 aptara The Content Transformation Company	AJPS	ajps'387	Dispatch: April 10, 2009	CE: n/a
	Journal	MSP No.	No. of pages: 16	PE: Martha

# Learning to Cooperate: Learning Networks and The Problem of Altruism

**John T. Scholz** Florida State University  
**Cheng-Lung Wang** National University of Singapore

*We explore how two populations learn to cooperate with each other in the absence of institutional support. Individuals play iterated prisoner's dilemmas with the other population, but learn about successful strategies from their own population. Our agent-based evolutionary models reconfirm that cooperation can emerge rapidly as long as payoffs provide a selective advantage for nice, retaliatory strategies like tit-for-tat, although attainable levels of cooperation are limited by the persistence of nonretaliatory altruists. Learning processes that adopt the current best response strategy do well only when initial conditions are very favorable to cooperation, while more adaptive learning processes can achieve high levels of cooperation under a wider range of initial conditions. When combined with adaptive learning, populations having larger, better connected learning relationships outperform populations with smaller, less connected ones. Clustered relationships can also enhance cooperation, particularly in these smaller, less connected populations.*

Can representatives in opposing parties learn to cooperate with each other in the legislature after decades of intensive partisan conflict? Can interests and agencies in overlapping but autonomous policy arenas establish collaborative relationships to deal with crises in the regional ecosystem (Feiock and Scholz 2009; Scholz and Stiffler 2005)? Can citizens develop mutually beneficial exchanges with immigrant workers from unfamiliar ethnic groups (Putnam 2007)? Can cooperation emerge across ethnic cleavages (Fearon and Laitin 1996)? Particularly in this era of globalization and crisis, rapidly expanding contacts between previously isolated groups, markets, policy arenas, and institutions have dramatically increased the potential for new wealth-enhancing but risky exchanges. But even when potential benefits are available, can cooperative exchange relationships emerge between potentially hostile populations in settings that offer no institutional support?

This article explores the conditions under which newly connected populations, groups, or authorities can learn to cooperate across such divides. Our model reflects, but is not limited to, our initial concern with the role of policy networks in enhancing cooperative interac-

tions between regulators and the regulated. In Scholz and Wang (2006) we argued that policy networks play a critical role in increasing collaboration between regulators and regulated entities in newly developing watershed policy arenas. But available data did not allow us to address the critical question of which network characteristics and learning processes within each group foster cooperation across groups. Do dense, overlapping, strong-tie relationships promote cooperation, as argued in the social capital literature (Coleman 1988; Putnam 1993)? Are the opposite characteristics of extensive, nonredundant weak tie relationships associated with finding jobs and efficient information transmission more effective in learning to cooperate (Granovetter 1973), or are bridging relationships that bring together otherwise isolated groups (Burt 2004; Schneider et al. 2003; Watts and Strogatz 1998) the most critical of all? Does adaptive learning that explores multiple strategies enhance cooperation more than learning that rapidly exploits successful strategies (Lazer and Friedman 2007)?

To test these propositions, we extend one-population agent-based studies of the evolution of cooperation (e.g., Axelrod 1997) by developing a model in which agents in

John T. Scholz is Francis Eppes Professor of Political Science, Florida State University, Tallahassee, FL 32306-2230 (jscholz@fsu.edu). Cheng-Lung Wang is Assistant Professor of Political Science, National University of Singapore, Singapore 117570 (polwcl@nus.edu.sg).

An earlier version was presented at The American Political Science Association Annual Convention, September, 2004, in Chicago. We thank discussants from that panel and Drew Fudenberg, Matt Golder, David Segal, Justin Esarey, and T.K. Ahn for helpful comments. Research was partially funded by NSF Grant SES-0125426.

*American Journal of Political Science*, Vol. 53, No. 3, July 2009, Pp. 571–586

©2009, Midwest Political Science Association

ISSN 0092-5853

one population play iterated prisoner's dilemma (IPD) games with agents in the opposite population, but learn about successful strategies from contacts within their own population. A two-population model provides the theoretical advantage of isolating the learning and game structures, allowing us to compare the performance of alternative learning network structures on the evolution of cooperation for fixed exchange relationships.

The model features the role of nice, retaliatory strategies (e.g., tit-for-tat in Axelrod 1984; Bendor and Swistak 1997) in learning to cooperate and incorporates the rapidly expanding Markov chain analyses of the evolution of cooperation in finite and spatially structured populations being made outside political science (e.g., Nowak 2006). In particular, we explore the limitations on achievable levels of cooperation imposed by more altruistic strategies, as suggested by Bowles (2004) and Imhof, Fudenberg, and Nowak (2005). Our results suggest that the dynamic relationship between retaliatory and altruistic strategies best accounts for the cooperative advantages observed in the simulation for a combination of exploration-oriented learning processes and exploitation-oriented learning networks.

## The Learning Context: Basic Assumptions

The simulation model explores the situation in which two populations come in contact that are each well-integrated internally but have been isolated from each other until recently, although our results turn out to be equally relevant for the problem of sustaining cooperation whenever learning only takes place within populations. The populations can represent a new ethnic group moving into an established neighborhood or village (Putnam 2007), stakeholders in two independent, well-established policy arenas whose decisions increasingly impinge on each other (Schneider et al. 2003) or representatives of two parties, of independent Congressional committees or of specialized local, state, and federal authorities forced by a recent crisis or by economic development and increased globalization to deal with the same issues (Scholz and Stiffler 2005).

In each situation, new opportunities for exchange between members of opposite populations hold out potential benefits involving some risk if the new partner does not cooperate. Exchanges could include straightforward exchanges of goods and services, but also of critical policy information, social or political favors, or joint policy projects or political agreements. Exchange is modeled as

iterated or finitely repeated two-person prisoner's dilemma (IPD); we focus on the initial phase of cooperation in which two-person interactions develop a foundation on which n-person coalition and institution-building activities can later be build (Lubell et al. 2002; Sabatier and Jenkins-Smith 1993).

Exchange opportunities involve random encounters of relatively short duration with an unknown population and are relatively unimportant compared with ongoing exchanges within each population. Individuals therefore economize on cognitive effort by utilizing trial-and-error imitation and unsophisticated innovation processes to learn simple strategies that perform well in the new exchanges (e.g., Bala and Goyal 1998; Bikhchandani, Hirshleifer, and Welch 1998; Fudenberg and Levine 1998).

The model assumes that learning takes place through established relationships within each population—through friends, relatives, neighbors among ethnic groups, or through political allies and authority relationships in political or policy arenas. Individuals can tell who among their contacts does best in exchanges with the opposing population and can generally learn what strategy that person is using. Network analysis provides a means of representing different patterns of relationships as *learning networks* that provide the primary factor manipulated in the simulation study.

Two different learning processes are analyzed; an imitation process that maximally exploits the existing repertoire of strategies and a more innovative process that explores a broader range of strategies. The *myopic best response* process always chooses the best performing strategy known to the player through learning network contacts. Imitation is particularly powerful in our study setting since all players confront the same set of players in the other population, so the best-scoring strategy represents the best response for any player. This process accelerates the spread of the highest-scoring strategy, but limits choice to the range of existing strategies and does not anticipate responses in the opposing population that may change the best response strategy in the future.<sup>1</sup>

The *adaptive learning* process provides for more innovation at the cost of slower dissemination of successful strategies. This process is based on proportional fitness selection models developed originally in evolutionary biology (e.g., Holland 1992) but used extensively to represent learning processes in evolutionary game theory as well

<sup>1</sup>Fudenberg and Levine (1998) note that the myopic learning process is not optimal because it fails to anticipate responses in the opposing population, but argue that the model still provides a reasonable representation for individual decision making in groups to the extent that any single individual's impact on the response of the other population would be miniscule.

(Fudenberg and Levine 1998). Individuals still seek the best response among learning network contacts, but their willingness to change their current strategy is now proportional to the difference in payoffs. Furthermore, individuals consider each component of the best response strategy and can adopt specific components rather than just copying the entire strategy. Finally, individuals have the capacity to adopt new strategic components not present in their own or the best response strategies, which corresponds to mutations in proportionate fitness models. In short, individuals can develop new strategies primarily by borrowing components from the most successful strategies, but also by occasionally developing a new component.<sup>2</sup>

A genetic algorithm is commonly used to model the adaptive learning process in studies of the evolution of cooperation (Gott, Polhill, and Law 2003). The genetic algorithm is particularly valued as a problem-solving search strategy because of its ability to find global maxima in large problem spaces (Mitchell 1996). The continuing ability to innovate reduces the risk of premature selection of an inferior local maximum, although innovation also imposes costs in terms of prolonging the search. Thus the tradeoff between these two learning processes reflects the observed conflict in complex systems between characteristics that enhance the exploration of new ideas and those that maximize the exploitation of existing ideas (Axelrod and Cohen 2001; March 1991). The best response process most rapidly exploits existing ideas in the population, while the adaptive process more broadly explores the full range of strategies potentially available to the population.

