ In a 10×10 version, and to focus on just one block of nine players, we would have Diagram 1, set forth on the following page. Player X interacts with her immediate eight neighbors. She plays the free-standing two-by-two game with each, but she only plays one strategy per round. That is, 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. So, for example, if our player played right, while all 8 of her neighbors played left, she would receive a payoff of $8b$. If 7 played left, while 1 played right, she would get a payoff of $7b + d$. This is a natural extension of the two-by-two

³⁸ Martin A. Nowak and Robert M. May, Evolutionary games and spatial chaos. 359 Nature 826 (1992); Martin A. Nowak and Robert M. May, The Spatial Dilemmas of Evolution. 3 Int'l J Bifurcation and Chaos 35 (1993); Martin A. Nowak, Robert M. May and Karl Sigmund. The Arithmetics of Mutual Help. Scientific American 76 (June, 1995); Martin A. Nowak, Sebastian Bonhoeffer and Robert M. May, More Spatial Games, 4 Int'l J Bifurcation and Chaos 33 (1994).

³⁹ For an introduction to the issues raised by embedding games, see Baird et al, supra note 4, at 191-95.

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 101x101, giving a total of 10,201 cells (and 10,201 players).

Diagram 1

I will discuss two kinds of neighborhoods: payoff neighborhoods and information neighborhoods. A payoff neighborhood is the local area that directly impacts one player's payoffs. This impact will be through the strategy choices made by the other players in the payoff neighborhood. In a basic two-by-two game, the payoff neighborhood for one player is simply the other player. In the game set forth in the diagram above, the payoff neighborhood for player X are her 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 of how linked we are to our neighbors. Of course, link-

age 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 are the latter, they are purely instrumental neighborhoods. My neighbors's strategy decisions set the environment, which, when coupled with my decision, create the consequences that flow to all of us.

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

Diagram 2

The grouping on the left is the Von Neumann neighborhood, that on the right the Moore neighborhood.⁴⁰ Focus on the cell at the center of each neighborhood. In the Von Neumann version, the payoff of each player is determined by her choice and that 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 that of her eight neighbors. (This is the version described in Diagram 1.) Of course, each player will be treated as centered at a payoff neighborhood, so the mosaic created is one of overlapping payoff neighborhoods.

An information neighborhood is the area over which a player observes results. You 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 be identical. But, as a general matter, it would be a mistake to assume that these need to be coextensive. Infor-

⁴⁰ For a general introduction, see Richard J. Gaylord and Kazume Nishidate, *Modeling Nature 4-7* (Springer-Verlag 1996).

mation 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 almost 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, obviously).

To look at this, I will vary the information neighborhoods. We will look at the coextensive cases first, then I will move to cases in which the information neighborhood is larger than the payoff neighborhood. A Von Neumann payoff neighborhood will be embedded 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. In similar fashion, we can 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, we will look at one version of a global information neighborhood, to let 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.

This lays out the basic physical setup. 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. How strategies change from round to round will be addressed in a moment, but, as is conventional in this literature, I will use a color-coding scheme to track these changes. This color-coding scheme will be used in the color pullout associated with this article and the website where videos of simulations are available.⁴¹ 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 red if right has

⁴¹ The website is at:

<http://www.law.uchicago.edu/Picker/AworkingPapers:norms.html>.

played in both of those 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.)

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

- *The Biological Model:* Two types of players exist, one can only cooperate, one can only defect (that is, only plays left only, one plays right only). Thus, these “players” make no decisions at all; they are by nature and instinct pre-programmed. 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 is a natural interpretation of the Darwinian model of adaptive success generating reproductive success.
- *The Full-Information/Full-Rationality Model:* Players know the strategy space of the game, the function giving rise to the payoffs, and have the ability to assess optimal strategy. This 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 Limited-Information/Incomplete Rationality Model:* Players lack full information, perhaps about the strategy space or the way strategies interact to give rise to payoffs. (They might know that they are playing a two-by-two game, but don’t know which one.) They may learn these 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 it isn’t 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.

It is important to understand where these approaches converge and diverge. The best way to see this is to consider a spatial version of the prisoner's dilemma. I would be remiss in not addressing the prisoner's dilemma, given its overall prominence, though for reasons that will become clear, I think the spatial prisoner's dilemma is something of 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	1, 1	0, b
	Right	b, 0	0, 0

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 left and right, while if she expected player 2 to play left, she would clearly play right, as $b > 1$. This reasoning holds for both players, so we should see (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.