Given an initial assignment of strategies to each player, the learning process and learning network fully specify an evolutionary process; as individuals learn from established contacts within their own population, the range and frequency of strategies coevolve in opposing populations. The coevolution of strategies thus determines the level of mutual cooperation achievable in the exchanges across populations.

## The Evolution of Cooperation

Our inquiry builds on the extensive social dilemma (reviewed in Gott, Pohill, and Law 2003) and evolutionary

<sup>2</sup>In our representation of innovation new components are adopted stochastically, not through analysis and planning. Even this stochastic innovation can be intelligent in the sense that unsuccessful innovations are rapidly abandoned. Innovation in our model implies a long term shift in strategy that differs from noise or error in implementing a strategy (Fudenberg and Maskin 1990), although both processes reduce the likelihood that a population will remain fixated on an inferior strategy (Kandori, Mailath, and Rob 1993; Wu and Axelrod 1995).

dynamics (Szabó and Fáth 2007) literatures that focus on the emergence of mutual cooperation among individuals playing IPD. A political scientist, Axelrod (1984) helped establish the evolutionary analysis of cooperation in his inquiry into international conflict, although our discipline has tended to ignore extensive new developments in biology, physics, and sociology. Axelrod noted that *nice* (never defect first), *retaliatory* (defect whenever a defection is encountered) strategies like tit-for-tat (TFT) provide a basis for the evolution of cooperation if the shadow of the future and cooperative payoffs are sufficient. We will call these strategies *retaliators*, will call all remaining nice strategies that do not retaliate *altruists*, and will call all nonnice strategies that defect first *exploiters*.

Axelrod transformed the study of cooperation from the choice of whether or not to cooperate in the prisoner's dilemma stage game to the choice of strategies to use in the IPD, as illustrated in Table 1. The upper panel presents a simplified representation of the underlying prisoner's dilemma (PD) stage game payoffs to row player. Any PD payoff matrix can be linearly transformed by first subtracting the sucker's (*S*) payoff from all payoffs to set  $S = 0$  and dividing by the resultant temptation (*T*) payoff to set  $T = 1$ . By setting the penalty payoff  $P = x$  and the cooperation reward  $R = 1 - x$  and limiting  $x$  to the  $[0, 0.5]$  interval, all PD games with symmetric differences between temptation and sucker payoffs ( $T - R = P - S$ )

TABLE 1 Payoffs for Altruists, Retaliators, and Exploiters

Stage Game Payoff in Prisoner's Dilemma			
	Cooperate	Defect	
Cooperate	$1 - x$	0	
Defect	1	$x$	
Strategy Selection Payoffs $r = 5$ repetitions, $x = .325$			
	ALLC	TFT	ALLD
ALLC	3.375	3.375	0
TFT	3.375	3.375	1.3
ALLD	5	2.3	1.625
Generalized Strategy Selection Payoffs			
	Altruist	Retaliator	Exploiter
Altruist	1	1	$x_a$
Retaliator	1	1	$x_r$
Exploiter	$y_a$	$y_r$	0

Conditions:  $x_a < x_r < 0, y_a > 1 > y_r > 0$

can be portrayed with just one parameter (Schmidt et al. 2001). The advantage of cooperation increases as  $x$  approaches zero.

The middle panel illustrates how payoffs in the IPD for any pair of opposing strategies can be calculated for any given values of  $x$  and  $r$ , the number of repetitions that determines the length of the learning period after which strategies are evaluated. The example uses  $r = 5$  to represent rapid learning cycles expected of the short, intermittent interactions in new exchange situations. Payoffs are given for row player when confronting the column player for three well-known strategies: an altruist who always cooperates (AllC), a retaliator who cooperates initially and only defects whenever the opponent defected in the previous round (TFT), and an exploiter who always defects (AllD). Total payoffs for adopting a given strategy for a given population in the transformed strategy selection game can be calculated by multiplying each payoff in Table 1, respectively, with the frequency of encounter with each respective strategy.

The strategy selection game transforms the original PD game into a coordination game with multiple equilibria. Note that TFT is both the payoff- and risk-dominant equilibrium in Table 1 when paired only with AllD, since the total payoff for TFT will exceed that of AllD as long as TFT represents over 50% of opponents. Axelrod (1984) introduced a spatial dimension by noting that clusters of TFT players can invade an AllD population as long as clustering increases the frequency of mutual TFT contacts enough to ensure that the invading TFT players have higher payoffs than the dominant AllD population, thus giving TFT a selective advantage; for payoffs in Table 1, TFT will outscore AllD whenever TFT is at least 34% of the population. Indeed, Bendor and Swistak (1997) demonstrate that nice, retaliatory strategies like TFT gain the greatest advantage of clustering compared with all possible strategies whenever selection processes reflect the standard proportional fitness models of evolution. They find that retaliators can ensure their survival under all potential games and payoffs if they account for at least half the encounters.

Since the three strategy types will be important for our analysis, the lower panel in Table 1 normalizes and generalizes the payoffs from the middle panel to illustrate our definition of the critical differences between retaliators, altruists, and exploiters.<sup>3</sup> First, altruists and retaliators both gain the full benefit of continued mutual

<sup>3</sup>The difference between the advantages of mutual cooperation obtained whenever altruists or retaliators meet ( $=1$ ) and the baseline obtained when exploiters meet ( $=0$ ) provides the basis for normalization. For example, the payoffs in the middle panel would be normalized by subtracting the AllD \* AllD payoff of 1.625

cooperation ( $=1$ ) whenever encountering their own or the other strategy. Exploiters encountering their own or the other strategy. Exploiters encountering other exploiters gain a lesser amount ( $=0$ ) reflecting their continued mutual defection, the Nash equilibrium of the stage game. Second, exploiters are their own best response because they do better against themselves than either altruists or retaliators do against them ( $x_a < x_r < 0$ ). However, retaliators are also their own best response as long as  $y_r < 1$ : in the central panel payoffs, for example, the retaliatory TFT will get a higher payoff than AllD when playing against TFT (3.375 vs. 2.3). Altruists, on the other hand, are dominated by exploiters ( $y_a > 1$ ) to reflect their exploitability; retaliators are not as exploitable as altruists ( $y_a > y_r$ ), although exploiters still do better against retaliators than against other exploiters ( $y_r > 0$ ).

## Markov Evolutionary Models in Finite Populations

Only exploiters are evolutionary stable (ESS) in standard evolutionary models (e.g., Nowak 2006), but in the evolution of cooperation, being discrete really matters (Szabó and Fáth 2007). That is, “traditional ESS and Nash conditions are neither necessary nor sufficient to imply protection by selection in finite populations” (Nowak 2006, 122). Finite-population models featuring stochastic replicator processes demonstrate that a single TFT can invade a full population of AllD (Nowak et al 2004), a finding replicated for several finite-population translations of the classic proportional replicator dynamic (Szabó and Fáth 2007). The proof is based on Markov chain analysis of replicator dynamics in which the probability  $\rho$  can be calculated that a single invading strategy can replace the full finite population.<sup>4</sup> Intuitively, a lone TFT in an AllD population fares very poorly, but still has some nonzero probability  $\rho_{TFT}$  of reproducing and hence adding

and dividing by the resultant payoff for TFT \* TFT encounters ( $3.375 - 1.625 = 1.650$ ).

<sup>4</sup>In the Moran selection process, nodes are selected to reproduce in proportion to their overall score after each round, and the selected node's strategy replaces the strategy of a randomly selected connected player (to keep population size constant). For this process, the probability  $\rho_{ji}$  that the no-mutation process starting with a full population of strategy  $i$  will be absorbed in a full population of strategy  $j$  is given by:

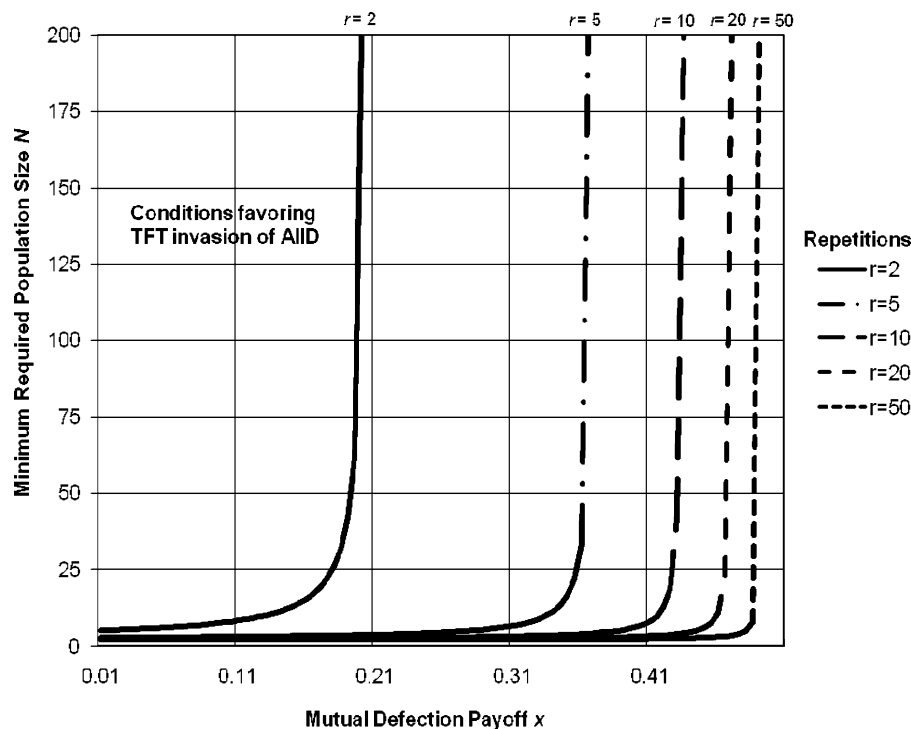
$$\rho_{ji} = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{l=1}^k \frac{g_l(i, j)}{f_l(i, j)}}$$

where

$$f_l = (i, j) = \frac{a_{ii}(l-1) + a_{ij}(N-1)}{N-1},$$

$$g_l = (i, j) = \frac{a_{ji}l + a_{jj}(N-l-1)}{N-1}.$$

FIGURE 1 The Impact of Payoffs, Repetitions, and Population Size on Boundary Condition for TFT to invade AllD



another TFT in the next period, and another in the following period, and so on until they take over the full population. For payoffs of the strategy selection game in the center panel of Table 1, the probability of taking over the whole population will be considerably higher for TFT than the probability  $\rho_{AllD}$  that a single AllD will take over a full population of TFT; TFT will outscore AllD as soon as TFT makes up at least 33% of the population, whereas AllD will outscore TFT only after it reaches 66% of the population. For population size  $N$ , selection favors cooperation as long as  $\rho_{TFT} > 1/N > \rho_{AllD}$ .

Nowak et al. (2004) calculate the boundary condition for this selection advantage in terms of payoffs and population size for any two strategies, as illustrated in Figure 1 for TFT and AllD.<sup>5</sup> The horizontal axis represents payoffs in the stage game,  $x$ , and the vertical axis represents the size of the population,  $N$ , up to 200, the range of interest for our study. Each solid line portrays the boundary condition for  $r$  repetitions of the stage game, with  $r$  labeled at the top of the figure above the line. Selection favors TFT

invading an AllD population for all values to the left and above the lines, labeled as “Conditions favoring TFT” for  $r = 2$  in Figure 1.

For a given population size, more repetitions (higher  $r$ , or greater “shadow of the future”) can compensate for less favorable payoffs; at  $r = 50$ , almost any payoff will favor the evolution of cooperation. For a given number of repetitions, the range of payoffs that favor cooperation increases rapidly with  $N$  for very small populations, but the marginal increase diminishes considerable after about  $N = 25$ . As population size increases beyond the range of Figure 1, however, the likelihood that a single invader could invade becomes increasingly smaller. Selection no longer favors cooperation as populations approach one thousand or more, as is consistent with the infinite population results (Nowak et al. 2004).

### Mutations and the Problem of Altruism

Markov chain selection processes with no mutations will eventually be absorbed into states with pure populations of a single strategy. In this case, when selection favors cooperation the state of full cooperation will be reached more frequently than when noncooperation is favored. With mutation pure populations are no longer absorbing, since mutations introduce new strategies that

<sup>5</sup>For a given  $N$  and sufficiently weak selection pressures, selection favors TFT replacing AllD if  $a(N-2) + b(2N-1) > c(N+1) + d(2N-4)$ , where the payoffs are  $a$  and  $b$  for TFT,  $c$  and  $d$  for AllD, and thus  $a = n(1-x)$ ,  $b = (n-1)x$ ,  $c = 1 + (n-1)x$ , and  $d = nx$ . Solving for  $N$  at the boundary condition gives  $N = (2x-1)/[x - \{(n-1)/(1-2n)\}]$ .

have some small potential of taking over the full population. In this case, the ratios  $\rho_{TFT}/(\rho_{TFT} + \rho_{AllD})$  and  $\rho_{AllD}/(\rho_{TFT} + \rho_{AllD})$  approximate the respective proportion of time spent in pure populations of TFT and AllD after all effects from initial conditions have been eliminated: that is, they provide the invariant frequency distribution of the Markov chain as mutation rates approach zero (Fudenberg and Imhof 2006). Using the payoffs in the middle panel of Table 1 for TFT and AllD, for example, a Markov chain with  $N = 10$  would average about 67% of its time in full TFT populations with full mutual cooperation and the remaining time in full AllD populations with no cooperation at all, so we would on average observe mutual cooperation 67% of the time. If the population grows to  $N = 20$ , the probabilities are even more favorable to cooperation, with mutual cooperation expected on average 96% of the time.

The presence of altruists compromises the selective advantage of retaliators because these neutral mutants do as well as TFT in mutual encounters, but provide support for exploiters when they meet. This support in turn reduces the selection advantage of retaliators (Bowles 2004; Imhof, Fudenberg, and Nowak 2005). The invariant distribution in a three-strategy process that includes the altruistic AllC strategy for the same payoffs and population  $N = 10$  would now include 55% AllD, 38% TFT, and 7% AllC. By including altruists, the average expected level of mutual cooperation by nice strategies drops from 67% to 45%. Since it is unclear how well these Markov chain processes reflect the learning processes and learning networks we model, we will return to this problem of altruists in the results section.

## Learning Networks and the Coevolution of Cooperation

To understand the impact of learning networks on the evolution of cooperation, we turn from homogeneous population models to those that use graph theory to represent local network structures. In these models, vertices or nodes generally represent players, and links connecting the nodes (called edges or arcs) indicate interactions (Ohtsuki and Nowak 2008; Szabó and Fáth 2007). The game network specifies who plays with whom; most models assume that players learn about other strategies only from those they play, so they assume that the learning network (who learns from whom) and game network are the same. Based on these assumptions, simulation studies have shown that game networks can provide sufficient clustering to favor cooperation for two-dimensional lattices (Axelrod (1997), lattices with random (small world)

connections (Macy and Skvoretz 1998), and even for stable random graphs (Cohen, Riolo, and Axelrod 1998)—the structures that we will analyze.

The limited contacts determined by local structure in graphs provide cooperative advantages much as do limited populations in finite models. Ohtsuki et al. (2006), for example, show analytically that cooperation can evolve even for simple cooperate or defect strategies in one-shot PD games on infinite lattices.<sup>6</sup> Ohtsuki, Pacheco, and Nowak (2007) extend this lattice-based analysis to separate networks for games and learning and demonstrate that when both networks overlap the network with the largest degree determines the required cooperative benefit necessary to attain cooperation. Thus greater connectivity or degree of the learning network does not affect cooperation if it is smaller than the game network, but works against cooperation if it is larger.<sup>7</sup> Both analyses suggest that poorly connected learning networks support higher levels of cooperation, although simulation studies of IPD on two-dimensional lattices have provided mixed support for this conclusion.<sup>8</sup>

These results cannot be applied directly to our learning problem because their Moran process and simple two-strategy evolution is not particularly suited for a learning model, because the regular graphs they rely on provide a very limited representation of learning networks, and particularly because there is no overlap between game and learning networks in our two-population learning problem; if at all applicable, the implication might be that cooperative benefits would have to approach infinity as the proportion of overlap approaches zero. To provide a more adequate foundation for hypotheses about the impact of learning network structure on cooperation, we next consider three of the most common characteristics of graphs—size, degree, and clustering (Wasserman and Faust 1993)—from the perspective of the exploitation-exploration tradeoff discussed previously for the learning process (cf. Lazer and Friedman 2007).

<sup>6</sup>Selection favors cooperation on lattices of degree  $k$  (all players play  $k$  games) whenever the payoff  $x < 1/(1 + k)$ , based on the same Moran selection process noted above. The  $b/c$  payoffs in Ohtsuki et al. (2006) equal  $(1 - x)/x$  for the PD parameters in Table 1, so the payoffs for  $x = .325$  would barely support cooperation for  $k = 2$  games per person in this one-shot setting.

<sup>7</sup>Selection favors cooperation for regular graphs when  $x < 1/(1 + h * g/l)$ , where  $h$ ,  $g$  and  $l$  represent, respectively, the degree of the game (“interaction”), learning (“replacement”), and the overlap between the two networks.