Now look at this 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, walk through the possibilities one-by-one. For example, imagine that you know that all eight of your neighbors are going to cooperate: what should you do? If you cooperate, you receive a payoff of 8, if you defect a payoff of $8b$, so you should clearly defect. Suppose that 7 neighbors were going to cooperate and 1 was going to defect. You get 7 from cooperating, $7b$ from defecting, so you defect. Approach the other extreme. Suppose that 7 of your neighbors were going to defect and 1 intended to cooperate. You will get nothing from your defecting neighbors, regardless of whether you cooperate or defect, 1 from your cooperating neighbor if you cooperate and b from him if you defect. So you defect. Finally, suppose that all 8 of your neighbors were going to defect. You would get 0 from cooper-

ating, 0 from defecting, so you are indifferent. Taking all of this 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. Here, 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: 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,⁴² a rich variety of behavior emerges. There is no good way to describe this work in general. In some cases, the model does converge to the all-defection outcome. 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 horizon, and the cycle will be all but undetectable to you or me.⁴³

This is a dramatic change from the full-information/full-rationality model. That model locks into the defection equilibrium immediately and is bereft of interesting behavior. In contrast, in the biological framework, we have an embarrassment of riches. Where does this put us? Biology is a great subject and all that, but we want models in which people make decisions and their behavior is not simply instinctive and pre-ordained. To get a handle on this, switch to the limited-information/limited-rationality framework. Players know only that they are playing a two-by-two game, so that they 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

⁴² See the articles cited in footnote 38.

⁴³ To see an example of this, set $b = 1.65$, and start with an initial random distribution of 10% defectors and 90% cooperators. A movie showing the first 100 generations of this model is available at my website, *supra* note 41.

simple observation. In each round, each player observes the payoffs that she receives and those received by her neighbors and the strategy that they play. Given that, she adopts the strategy that did best, meaning, the strategy that resulted in the single highest payoff in the nine cells that she sees.

This matches the biological framework exactly; indeed, that is the point. The absence of learning means that our player never assesses whether defection is a dominant strategy. Instead, she uses a more basic decision rule and one 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 moves too quickly. It doesn't 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 plays cooperate 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:

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

(p_{cc} and p_{cd} are the payoffs in the model when, respectively, both cooperate and player 1 cooperates and player 2 defects). Obviously, an infinite number of pairs will satisfy this, including (1,0) and (0,3). She doesn't know much, yet. Suppose that in the next round, she again cooperates, while 5 of her neighbors cooperate and 3 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 in two 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} . She then has full-information about the game, and should be able to assess that defection is her dominant strategy.

So, the limited-information model, which tracks the interesting biological model when there is no learning, quickly converts into the full-information model. To give a richer sense of this, I have simulated the

model 100 times on the assumption that 25% of the players learn the game and defect. In addition, I have assumed that 25% of the remaining players defect initially, while 75% cooperate. After that, they choose the strategy that yields the highest payoff in the nine cells that they see. Under those parameter settings, 85 out of 100 games ended in the all-defection equilibrium. In the other 15 simulations, cooperation did poorly.

For me, all of this means the following. We can generate complex, interesting behavior in the spatial prisoner's dilemma, and this 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 is a result worth noting, but it also means that we may learn very little by playing the spatial prisoner's dilemma.

This 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 rationality approach into the traditional rationality approach. All of this suggests that we should look elsewhere to model norm competition.

III. How Norm Competition Evolves

Turn 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. What does game theory say about this? Very little, it turns out. Both (left, left) and (right, right) are Nash equilibria: neither player wants to switch strategy given the other player's strategy. Nonetheless, we have no good way of choosing between these equilibria. We do not reach a determinate result, meaning that I do not explain how the players would actually make decisions.⁴⁴ This means that we are necessarily pushed to use our limited-information/limited-rationality framework. Once we are there, we may be able to weaken our usual information assumptions without changing results in any material way.

To see this, 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 of 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, 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, because we do not have a determinate way of playing co-

⁴⁴ This overstates somewhat. Harsanyi and Selten have emphasized the idea of risk dominance to resolve these games. The 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 Press, 1988).

ordination games, we necessarily must reach outside the model for an operational decision rule. Learning doesn't convert the limited information framework into the full-rationality model, and indeed, we remain quite close to the biological approach.