<sup>8</sup>Hoffmann and Waring (1998) support Ohtsuki, Pacheco, and Nowak’s analysis by showing that a larger learning network coupled with a smaller game network produces the least cooperation, while Eschel et al. (2000) found greater cooperation when the learning network was larger than the game network.

## Network Size and Degree Increase Exploitation, but Clustering Does Not

Graph characteristics that enhance the set of available strategies and accelerate the diffusion of highest-scoring strategies would provide a network function that maximally exploits the best available strategy. Sampling theory suggests that a larger *size* (number of nodes in full network) would provide a higher probability of including high-scoring strategies, assuming as the model does that the initial strategies in the learning process are independently and randomly distributed across nodes. Similarly, a higher *degree* (or number of learning links per node) increases the likelihood that the highest-scoring strategy will be in each node's learning set. Put differently, a higher degree decreases the average number of steps (or path length) required for the highest-scoring strategy to reach every node. Hence we would expect larger size and higher degree to be associated with the ability to rapidly exploit existing high-scoring strategies.

Clustering for a given node is defined as the proportion of all linked nodes that share links with each other, or the proportion of a player's learning contacts that learn from each other. Clustering provides redundancy at the cost of a broader reach: if all nodes have a degree of 4, for example, a node with fully clustered relationships can reach only eight nodes in two steps, compared with 16 nodes if none of the connected nodes were connected with each other. Redundancy reduces transmission efficiency, although it may enhance accuracy when transmission errors or congestion cause problems. Granovetter (1973) found that successful searches for new jobs relied on nonredundant "weak ties" that brought in a wider range of information. In short, more clustering leads to slower diffusion and hence less exploitation.

## Clustering Increases Exploration

The featured role of highly clustered, overlapping networks of reciprocation in the resolution of social dilemmas (Coleman 1988; Granovetter 1985; Putnam 1993) is perhaps most commonly understood in term of the game networks, where clustering provides the previously discussed and well-known advantages for cooperative strategies. Yet some of the common arguments in support of social capital refer also to learning effects relevant to the two-population problem. Clusters of individuals interlinked by overlapping strong ties can develop shared values, mutual cooperation, and trust. Within clusters, like-minded neighbors can more readily share experiences about encounters with the other population and explore

alternative strategies for dealing with them. In conjunction with individual learning processes that emphasize innovation and exploration, clustered learning networks in which the local best-response strategy is altered and recombined in multiple ways may provide creative crucibles required for developing robust cooperative strategies where none exist.

## Small Worlds Enhance Exploitation and Exploration

Watts and Strogatz (1998) and Watts (1999) argue that small-world networks combine the clustering advantages of ring networks with the efficient transmission advantage of random graphs. They also provide a more realistic representation of observed social networks than either of these theoretically interesting graphs. The *ring* network links each node symmetrically to its  $k/2$  nearest neighbors on each side in a ring or one-dimensional lattice, where  $k$  is the number of links per node. Whenever nodes have more than two links, the overlapping links provide a clustered neighborhood in which neighbors of neighbors are linked. The *random* network links each node to every other node with an equal probability, which provides much shorter path lengths than rings with the same number of nodes and links. Randomly assigned links also ensure low clustering since one node's neighbors have no more than average probability of being linked together.

The *small world* network begins with the ring structure, but then randomly assigns each existing tie with probability  $\beta$  to a randomly selected node to which the initial node is not already linked.<sup>9</sup> Watts (1999) shows that, as  $\beta$  increases, average path length initially decreases very rapidly while clustering decreases only slowly. Thus small-world networks with low values of  $\beta$  may simultaneously offer the global exploitation advantages of short path lengths in random networks and the local exploration advantages of higher clustering in ring networks. If the exploration-exploitation tradeoff affects the evolution of cooperation, small-world learning structures may provide relatively high levels of both.

We investigate the relative importance of exploration and exploitation by comparing these three types of networks for different population sizes and degrees. The exploitation hypothesis implies that cooperation will be more prevalent in larger, better connected, less clustered populations. On the other hand, the exploration hypothesis implies that cooperation will increase with clustering,

<sup>9</sup>See Newman (2003) for discussion of this and alternative approaches.

while the exploration-exploitation tradeoff hypothesis suggests that the small-world network will outperform the other two.

## The Agent-Based Simulation Experiment

Agent-based models represent the decision and learning processes for each agent in a computer program, using fixed parameters that reflect the model's assumptions and variable parameters that represent the experimental conditions to be manipulated (Axelrod 1997). In our model, we vary the learning process and learning network structure to evaluate their impact on levels of mutual cooperation in IPD games played across populations. Computer simulation has played a particularly useful role the analysis of social dilemmas in network settings that are too complex for currently available analytic models (Gott, Pohill, and Law 2003; Szabó and Fátih 2007). Our simulation experiment is based on a discrete-time, autonomous, finite-population, stochastic adjustment model (Fudenberg and Levine 1998, 137) involving games played between two populations where strategies are learned within each population. We first review how the model represents the basic assumptions discussed previously and then describe the experimental design.

*Strategy.* To represent an initial learning situation in which many potential strategies may exist, including a variety of retaliatory, altruistic, and exploitative strategies, we randomly assign to each player (the agent) with equal probability one strategy from the complete set of 32 pure strategies with a one-period memory (Gott, Pohill, and Law 2003). The learning process then potentially modifies each player's strategy in each learning cycle, which we will refer to as a learning *period*.

The genetic coding for each strategy specifies whether to cooperate ( $C = 1$ ) or defect ( $D = 0$ ) in the first period and then whether to cooperate or defect after each of the four possible outcomes in the previous period  $\{DD, CD, DC, CC\}$ , with the player's choice indicated first and the opponent's choice second. We can therefore code a player's strategy with five binary variables or bits to represent the total strategy space of  $2^5 = 32$  strategies. The first bit indicates the first-round choice, and the remaining four bits encode instructions for the four possible outcomes of prior interaction, giving a five-bit representation of the strategy (*Nice*, *DD*, *CD*, *DC*, *CC*),

For example, (0, 0, 0, 0, 0) represents the AllD strategy that always defects, while (1, 1, 1, 1, 1) represents the AllC strategy that always cooperates. The eight nice strategies

that will never be the first to defect are represented by (1,  $z$ ,  $z$ ,  $z$ , 1), where  $z$  can take the values of either 0 or 1. Of these nice strategies, the two retaliatory strategies that always retaliate after the other side defects include TFT (1, 0, 0, 1, 1), which cooperates in the first round and then cooperates whenever the opponent cooperated in the previous round, and GRIM (1, 0, 0, 0, 1), which cooperates in the first round, and then cooperates only if both players cooperated in the previous period.

*Learning.* The two learning processes outlined above are both based on each player's fitness score for the learning period, which is determined by the player's average payoffs for all  $r$  repetitions of the stage game across all games with the opposing population. The myopic best response process adopts the best-scoring strategy among a given player's contacts, as determined by the learning network. The adaptive learning process, on the other hand, copies each gene in the best-scoring strategy with a probability determined by

$$P = \frac{\text{Maxfitness}_i - \text{Fitness}_i}{\text{Maxfitness}_i} \quad (1)$$

where  $\text{Fitness}_i$  is the fitness of player ( $i$ ), and  $\text{Maxfitness}_i$  is the highest fitness score among the player's network contacts. This probability is zero if the player is the highest scoring or fittest among her contacts and increases monotonically as the gap between the player's fitness and the highest score expands. To further represent exploration and innovation, the genetic process includes a mutation at the end of the learning cycle (Gott, Pohill, and Law 2003; Kandori, Mailath, and Rob 1993) in which each gene may flip to its opposite value with a small probability (0.01 in our model). The transition probabilities defined by each process in conjunction with the a learning network structure and payoffs determines a unique Markov chain that determines the spatial distribution of strategies in each population at the end of each learning period based on the spatial distribution of strategies at the beginning of the period.

*Experimental Design.* There are three nested stages in the experimental design. Stage 1 consists of setting all fixed and variable parameters for a given experimental condition and then running stages 2 and 3 for each condition. We fix the payoff for mutual defection at  $x = .325$ ,  $r = 5$  after initial analysis discussed in the next section. To maintain tractability and focus on learning networks in the subject population, we also fix one population (representing authorities in our initial formulation) and vary parameters only in the other population (representing subjects). The number of authorities is fixed at  $N = 10$

with a modestly connected ( $k = 2$ ) ring learning network. We then systematically vary each network parameter for the subject population using four values for  $N \in \{25, 50, 100, 200\}$  and three for  $k \in \{2, 4, 8\}$  for the ring, small-world, and random networks, providing  $4 \times 3 \times 3 = 36$  experimental conditions to be evaluated. The range of parameter values were set to reflect the likely size and communication structures of regulatory authorities and firms within local jurisdictions analyzed in Scholz and Wang (2006).