This lays out the game and its setup. Next, we need to specify a choice rule for the players. For now, we will assume that each player uses the same rule, though asking what happens if a mix of rules is used is obviously important. 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 so far. 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. So the player looks at the payoffs obtained by herself and her eight neighbors, figures out which is highest and adopts the strategy played by that player. As noted before, this scheme doesn't work in the first round—there are no prior payoffs to evaluate obviously—so strategy choices will be assigned at random. Do note that this decision rule—while perfectly plausible to me at least—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, these are not the same idea: we will have absorbing states in the models which are clearly not Nash equilibria.) To see that these two outcomes are Nash, if a given player thought that every other player was going to play right, she would clearly play right. She would get zero if she played left and a payoff of 8b from playing right. The same holds for left: if everyone else was expected to play left, she would play left and get 8 rather than right and 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, that is, 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, again, all observe only left outcomes, and choose 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, 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 xx to zz on the color foldout show six snapshots of the evolution of this model.⁴⁵

As is evident, this model converges to a mixed equilibrium, with large numbers of players adopting each strategy. (As should be clear, I mean mixed in the sense of having both strategies played at the same time by different players; this is not mixing in the sense of one player playing both strategies with some positive probability.) Of course, this is inefficient, as the social optimum is achieved when all players play right. Nonetheless, this isn't too surprising. The value of getting to the right equilibrium is low— 1 vs. 1.05— and the initial starting conditions do not tilt the tables 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 xx to zz on the color foldout show six snapshots of the evolution of this model.⁴⁶ 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 isn't enough to assure convergence to the right equilibrium. Let 80% of the players start with the left strategy and 20% with the right, and consider the six snapshots of the model given on the color foldout as Figures xx to zz. Once again, the model fails to converge completely.⁴⁷

This is a good point to highlight the difference between absorbing states— fixed points— and Nash equilibria. The result in Figure zz is an absorbing state, but it is not a Nash equilibrium. The diagram below sets out the relevant chunk of the final result:

B B B B B B B B
 B B B B B B B B
 B B B B B B B B

⁴⁵ To see this directly, play the video at my website, *supra* note 41.

⁴⁶ Again, To see this directly, play the video at my website, *supra* note 41.

⁴⁷ To see this directly, play the video at my website, *supra* note 41.

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

Diagram 3

To see that this isn't Nash, focus on the red cell at the corner of the cluster of nine reds. Round after round in this model, this cell continues to play right— remains red— because it observes the red at the center of the cluster of nine reds receiving a payoff of $8b$. This, of course, is the best possible outcome in the model, and, on my assumed strategy for making decisions, any player who observes that outcomes adopts right as her strategy. Nonetheless, for this to be Nash it would have to be the case that right is the best strategy for this player given all of the other strategies. Our corner red expects, round after round, 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? This particular mixed play absorbing state can be sustained only if we have “irrational” play round after round. Our corner red continues to play red, because of how well the center red does, even though the context in which the center red plays is quite different from her 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 play in perpetuity that seems implausible. Some middle ground may be more appropriate,⁴⁸ though the fixed environment makes the best case for an eventual move to a Nash approach. Introducing a background level of change to these models— birth, death and spontaneous mutations or experiments— might be enough to make it difficult to bring the Nash scheme to bear (and almost might alter the mixed play results for ab-

⁴⁸ See the suggested directions for future research at text at p.65.

sorbing states). In any event, in this paper, the approach taken 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 obtained.⁴⁹

B. The Emergence of a Phase Transition

These examples give a flavor for the range of behavior that arises in the model. To get more systematic, I set $b = 1.65$ — about which more momentarily— 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). Three possible results are captured in the three graphs of Figure 1. All of the players could converge on playing right (“Red”); all could converge on left (“Blue”); some could converge on left while others played right (“Mixed”).

Figure 1

⁴⁹ This parallels— indeed, overlaps with— the recent work in evolutionary game theory. See note 34 supra.

The graphs chart the number of times each possible outcome occurs in the 100 simulations for each initial value. So, if we start with 1% of the players playing left and 99% 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% of the players playing left and 1% playing right, then in 100% of the cases we converged on the inferior left-left equilibrium.

Neither of these results is particularly surprising. What is more interesting is to note how robust the good equilibrium is. Even if we start in tough conditions—say with 80% of the players playing left and 20% playing right—we still converge on the good equilibrium in 100% of the cases. As we push the initial density of players playing the inferior choice ever higher though, 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% playing left initially, the graphs cross: as many simulations converge on the bad equilibrium as converge to the good equilibrium. Push the initial densities still higher—have more players choose initially the inferior strategy—and 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 to the inferior equilibrium.

The shape of these graphs is characteristic of a phase transition in physics or a model of punctuated equilibria in biology. The system has two natural equilibria and shifts from one to the other occur over a very 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% of the players initially choose the inferior play? To understand this, 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 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

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 8 players playing the strategy that she has played, so she is perfectly coordinated with her neighbors). The best blue that that our border red could see would have a payoff of 7. No border red will switch. 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 blues bordering the cluster of 9 reds. There are 4 corner blues. A corner blue will see a best blue receiving a payoff of 8. The only red it sees receives $3b$. Given our numbers, a corner blue will switch only if $b > 8/3$. There are eight off-center border blues. Each will see a blue receiving 8, while the best red seen— the middle border red— will be getting $5b$. The off-center border blue will switch if $b > 1.6$. Finally, there are 4 center border blues, but the analysis for them tracks that for the other border blues. So, if $b > 2.66$, all of the border blues switch strategies; if $1.6 < b < 2.66$, all of the non-corner blues switch, while if $b < 1.6$, no blue cells change over.

Already, under this decision rule, we see that we can have pockets of players playing blue and red simultaneously, if the benefits of coordinating on strategy right are not sufficiently great. This 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 role 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 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 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

Diagram 5

For the reds, nothing has changed. The original 9-cluster of reds play as before. For the yellow cells— recall that these are players who had been playing strategy left and who have now switched to strategy right— each yellow will see a red cell receiving a payoff of $8b$. Each yellow cell will play strategy right again (and will turn red in the next round).

As to the blue cells, again, go blue type by blue type. The former corner blues still see a blue playing with 8 neighboring blues, but now the best red seen plays with 7 neighboring reds. The old corner blue will switch if $8 < 7b$, or if $b > 1.14$. The new non-corner border blues will each see a red playing with five reds, and thus will switch, given that $b > 1.6$. 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 2.⁵⁰

As is clear, with $b > 1.6$, there is a powerful drive in the system to converge on the correct equilibrium. That is not to say that the model necessarily does get there. I started the cluster analysis with a block of 9 red cells, and depending on how relatively scarce players of strategy right are, we may not have such a cluster to start with. 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 red cell cater-corner to it. That blue cell will touch 7 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.⁵¹

⁵⁰ To see this directly, play the video with a nine red cell starting point at my website, *supra* note 41.

⁵¹ To see this directly, play the video with a four red cell starting point at my website, *supra* note 41.

When $b = 1.65$, a 9-cluster grows until it spreads through the entire domain, while a 4-cluster withers. A little bit more analysis makes clear that a 2 by 3 cluster of 6 suffices when $b = 1.65$. Again, surround our 6-cluster 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. (That is just to say that there are two red cells with 5 red neighbors.) 8.25 exceeds the highest payoff that a blue cell could obtain—namely, 8—and thus no red cell will switch in the next round. What will the neighboring blue cells do? It is enough to notice that each of these neighboring blue cells sees a red receiving the 8.25 payoff. Again, they will see no blue doing better, so they will all switch. That means, we now have a red 9-cluster, and the analysis above applies.⁵² And, to complete the analysis, consider a cluster of 5-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, and the fifth red cell will follow in the next round.⁵³ So six sustains growth while five does 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 six cluster grows indefinitely until every player plays the socially-preferred strategy. If we assign strategies at random initially, what is the probability that we will have one or more of the growth clusters? We could do this algebraically, but the chart itself provides the answer. For example, when we start with 90% of the players playing left and 10% playing right, then 55 times out of 100, we do not get at least one of our growth clusters. This is just to restate the result that, absent a growth cluster, the small, scattered clusters of red die, and we converge to the inefficient blue equilibrium.⁵⁴

To put this in a different language— that of dynamic systems— we have mapped the basin of attraction for each of our point attractors.

⁵² To see this directly, play the video with a six red cell starting point at my website, *supra* note 41.

⁵³ To see this directly, play the video with a five red cell starting point at my website, *supra* note 41. Note that there is not a unique way of laying out the five cells. This was true as well for our prior examples as well, but I emphasized relative compactness of the cluster, and that gave a simple representation of the cluster.

⁵⁴ To see an example of this, play the video at my website, *supra* note 41.

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.⁵⁵ (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% left to 85% left, while the corresponding basin of attraction for the blue equilibrium covers 96% left to 99% left. Between 86% left to 95% left it is 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.

Step back and now ask what we should make of this. First, this is path dependence writ large. Initial starting conditions matter in an important way for where this system converges. There is a burgeoning literature on path dependence⁵⁶ and this model gives a nice, crisp example of this phenomenon. Second, and perhaps of more interest given what we already know about path dependence, this setup converges quite nicely to 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 that equilibrium.⁵⁷ 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. Each if the choice is substantially biased against the good strategy, we still converge on the best

⁵⁵ For an introduction to these ideas, see J.M.T. Thompson and H.B. Stewart, *Non-linear Dynamics and Chaos* 9-12 (Wiley, 1986).