Since ring networks have only limited clustering and therefore may not adequately test the impact of clustering, we also include two “caveman” or village networks for each population size, one with five and the other with 10 nodes per village;  $N = 25$  is omitted for villages of 10 because it cannot be constructed.<sup>10</sup> Within each village all nodes are interconnected, and two “gatekeeper” nodes are also linked to different neighboring villages, creating a ring of densely clustered villages each linked to two adjoining villages.

For each experimental condition in stage 1, stage 2 creates the appropriate network structure and assigns each player an initial randomly determined strategy. Since the model contains stochastic elements in network structures, strategy assignments, and the learning process, stage 2 is repeated multiple times for each experimental condition. The main reported results are averaged over 900 simulations per experimental condition and thus represent the expected level of cooperation for each experimental condition given the initial uniform distribution of strategies across nodes.

Stage 3 performs the individual simulation for each assignment of networks and strategies from stage 2. All players initially play five repetitions of the stage game with every player in the opposite partition. They then observe the outcomes associated with strategies of players they are linked to and use this information to update their strategy. All players again play five repetitions in the second period, again observe linked players and update their strategy, and so on. The learning cycle is repeated for 3000 periods to ensure that we reach a relatively steady state, and information about the level of mutual cooperation is recorded in each period. Note that the relatively small number of iterations in the IPD and the large number of learning cycles is consistent with our assumption that adaptive feedback learning rather than strategic calculation better characterizes the early contact situation

<sup>10</sup>The clustering coefficients for village networks are .86 and .96 for villages of 5 and 10 respectively, compared to .5 for rings with  $k = 4$  and .64 for  $k = 8$ , and compared to clustering coefficients ranging from .01 to .33 for random networks.

we model: players use simple strategies but adapt rapidly over a long period.

## Results

Our initial analysis probes whether stable levels of cooperation could emerge under any payoff conditions. We systematically vary payoffs in the range  $.20 \leq x \leq .40$  and compare the mean level of mutual cooperation (the proportion of *CC* outcomes) in the subject population achieved after 3,000 learning periods, averaged in this preliminary test over 500 simulations for each value of  $x$ .<sup>11</sup> The four values of  $x \leq 0.3$  achieve a proportion of mutual cooperation between .8 and .9 within the first 200 periods and remain at these levels, revealing the typical Markov chain pattern of an initial burn-in period influenced by the initial conditions followed by stable oscillations representing a typical Markov chain pattern.

When  $x > 0.35$ , cooperation never rises substantially, confirming an invariant distribution in which selection does not favor cooperation. The values between  $0.3 < x < 0.35$  define the most interesting transition zone in which cooperation is only moderately favored. Strong selection effects outside this transition zone leave little room for learning networks to make a difference; why consult with friends if the gains or losses from cross-population exchanges are readily apparent from personal experience? We select  $x = .325$  for the remaining experimental manipulations to analyze the effect of networks within the transition zone. For this payoff, selection is favored for  $N > 9$  in Figure 1, so even the fixed population of  $N = 10$  is just sufficient to favor cooperation in the finite population model. Network manipulations produce levels of cooperation ranging from a low of 59.5% ( $N = 25$ ,  $k = 2$ , *random*) to a high of 71.9% ( $N = 200$ ,  $k = 8$ , *ring*), which indicates the relative differences in cooperation attributable to differences in learning networks within the transition zone. Although specific levels of cooperation would undoubtedly differ, the observed patterns of results reported below should be generalizable to the comparable transition zones determined by other values of  $x$ ,  $r$ , and  $N$ .

### Myopic Best Response Does Not Support Cooperation

We next test the myopic best response learning process on all ring, random, and small-world networks, and find that

<sup>11</sup>Other parameters are fixed at  $r = 5$ ,  $N = 200$ ,  $k = 8$ , for ring network using the adaptive learning process.

achievable levels of mutual cooperation never rise above 15%. This learning process rapidly eliminates altruists and retaliators alike; on average, exploiters comprise 86% of the final population, and the remaining 14% is evenly split between retaliators and altruists. The evolution of strategies stop well before the 25<sup>th</sup> period; with rapid dissemination and no mutation, the initial uniform distribution of strategies determine which absorbing state is reached. With only 25% of the strategy types being nice, initial distribution apparently favored exploiters 86% of the time. Since low levels of cooperation associated with myopic best response provide little room for learning networks to matter, we only use the adaptive learning model to analyze the impact of learning network structure.

### Size, Degree, and Clustered Learning Networks Enhance Mutual Cooperation

The two graphs in Figure 2 report the proportion of mutual cooperation in the final 500 periods, averaged over 900 simulations for each experimental condition or 450,000 observations; this reflects the time-averaged invariant distribution of strategies associated with each condition's unique Markov chain. The symbols indicate the observed level of cooperation and lines connect observations of the same type. The upper graph reports the results as a function of population size and degree (averaged over all types of networks), while the lower graph reports the results for village, ring, small-world, and random networks for the indicated population size (averaged over all degrees).

The upper graph shows that cooperation increases with both population and degree, but at a declining rate. Population size consistently accounts for the greatest variance in both graphs. Holding degree constant, the level of cooperation rises with each increase in size, although the rate of change decreases as the population becomes larger and appears to level off around  $N = 100$ .<sup>12</sup> Cooperation also tends to increase with degree when holding population constant, although the differences between degrees of four and eight are small, suggesting a limiting advantage as degree increases beyond eight. Villages of degree five and 10 both outperform other networks with degree eight, and both of these village networks perform at very similar levels, suggesting that clustering also plays a role in promoting cooperation.

The lower graph confirms the importance of clustering based on a comparison of different network struc-

tures. For all sizes, the clustered village and ring networks provide the strongest support for cooperation. The advantage of the ring is more evident for smaller sizes, with differences that tend to disappear as the size increases to  $N = 100$ , as in the upper graph. The small-world results are not much lower than the ring, but clearly are not better; the expected advantage of the small world for the exploration-exploitation hypothesis is not supported.

Overall, the two graphs show that cooperation varies with all three dimensions, with the largest differences provided first by size, then by degree, and finally by the clustering of the network structure. Achieving higher levels of cooperation apparently require some blend of exploration and exploitation, although not the blend associated with the small-world network. The strong performance of the adaptive learning process and the clustered village network underscore the advantages associated with more global exploration. The gains associated with greater size and degree, on the other hand, support the role of rapid exploitation. To explain these results, we analyze the impact of each factor on the selection advantages of retaliators, altruists, and exploiters.

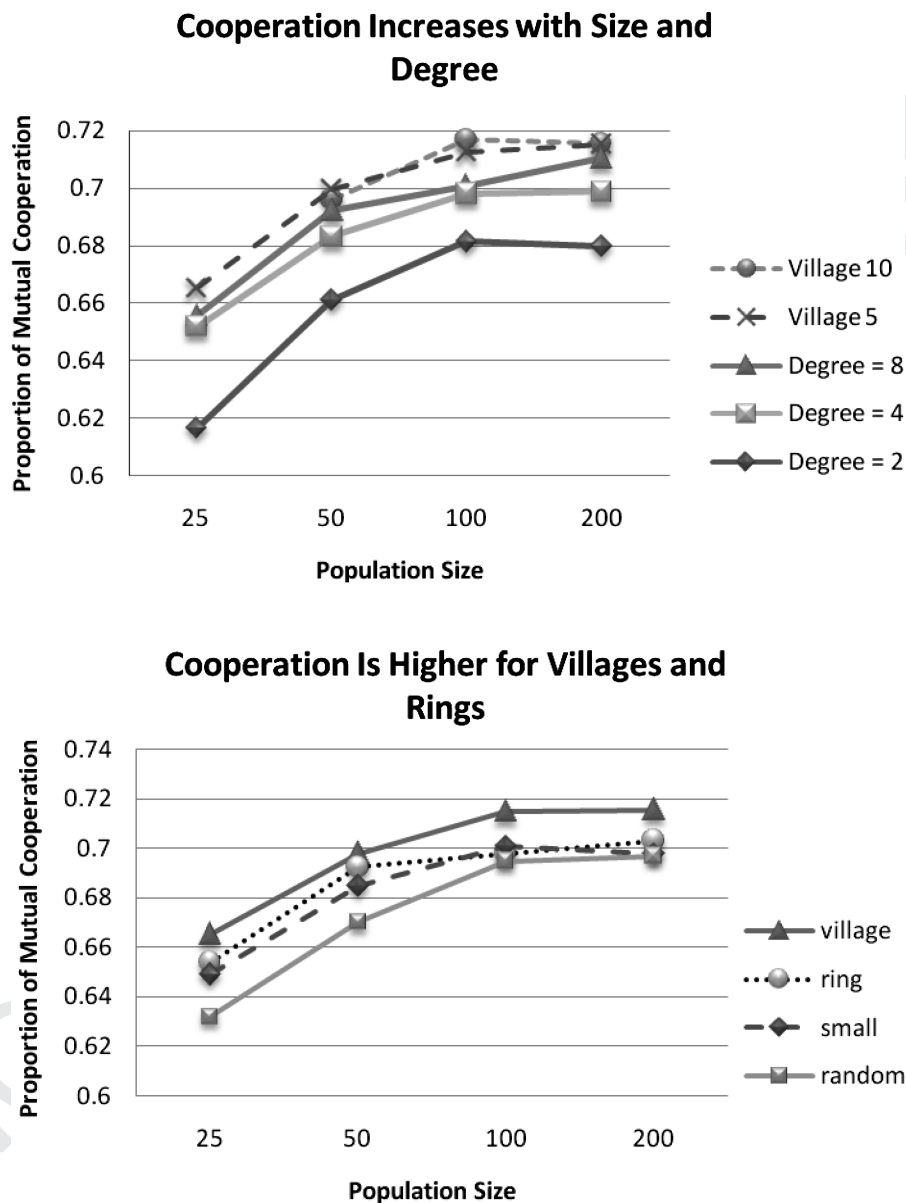
### Retaliatory Strategies, Exploration, and the Problem of Altruism