⁵⁶ See W. Brian Arthur, *Competing Technologies, Increasing Returns, and Lock-In By Historical Events*, 99 *Econ J* 116 (1989).

⁵⁷ 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, *supra* note 34, at 44; see also Ellison, *supra* note 34.

norm. (Replace the coin with a single die, and play right only if six comes up, and 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 one favoring the good strategy.

This is all 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 which in turn often are used to justify legal intervention. At least in this particular framework and on these values, my guess is that it is highly unlikely that we would end up in the inferior equilibrium. We need either really 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 unrelenting sameness until we get to particularly skewed distributions of initial choices at the end. Value after value ends up in the same place. But this omits— importantly— 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 it took for the model 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.⁵⁸

⁵⁸ 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, *supra* note 34, at 1060-62.

Figure 2

C. Two Phase Transitions and Mixed Play

So far we have looked at nearly 10,000 simulations of a 10,000 player model for a single value of $b = 1.65$. The next step is to understand how these results change as we alter 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 support smaller initial growth clusters, smaller b values will require larger initial clusters. So, for example, a cluster of 4 red cells will 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 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. This is not to say that we cannot end up in the inefficient norm equilibrium, but now that would occur only if no cluster of 4-red cells formed in our initial distribution of strategies. As we increase b beyond 1.65, we keep pushing towards 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 9-red cluster. As noted above, when $b < 1.6$, the neighboring blue cells do not switch over. The blue most likely to switch over sees at least one blue earning 8, while the best red seen will earn $5b$. When $b < 1.6$, this

is less than 8, and the blue cell stands pat. This raises the possibility of a mixed outcome, that is, one in which, in perpetuity, some players play left, while others players play right. Indeed, were we to start with a single 9-red cluster and all others blue, that model wouldn't 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, and we see, as predicted, a region in which the model converges to an outcome in which some players are playing left, while others play right.

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 don't 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, where 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, say, 87 and is over by the time we get to 92. We move away from the mixed absorbing states to the bad all blue-equilibrium. Finally, for initial blue densities of 92 to 99, we do converge to the blue equilibrium.⁵⁹

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

Figures 4 to 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 this means that the all-right equilibrium is still always better than the all-left equilibrium, but when $b = 1.05$, the differences are quite small. A quick glance at these figures reveals a number of facts. First, in each, 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% of the simulations converges to the good norm is reduced from 65% when $b = 1.35$ to 60% when $b = 1.15$; to 17% when $b = 1.14$; to 16% when $b = 1.10$; and, finally, to 17% 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

⁵⁹ I should note that at an initial density of 76, 4 of the 100 simulations did not converge before reaching the then-applicable limit of 250 generations.

then dramatically while that for the bad equilibrium grows slowly over time.

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

As this should make clear, we can remain fairly confident that the model will converge on the good norm so long as b is at least 1.15. To be sure, the chance that we will end up elsewhere, in one of the mixed play outcomes or the all-blue outcome, is rising, but even with $b = 1.15$, so long as not more than 60% of the players play left initially, we will converge on the good norm. This again should be quite reassuring. If two norms are competing and one is better for everyone than the other by at least 15%, 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. That is not a result that we could get out of our free-standing coordination game, where we could only identify our two Nash equilibria and then punt. The result that we get to the good equilibrium when there is a substantial advantage to doing so instinctively seems right and it is nice to see this result emerge in the model.

But, and this is another point of interest, there is a sharp break between 1.15 and 1.14. The probability of ending up in the good norm equilibrium in all cases plummets. 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 results holds as we move b towards 1. There are two natural questions. First, why the sharp change at this point? Second, are all the mixed outcomes identical, or, put differently, are there meaningful differences for different initial

densities, even if we know that we end up in mixed play in 100% 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 100 simulations for each initial density. Obviously, if all 100 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 is substantial and unsurprising path dependence.

Now consider the first question, namely, why the sharp transition between 1.15 and 1.14. The simple answer is that 7×1.14 is 7.98, which is less than 8, while 7×1.15 is 8.05, which is more than 8. To see why that matters, consider a 9-cluster of blue cells surrounded by a sea of red (this is the flip of our prior example, where we started with a 9-cluster of red):

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 are all rock-solid here, so long as $b > 1$. Each red either will see a red receiving $8b$ or be such a red, and the best blue that could be observed would do no better than 8. No red will switch. Next consider the blues, 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 a best red— the one cater-corner to it— receiving a payoff of $7b$. That blue will stay blue is $8 > 7b$, or $b < 8/7$. The corner blue stays at $b = 1.14$ and switches at $b = 1.15$. That is, a 9-blue cluster cannot be invaded if $b < 8/7$, and this will give rise to a mixed play outcome. 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 all this? In one way, this should be quite comforting. If we have “reasonable” initial blue densities and sufficient benefit from the superior equilibrium, we converge to the good equilibrium. In many cases, there may be good reason to expect this to hold. 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— we will be at one end of the initial densities or the other. The low values in the parameter space would 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

It is essential to alter basic features of the model to evaluate how important they are to driving the results. In this section, I will discuss four variations on the original model: (1) a different decision rule that focuses on the strategy seen resulting in the highest average payoff; (2) a different payoff neighborhood; (3) different information assumptions; and (4) an alternative approach to parallel decision processing. Take these one by one.