Figure 3 illustrates the critical role of retaliators and altruists in the evolution of mutual cooperation for one experimental condition. The solid line in Figure 3 reports mutual cooperation in each of the 3,000 periods averaged over 500 simulations, with greater detail shown in the left-panel for the first 25 periods. As in all conditions, mutual cooperation starts at an expected 25% that reflects the randomized uniform distribution of strategies, 25% of which are nice. Cooperation quickly drops to a very low level as exploiters rapidly replace exploitable nice strategies in the first 25 periods, after which it recovers and reaches a relatively stable plateau after the first 1,000–1,500 periods. Note that each period reports averages—most initial conditions lead rapidly to a full extinction of nice strategies, while the few favorable ones lead rapidly to high levels of cooperation.

The dashed line in Figure 3 plots the proportion of the population represented by the eight nice strategies that always cooperate with each other, which account for most of the observed mutual cooperation. Their proportion begins at the expected .25 (representing their proportion of strategy types), drops rapidly in the first few periods, and then expands following a path similar to that of cooperation. The dotted line charts the evolution of the retaliatory strategies, TFT and GRIM. Retaliators on average maintain their proportion of the population as the

<sup>12</sup>Nowak et al. (2004, Figure 1a,  $w = 1$ ) find that the selective advantage for cooperation peaks around  $N = 200$  for the related Moran process when selection effects are strong, as they are in our simulation.

FIGURE 2 Observed Levels of Mutual Cooperation

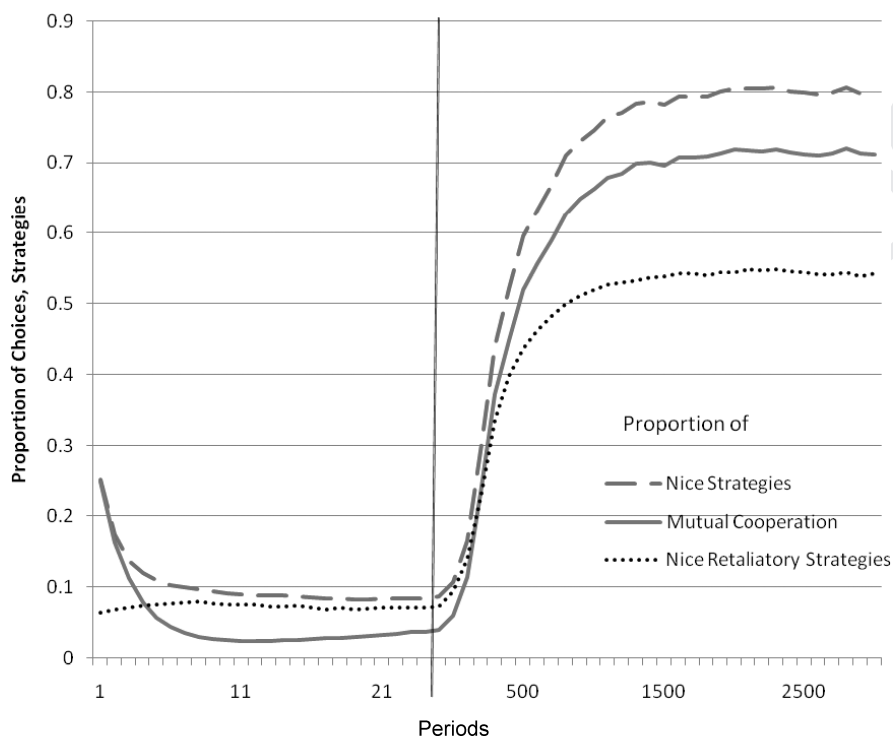


Note: Mean values for last 500 periods, averaged over 900 runs per condition.

altruistic nice strategies are eliminated and account for almost the entire population of nice strategies in this initial hostile environment dominated by exploiters. Once favorable distributions emerge, the selection advantage of retaliators increases their proportion to just over half the population.

Altruists expand almost as rapidly as retaliators in the more favorable environment created by retaliators. In the final conditions reflecting the invariant distribution, the observed population consists of 54% retaliators, 26% altruists, and 20% exploiters. To explain this expansion of altruists, Figure 4 shows the probability that TFT

genes will be copied by AllC and AllD for different proportions of each strategy, calculated using the adaptive learning algorithm (equation 1). The horizontal axis represents the proportion of altruists (AllC), and the vertical axis indicates the corresponding probability of copying a gene from the retaliatory strategy (TFT). Retaliators outscore exploiters for all values in the figure, so selection always favors retaliators. The lines indicate the probability of copying by altruists (the three dotted lines) and exploiters (the three solid lines) for populations with proportions of retaliators  $p(t) = .3$  (indicated by circles),  $.5$  (triangles), and  $.7$  (squares), respectively.

**FIGURE 3 Nice Strategies Explain the Evolution of Cooperation**

Note: Solid line represents the mean proportion of mutual cooperation averaged over 500 simulations for small world with  $N = 200$  and  $k = 8$  for each of the 3,000 generations. The dotted line presents the proportion of all strategies that are retaliators, and the dashed line presents the proportion that are nice strategies (including retaliators). The vertical line separates the expanded scale for the early periods from the compressed scale for the later periods.

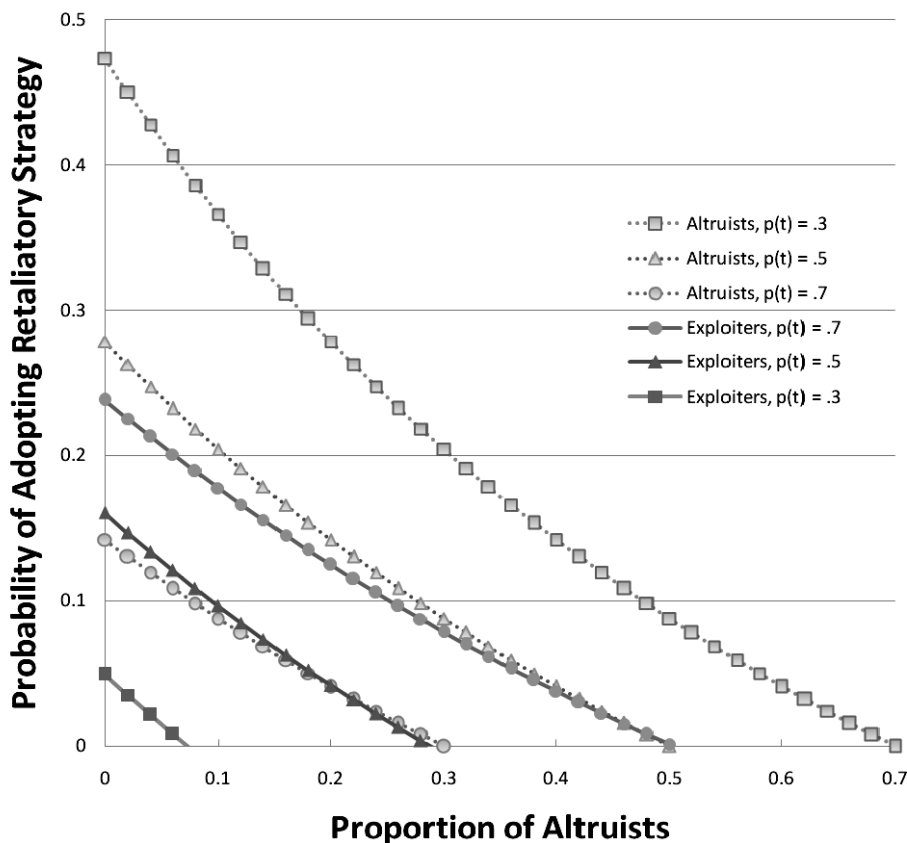
The downward-sloping lines indicate that selection pressure against altruists and exploiters decrease as altruists increase. More importantly, selection pressure increases against exploiters but decreases against altruists as retaliators increase from .3 (circles) to .5 (triangles) and .7 (squares). When retaliators compose half the population and retaliators and altruists are equally divided at 25% each in Figure 4, the probability of copying from TFT is .11 for altruists versus .02 for exploiters. If the proportion of retaliators increases to 70% and altruists remain at their final average of 25%, the probability of conversion reverses to .02 for altruists and .10 for exploiters.

Any deterministic selection process would produce a full population of retaliators from any distribution in Figure 4, but the mutation and proportionate fitness features of the adaptive learning process at some point counterbalance the retaliators' selection advantage over altruists.<sup>13</sup>

<sup>13</sup>Overall, mutation and selective learning appear equally important in changing strategies: 10 randomly selected agents made on average 142 learning and 131 mutation changes in strategy over the 3,000 periods, or a rate of .047 and .043 per period, respectively.

Selective learning allows at least some exploiters to copy nice genes but not retaliatory ones, resulting in an increase in altruists rather than retaliators at the expense of exploiters; the increase in altruists reduces selection pressures on both exploiters and altruists. In addition, as retaliators grow in number, mutation converts more retaliators to altruists than altruists to retaliators; at some point, mutation conversions outweigh selection conversions, thereby limiting the growth of retaliators. Mutation also produces exploiters, but selection pressures increase against exploiters as retaliators increase and are not counterbalanced by mutation. In sum, the combination of diminishing selection pressures, selective learning, and mutation allow the population of altruists to survive, which in turn allows evaders to survive as well. The stochastic learning process produces oscillations driven initially by the increase in retaliators, with occasional expansion of altruists to the point where exploiters outscore retaliators, followed by decreased in the altruist and retaliator populations until retaliators again outscore exploiters, and so on.

FIGURE 4 Probability of Adopting Retaliatory Strategy Decreases with the Proportion of Altruists



Lines indicate the probability that Altruists and Exploiters will adopt genes from the retaliatory strategy for three fixed proportions of retaliatory strategies  $p(t)$ , based on payoffs for Altruist = AllC, Retaliator = TFT, Exploiter = AllD for  $x = .325$ ,  $r = 5$ .

Thus innovative learning processes provide both advantages and limitations to the evolution of cooperation. Although innovation in the form of mutation aids and abets altruism in limiting achievable levels of cooperation, it is critical in early periods to escape the dominance of exploiters. Most of the runs we observed crashed to all-exploiter populations several times in the first thousand periods, and hence relied on mutation to reintroduce retaliators and eventually achieve higher levels of cooperation. Far from being a peculiar artifact of evolutionary models, mutation represents a critical component of the process of learning to cooperate that deserves further exploration.

### The Role of Learning Networks

Learning processes that enhance exploration rather than exploitation appear to be more conducive to cooperation, but structures that enhance exploitation appear to

be more advantageous for learning networks based on the results in Figure 2. Population size and degree increase cooperation, and basic probability theory suggests the reason that they can enhance the selection advantage for retaliatory strategies. The likelihood that mutation will produce a retaliator in an all-exploiter population is proportional to the population size times the mutation rate, so larger populations provide the higher likelihood of retaliators that is critical for successful invasion (Nowak 2006). Furthermore, the likelihood that a retaliator will be in the learning set of the average player increases exponentially with the number of partners, accelerating the rate of expansion whenever retaliators have achieved a selection advantage. For example, once TFT exceeds the threshold frequency and becomes the highest scorer, the probability that a randomly drawn learning set will NOT include TFT is .75 for only one partner,  $.75 * .75 = .56$  for two partners, .1 for four partners, and miniscule for eight partners.

Yet even in learning networks, exploration appears to play some role in enhancing cooperation, since we have seen that clustering increases cooperation particularly for lower values of size and degree. Clustering creates redundant links, which reduces the spread of successful strategies for a given degree. Although this retards the spread of retaliators when they are in the minority, it also stabilizes the population when they are in the majority. Small populations with small learning sets are more prone to exaggerate stochastic disturbances, so clustering can dampen the oscillations between expanding altruist and exploiter populations in these conditions once high levels of cooperation have been achieved.<sup>14</sup> Of course, clustering also retards the initial growth of cooperation and recovery from occasional successful invasions of exploiters,<sup>15</sup> but the overall dampening effect still favors cooperation since the evolutionary process spends more time at higher levels of cooperation once initial conditions become less important. This defensive role of clustering appears similar to the role of conformist norms in Bowles (2004), which helped to prevent the invasion of “grabbers” in more cooperative populations.

We believe that these findings apply well beyond our simulation model. The simulation modeled the initial development of cooperation, but the stability of cooperation around an apparent Markov chain invariant distribution suggests that the results apply equally to the maintenance of cooperation. The simulation included two populations, but the ten encounters with agents in the “authority” population could just as well represent random encounters with other agents within a single population, particularly since observed average frequency distributions of strategies in both populations were almost identical. The simulation explored only one value of  $x$  and  $r$ , but the same factors should influence the evolutionary process at other transition zones as indicated in Figure 1. Similarly, an increase in the size of the authority population (or equivalently an increase in the number of agents encountered each period) should also influence the values of  $x$  and  $r$  that determine the transition zone, but not the

<sup>14</sup>In the last 1,000 periods, the two smaller ring networks have a higher average proportion of retaliators (.50 vs. .48) and also a lower variance in population frequencies ( $7 * e^{-5}$  vs.  $9 * e^{-5}$ ) compared with the random networks, suggesting that clustering in ring networks constrains the oscillations around the mean level of cooperation. Across all conditions, mean cooperation is negatively correlated to variance ( $-.54$ ) in the last 1,000 periods, suggesting that tighter constraints play an important role in sustaining higher proportions of retaliators.

<sup>15</sup>Ring networks have lower proportions (.07 vs. .09) and lower variance ( $3 * e^{-6}$  vs.  $8 * e^{-6}$ ) of retaliators in the earlier periods dominated by exploiters (periods 20 to 30) compared to the less-clustered random networks.

evolutionary process within that zone. Finally, the limitation on population size capable of supporting cooperation suggested in finite-population models apply only when every agent at least potentially exchanges with every other agent; the structuring of exchange interactions into game networks can provide an alternative means of partitioning even large populations in a manner capable of supporting cooperation. To further explore the interaction between such game network structures and the learning networks analyzed here, our findings suggest that models including retaliators, altruists, and exploiter payoffs as defined in Table 1 would be sufficient to explore the critical evolutionary process.

## Conclusion

We find that the way individuals get and process information about exchanges strongly influences the coevolution of cooperation as two populations warily seek to establish and maintain mutually advantageous exchanges. An adaptive learning process that is not too quick to exploit existing knowledge is particularly important for achieving high levels of cooperation, given the relatively hostile and diverse initial environments where learning commences. The best-response learning process in which individuals immediately adopt the highest-scoring strategy available to them produces little cooperation here; everyone quickly adopts the exploitative strategies favored in most initial distribution of strategies, and the learning process has no innovative feature to reintroduce the retaliatory strategies necessary to support cooperation. Best-response learning can accelerate the evolution of cooperation, but only if retaliatory strategies are dominant; adaptive learning, on the other hand, can achieve high levels of cooperation even when exploitative strategies dominate initially because it is both innovative and more selective in learning from higher-scoring strategies. Intelligent innovation and learning may be helpful, but even the simple-minded stochastic process in our model of trying new strategies and retaining successful ones plays a critical role in learning to cooperate. In our earlier examples, membership turnover in the legislature, newly activated stakeholders in policy arenas, and increased rate of immigration may all accelerate the initial evolution of cooperation to the extent that they enhance exploration through the introduction of new strategies and the enhanced willingness of new arrivals to consider a range of strategies.

Innovation is critical in hostile early environments, but eventually limits the average level of attainable cooperation by abetting the problem of altruism. Without altruists, retaliators could fully dominate the population

and sustain higher levels of cooperation. But altruists are neutral mutants of retaliators who do as well against retaliators, but fail to retaliate against exploiters; neutral drift through mutation expands the altruist population, creating cycles in which exploiters subsequently expand and decline. The rate at which nice strategies lose their retaliatory edge thus determines the attainable level of cooperation. Institutions that encourage conformity and hence constrain innovation could enhance cooperation once retaliators are dominant (Bowles 2004), but conformity would also diminish the likelihood that sustainable cooperation would coevolve in the first place. Thus stability rather than turnover in the legislature, in policy arenas, and in existing immigrant communities becomes valuable to sustain cooperation once it has been achieved.