Different Decision Rule. I have used a particular decision rule so far without offering any particular justification for the choice. I believe it a plausible rule, but surely not one that is uniquely so. It might be worth exploring this systematically, but, to provide just one source of comparison, consider the following rule. Suppose that instead of choosing the strategy from the nine results seen that results in the highest payoff our players choose the strategy that results in the highest *average* payoff. This imposes a much more severe calculation burden, but not one that we should think of as too daunting.⁶⁰ How do the results change with the new decision rule?

Figures 10 to 14 set out results for b set to, respectively, 1.65, 1.35, 1.15, 1.10 and 1.05. A quick look makes clear 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 is that we do see 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. That other change of interest is that we sustain 100% convergence to the good norm for lower values of b . For example, when $b = 1.10$, we get 100% convergence even with initial blue densities of 54%. Recall that under the highest decision rule, when b was 1.10, we lost 100% convergence when the initial blue density exceeded 16%. This is a

⁶⁰ To see an example of this, return to diagram 4 and focus on the decisionmaking of the 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, respectively, 5, 6, 7, 6 and 5, for an average payoff of 5.8. In contrast, the observed red cells have payoffs of, respectively, $5b$, $8b$ and $5b$, plus the corner red cell itself obtains a payoff of $3b$. The average red payoff is therefore $21/4 * b$, or $5.25b$. The corner red will switch if $5.8 > 5.25b$, or $b < 1.105$, and will stay otherwise.

substantial difference, not only in absolute terms, but relative to where we think we are likely to be. If you believe that at least 50% of the players should adopt the good strategy, the highest average rule provides real evidence that we will converge to the good norm even when the good strategy provides relatively insignificant benefits (10%). We do lose this result when b drops to 1.05– 100% convergence to the good norm occurs last at an initial density of 43%– but this 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.

Figure 10

Figure 11

Figure 12

Figure 13 **Figure 14**

Different Payoff Neighborhood. A second way to change the model is to switch payoff neighborhoods. We have focused on the Moore neigh-

borhood, and there is no doubt that particular features of the results are driven by that fact. 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 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

Figure 16 **Figure 17**

Figure 18

We do see important differences in the results. The mixed play region exists even when $b = 1.65$, and the last 100% convergence to the good norm drops from 85% in the original Moore version to 78% in the Von Neumann version. The results at $b = 1.35$ are almost identical for the two versions. Note also that we get a break point at 1.33 rather than 1.14, so when $b = 1.30$, we have a large mixed play region emerge. Again, this is important, for it suggests that even with gains as large as 30%— with $b = 1.30$ the good norm is that much better— we cannot be confident that the model will converge to the good equilibrium.⁶¹ Compared with the results using the Moore neighborhood, this suggests that a bigger payoff neighborhood increases the chance that the model when converge on the good norm.⁶²

Different Information Assumptions. To see a simple example of the way that a small change can alter the results substantially, alter the information neighborhood. Recall that the payoff neighborhood is made up

⁶¹ As to why $b = 1.33$ matters, see *infra* note 65.

⁶² This result needs more analysis. Berninghaus and Schwalbe reach the opposite conclusion in a related model— “[t]he smaller this reference group, the higher the probability that an efficient equilibrium will be reached.” See Berninghaus and Schwalbe, *supra* note 34, at 79.

of the neighboring cells whose actions directly influence payoffs. The information neighborhood is the neighborhood seen by a given cell and provides information that can be used 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.