Learning networks determine the spread of successful strategies; a larger, more connected population with more overlapping clusters of relationships averages 72% mutual cooperation in our simulation, a 12 percentage point gain over the 60% averaged by the smallest, least connected, least clustered population. Larger populations and more contacts both increase the chance of learning from successful retaliatory strategies and hence increase the growth of retaliatory strategies whenever selection favors cooperation. Thus larger populations with more developed learning relationships should on average exhibit higher levels of cooperation. This would suggest that cooperation across longstanding cleavages would develop more rapidly in larger interactive legislatures, larger more established policy arenas, and larger integrated immigrant communities—in the latter case, however, analytic results for finite populations suggest that at some point larger populations may impede the evolution of cooperation unless interactions are restricted into structured game networks.

Clustering in learning networks does not hasten the development of superior strategies, but appears to decrease the ability of exploiters to invade once high levels of cooperation have been established, an advantage that is most apparent in smaller, less-connected populations that are more prone to such invasions. Thus clustering, like conformity constraint, is more valuable in the later rather than earlier stages of learning to cooperate and is most important for smaller population size and degree. In our examples, the clustering of legislators into geographic or policy subgroups within each party, of stakeholders into multiple policy study groups and planning venues within each policy arena, of immigrants and existing citizens into segregated neighborhood organizations or church groups may potentially serve this function, at least to the extent

that these clusters increase the redundancy of existing ties within the population.

Of course, the evolutionary advantages of cooperation identified in our individual learning model can be accelerated or retarded by existing institutions and coalitional activities. For example, if the era of partisan confrontation were to end in the U.S. Congress, our model suggests that bipartisan cooperation would be more likely to emerge in the larger House than in the smaller Senate, *ceteris paribus*. Differences in formal authority and party strategies that enhanced partisan conflicts to begin with, however, are likely to overshadow individual learning processes in shaping any possible bipartisan cooperation. Similarly, in particularly volatile situations in which even a small number of individual defections from either group will trigger an escalation of ethnic conflict and a complete breakdown of intergroup relations, the development of effective “ingroup policing” mechanisms or a coordinated fear of triggering the “spiral equilibria” to prevent defections (Fearon and Laitin 1996) may be necessary as a parallel process in support of the individual learning emphasized in our model. Contextual features such as these are important to consider when applying our basic model and incorporating important classes of features into future models can provide a stronger foundation for understanding the evolution of cooperation in the broad range of settings encountered in political science.

## References

- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.
- Axelrod, Robert, and Michael Cohen. 2001. *Harnessing Complexity: Organizational Implications of a Scientific Frontier*. New York: Free Press.
- Bala, Venkatesh, and Sanjeev Goyal. 1998. “Learning from Neighbours.” *Review of Economic Studies* 65(3): 595–621.
- Bendor, Jonathan, and Piotr Swistak. 1997. “The Evolutionary Stability of Cooperation.” *American Political Science Review* 91(2): 290–307.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1998. “Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades.” *Journal of Economic Perspectives* 12(3): 151–70.
- Bowles, Samuel. 2004. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton, NJ: Princeton University Press.
- Burt, Ronald S. 2004. “Structural Holes and Good Ideas.” *American Journal of Sociology* 110(2): 349–99.

- Cohen, Michael D., Rick L. Riolo, and Robert Axelrod. 2001. "The Role of Social Structure in the Maintenance of Cooperative Regimes." *Rationality and Society* 13(1): 5–32.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(Suppl.): S95–S120.
- Eschel, Ilan, Dorothea K. Herreiner, Larry Samuelson, Emilia Sansone, and Avner Shaked. 2000. "Cooperation, Mimesis and Local Interaction." *Sociological Methods and Research* 28(3): 341–64.
- Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90: 715–35.
- Feiock, Richard C., and John T. Scholz, eds. 2009. *Self-Organizing Federalism: Collaborative Mechanisms to Mitigate Institutional Collective Action Dilemmas*. New York: Cambridge University Press.
- Fudenberg, Drew, and Eric Maskin. 1990. "Evolution and Cooperation in Noisy Repeated Games" *American Economic Review* 80: 274–79.
- Fudenberg, Drew, and David K. Levine. 1998. *The Theory of Learning in Games*. Cambridge, MA: MIT Press.
- Fudenberg, Drew, and Loren A. Imhof. 2006. "Imitation Processes with Small Mutations." *Journal of Economic Theory* 131: 251–62.
- Gott N. M., J. G. Polhill, and A. N. R. Law. 2003. "Agent-Based Simulation in the Study of Social Dilemmas." *Artificial Intelligence Review* 19: 3–92.
- Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6): 1360–80.
- Granovetter, Mark S. 1985. "Economic Action, Social Structure, and Embeddedness." *American Journal of Sociology* 91(3): 481–510.
- Hoffmann, R., and N. C. Waring. 1998. "Complexity Cost and Two Types of Noise in the Repeated Prisoner's Dilemma." In *Artificial Neural Nets and Genetic Algorithms*, eds. G. C. Smith, N. C. Steele, and R. F. Albrecht. Vienna: Springer-Verlag.
- Holland, John H. 1992. "Genetic Algorithms." *Scientific American* 267(1): 66–72.
- Imhof, Loren A., Drew Fudenberg, and Martin A. Nowak. 2005. "Evolutionary Cycles of Cooperation and Defection" *Proceedings of the National Academy of Sciences* 102(31): 10797–800.
- Kandori, Michihiro, George J. Mailath, and Rafael Rob. 1993. "Learning, Mutation, and Long Run Equilibria in Games." *Econometrica* 61(1): 29–56.
- Lazer, David, and Allan Friedman. 2007. "The Network Structure of Exploration and Exploitation." *Administrative Science Quarterly* 52(4): 667–94.
- Lubell, Mark, Mark Schneider, John T. Scholz, and Mihriye Mete. 2002. "Watershed Partnerships and Emergence of Collective Action Institutions." *American Journal of Political Science* 46(1): 148–63.
- Macy, Michael W., and John Skvoretz. 1998. "The Evolution of Trust and Cooperation between Strangers: A Computational Model." *American Sociological Review* 63(5): 638–60.
- March, James G. 1991. "Exploration and Exploitation in Organization Learning." *Organization Science* 2(1): 71–87.
- Mitchell, Melanie. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.
- Newman, M. E. 2003. "The Structure and Function of Complex Networks." Santa Fe Institute Working Paper (March 23).
- Nowak, Martin A. 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Cambridge, MA: Belknap Press.
- Nowak, Martin A., Akira Sasaki, Christine Taylor, and Drew Fudenberg. 2004. "Emergence of Cooperation and Evolutionary Stability in Finite Populations." *Nature* 428(April): 646–50.
- Ohtsuki, Hisashi, Christoph Hauert, Erez Lieberman, and Martin A. Nowak. 2006. "A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks." *Nature* 441: 502–05.
- Ohtsuki, Hisashi, and Martin A. Nowak. 2008. "Evolutionary stability on graphs." *Journal Of Theoretical Biology* 251(4): 698–707.
- Ohtsuki, Hisashi, Jorge M. Pacheco, and Martin A. Nowak. 2007. "Evolutionary graph theory: Breaking the symmetry between interaction and replacement." *Journal of Theoretical Biology* 246: 681–94.
- Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-First Century. (The 2006 Johan Skytte Prize Lecture)" *Scandinavian Political Studies* 30(2): 137–74.
- Sabatier, Paul, and Hank Jenkins-Smith, eds. 1993. *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder, CO: Westview Press.
- Schmidt, David, Robert Shupp, James Walker, T. K. Ahn, and Elinor Ostrom. 2001. "Dilemma Games: Game Parameters and Matching Protocols." *Journal of Economic Behavior & Organization* 46(4): 357–77.
- Schneider, Mark, John Scholz, Mark Lubell, Denisa Midruta, and Matthew Edwardsen. 2003. "Building Consensual Institutions: Networks and the National Estuary Program." *American Journal of Political Science* 47(1): 143–58.
- Scholz, John T., and Cheng-Lung Wang. 2006. "Cooptation or Transformation? Local Policy Networks and Federal Regulatory Enforcement." *American Journal of Political Science* 50(1): 81–97.
- Scholz, John T., and Bruce Stiftel, eds. 2005. *Adaptive Governance and Water Conflict. New Institutions for Collaborative Planning*. Washington, DC: Resources for the Future Press.
- Szabó, György, and Gábor Fáth. 2007. "Evolutionary Games on Graphs." *Physics Reports* 446: 97–216.
- Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Watts, Duncan J., and Steven H. Strogatz. 1998. "Collective Dynamics of Small-World Networks." *Nature* 393(June): 440–42.
- Watts, Duncan J. 1999. *The Small-Worlds: the Dynamics of Networks between Order and Randomness*. Princeton, NJ: Princeton University Press.
- Wu, Jianzhong, and Robert Axelrod. 1995. "How to Cope with Noise in the Iterated Prisoner's Dilemma." *Journal of Conflict Resolution* 39(1): 183–89.