Switch this. Set the payoff neighborhood as the Moore neighborhood and let the information neighborhood be the 5-by-5 cluster of 25 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 extend out 2 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 25 cells seen. Figures 19 through 21 set our runs with b set at, respectively, 1.65, 1.35 and 1.10:

Figure 20

Figure 21

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

Why the dramatic shift? Return to diagram [xx], where we have a 9-cluster of 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 held firm. If 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 9-cluster of red is as before: each sees the center red receiving $8b$, and no one will switch. But, and here is what matters, the border blue cells now see the red cell in the center of the 9-cluster. When the blue cells could see out only one level, they saw only their immediate neighbors. Seeing out two levels brings the center red within view; it does the best that anyone could do; and the blue cells shift. Note that this is true regardless of how big b is, so long as $b > 1$.⁶³

This is not to say that we always reach the good equilibrium or that the size of the gain from getting to that equilibrium is irrelevant. The graphs make clear that if we start with too many players playing the inferior strategy, we will converge to the inferior equilibrium. And, the size of the minimum cluster necessary to sustain growth is still determined by b . But, so long as we have at least one cluster of 9-red cells, 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 .

⁶³ To see this, play the video that starts with a single 9-red cluster and $b = 1.01$ at my website, *supra* note 41.

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

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. What is noteworthy is that, save for one density, we converge to the good equilibrium more frequently when we have less information. To say it again, when we give the players more information, we end up at the wrong equilibrium more often.

Somewhat shockingly, more information equals worse results. The double Moore results move to the inefficient blue equilibrium more often, sometimes substantially so (look, for example, when initial blue density is 88). This result is in spirit akin to that seen in the herd behavior literature,⁶⁴ but is less obviously tied to hidden information.

⁶⁴ See Abhijit Banerjee, A Simple Model of Herd Behavior, 107 QJ Econ 797 (1992); see Baird et al, *supra* note 4, at 213-17.

This is not to say that our players are worse off with the additional information, as the double Moore model converges much more quickly (see Figure 23). This is a good thing, obviously, and we need to explicitly weigh faster outcomes against the possibility of more wrong outcomes.

Figure 23

As a second example of how information changes alter the results, consider Figures 24 to 26. These runs use a Von Neumann payoff neighborhood and a Moore information neighborhood meaning, of course, that payoffs are 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% red convergence at an initial blue density of 78%. The revised model (see Figure 24) has no mixed play region at all and converges successfully to the good equilibrium 100% of the time up to an initial left density of 89%. The same result holds at $b = 1.35$. Before, see Figure 16, we had mixed play and 100% red convergence through 67%.

Figure 24

Figure 25

Figure 26

In the revised model, see Figure 25, no mixed play occurs and 100% red convergence runs through 88%. 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%. This change in the basin of attraction for the good equilibrium is a sizable improvement over the Von Neumann run (see Figure 18), where all-red convergence was lost at 8%.⁶⁵

⁶⁵ If you have followed this so far and care why the results change with the switch to the Moore neighborhood, the key is to note that when $b > 1.33$, a red Von Neumann cluster will be a growth cluster when the Moore information neighborhood is used. The key is that when we start with five red cells in the form of a cross— a red cross literally— 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 now observes the center red as well. Note importantly that they would not have seen this red with Von Neumann vision— if they were using the Von Neumann information neighborhood— as the center red is on their diagonal. With the Moore information neighborhood, they too see the center red, and they switch over. This creates a cluster of 9 red cells, surrounded by 16 blue cells. The non-corner blue cells— all but 4 obviously— will each see one red cell receiving a payoff of $3b$, the center-edge red. The best blue cell they will observe will receive a payoff of 4, and the non-corner border blue cells will switch if $3b > 4$, or $b > 1.33$. This drives the model round after round to the red equilibrium. To see the directly, play the video at my website, su-

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

Figure 27

Figure 28

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 up is more complicated than this, 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 with Figure 4:

Figure 29

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

These are 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 with $b = 1.35$ that we will converge on the good equilibrium. That comes at the expense of losing the mixed play region.⁶⁶

⁶⁶ 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 24-days of computer time. And, there is substantial reason to doubt in any event the overall stability of the mixed play region. A base level of mutations in the players would tend to disrupt the boundaries that define the mixed play outcomes. Compare Kandori, Mailath and Rob, *supra* note 34.

IV. Speculations on Law: Decentralized Rule-Making, 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 more generally nor have I addressed the range of infirmities associated with government action. That caveat issued, I will nonetheless offer a few speculative thoughts:

Echoing the Free-Standing Coordination Game. We have used the basic coordination game as our generic model of the problem of coalescing around a Pareto-superior norm. I had very little to say about the original free-standing coordination game. For better or worse, there has been quite a bit to say about the spatial version of this game. As the extended analysis of the basic spatial game and possible variations makes clear, 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 improved matters, 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. For me, at least, this makes me much less concerned about the problem seen in the free-standing coordination game.⁶⁷

The Importance of Phase Transitions for Policymakers. I find it striking how small a range matters for the outcome 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 this is a good thing. If a policymaker were to try to switch this system from one equilibrium to the other, the narrow band effect would work in our fa-

⁶⁷ Again, see the prior results to this effect in Kandori, Mailath and Rob, *supra* note 34; Ellison, *supra* note 34.

vor. A policymaker who mistakenly sought to switch from the red equilibrium to the blue and who took blue density as the variable to work on might do real damage, if we were near the phase transition. That, however, would be exceedingly bad luck— unless we have a story as to why these systems should gravitate towards the point of a phase transition.⁶⁸ 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. (That said, my guess is that politicians at either end of the spectrum would say that their worthy opponents routinely display such bullheadedness in promoting policies!)

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— correctly— might want to try to trigger a norm cascade, which looks a lot like swooshing down the phase transition to the good equilibrium. Here, very little effort would be rewarded quickly, and the reward might be vastly disproportionate to the effort expended.

When might this 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. (This may be where we are today in recorded video technology, with VCR tapes representing the old, locked-in standard, and DVD the new superior technology.)

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 convention, or custom, to evolve. Indeed, the status attached to the pre-existing convention further entrenches that convention against newcomers.

⁶⁸ We do get stories in the chaos literature of models that move to the edge of chaotic behavior, see, e.g., Kauffman, *supra* note 26, at 26-29, but that is a far cry from where we are now.

The models proper won't let us get from the old standard to the new standard, and even if we allow in an ad hoc fashion small numbers of folks start to 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% or 99% playing the old, inferior standard, and only 1% or 2% playing the new standard. As the charts demonstrate, even in the face of a substantial improvement— $b = 1.65$, after all, represents 65% 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.

Seeding Norm Clusters. If we take the model literally, there is a more direct route open to the government: seed norm or standard clusters. Given a cluster of the right size— for example start with 6 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 foldout at Figures xx to zz.⁶⁹

Gerry Mackie provides a striking example of the power of seeding norm clusters in an account of the end of footbinding in China.⁷⁰ 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. Notwithstanding this, the practice vanished in a generation. Mackie cites data showing, for example, that in Tingsien, 99% of the women were footbound in 1889, 94% 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 antfootbinding society in 1874. Families pledged that they would not footbind their daughters and that they would not let

⁶⁹ To see this directly, play the video at my website, *supra* note 41.

⁷⁰ Gerry Mackie, *Ending Footbinding and Infibulation: A Convention Account*, 61 *Amer Sociological Rev* 999 (1996).

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 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, government to be sure, but so to charities and for-profit entities.⁷¹

This approach is probably more important in cases in which the appropriate norm or meaning that is in place changes and we need to navigate from the formerly-appropriate norm to the new norm. The idea of norm seeding is a low-risk strategy. If the government seeds an inefficient cluster, it will die, and little will be lost. If, though, the new norm is superior to the old norm, the artificially-created norm cluster will thrive and spread. This 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.

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 outcomes of the system. First, it appears to be the case that larger payoff neighborhoods do a better job of converging on norms that represent modest improvements (say, $1.15 < b < 1.33$).⁷² As discussed before, creating instrumental connectedness among individuals within a neighborhood may be difficult, but this model suggests that there may be substantial returns from succeeding. Second, more refined decisionmaking— averages rather than a single, highest value— also supported success for lower values of b . Third, additional information usually increase the likelihood of success for the

⁷¹ 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 (1988), cited in Posner, *supra* note 7, at 1740 n. 108, and in the Old South, see Jack Kenny Williams, *Dueling in the Old South: Vignettes of Social History* 64-64 (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 n27 (1984).

⁷² But see Berninghaus and Schwalbe, *supra* note 34, at 79.

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.

V. Conclusion. Limits and Future Directions

The recent interest in norms and law almost certainly dates from Ellickson's work on Shasta County. This work started as an inquiry into the Coasean Irrelevance Proposition and emerged as 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 preexisting transaction costs. Ellickson's investigation of cattle in Shasta County led him to conclude that law was irrelevant when robust local norms could evolve. Neighbors didn't know the law, didn't 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 didn't 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. 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 interconnectedness (my payoff neighborhoods); the information available to the players (the information neighborhoods); and the manner in which in-

dividuals 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.

The current paper suffers from at least three weaknesses, all of which suggest directions for future work. First, there is no conflict among the players over 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. Migration will become an important feature of these models as well, as players may reduce the conflict over norms through separation by type. Second, only competition between norms has been restricted to only two norms at a time. As suggested before, this may capture many social contexts accurately, but it almost certainly doesn't track all commercial dealings.⁷³ Third, 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.

⁷³ 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 Nat'l Acad Sci* 7716 (1993).