International Studies Quarterly (2004) **48**, 455–470

Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations

STEPHEN J. MAJESKI

University of Washington

The question addressed in this analysis is whether endowing agents with various forms of asymmetric power makes cooperation more likely across a variety of structural settings of conflict and cooperation present in international relations. To address this question, an agent-based model incorporating asymmetric power among agents in a set of (2 × 2) games that represent different forms of conflict and cooperation prevalent in international relations (Chicken, Stag, Assurance, Deadlock, and Prisoner's Dilemma) is developed and analyzed via simulation. Simulation results indicate that the introduction of asymmetric power substantially increases the chances that both cooperative agents survive and cooperative worlds evolve. This is particularly the case when agents are endowed with the ability to selectively interact with other agents. Also, anticipated variations in outcomes across the game structures regarding the likelihood of cooperation are supported.

Whether and how cooperation evolves in social settings characterized by the presence of selfish agents engaged in repeated relations without central authority has been of considerable importance to scholars of international politics and of interest to scholars across all the social sciences as well as philosophy, biology, and computer science.[1] International relations scholars have been particularly interested in various features of nation-states, the relations among nation-states, and the structural environment in which nation-states are embedded that make cooperation either possible or more likely. Studying the evolution of cooperation in the context of the Repeated Prisoner's Dilemma (RPD) has proven to be quite fruitful for international relations scholars.[2]

Yet, the RPD framework is also restrictive in a variety of ways.[3] For instance, while the RPD captures one important type of relationship among nation-states in the international system, there are a number of other structural settings that

---

Author's note: An earlier version of this article was presented at the 96th Annual Meeting of the American Political Science Association in Washington, D.C., August 31–September 3, 2000. I would like to thank Scott Paige, David Sylvan, Lars-Erik Cederman, Maurits van der Veen, Phil Schodt, Paul Johnson, and the anonymous reviewers for helpful comments and criticisms.

[1] Based upon the seminal works of John Maynard Smith (1982), Michael Taylor (1976), and Robert Axelrod (1984), the Repeated Prisoner's Dilemma (RPD) has become the central metaphor for the evolution of cooperation in populations of selfish agents without central authority.

[2] See Axelrod (1980a, 1980b), Axelrod and Keohane (1985), Dacey and Pendegraft (1988), Bendor, Kramer, and Stout(1991), Bendor (1993), Busch and Reinhardt (1993), Signorino (1996), and Majeski et al. (1999).

[3] Indeed, Axelrod and Keohane (1986:231) argued that the games of Stag, Chicken, and Deadlock predominate in international relations.

Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, USA, and 9600 Garsington Road, Oxford OX4 2DQ, UK.

scholars have found of relevance to understanding conflict and cooperation in international relations.[4] Also, agents in the RPD context typically have been treated as essentially undifferentiated units where only the type of plan or strategy they employ dictating how they interact with other agents varies.[5] Nation-states, of course, are differentiated on many dimensions and it is not surprising that differentiation among agents in terms of power and capability has been of particular interest to international relations scholars because such asymmetries have been an enduring feature of international systems.[6]

The question addressed in this analysis is whether endowing agents (nation-states) with various forms of asymmetric power makes cooperation more likely across a variety of conflict and cooperation structural setting types present in international relations. To answer this question, an agent-based model incorporating asymmetric power among agents in a set of (2 × 2) games that represent different forms of conflict and cooperation prevalent in international relations (Chicken, Stag, Assurance, Deadlock, and Prisoner's Dilemma) is developed and analyzed via simulation.[7] Before elaborating the particular forms of asymmetric power considered, the different game structures to be examined, and the simulation analyses of these agent-based models, the basic structure of the agent-based model is introduced.

Structure of the Agent-Based Model

All agent-based models have two components: agents, and an environment or world the agents inhabit. In abstract terms, agents are bundles of two types of rules: those that define various internal states of the agent and those that dictate how the agent responds to various stimuli from other agents and the environment (Holland, 1995). The basic structure of the agent-based model developed here is a repeated (2 × 2) game.[8] Agents have two choices: cooperate (C) and defect (D). When an agent interacts with another agent, there are four possible outcomes: both cooperate (CC), both defect (DD), one agent cooperates and the other defects (CD), or one agent defects while the other cooperates (DC).

Each agent is represented by a strategy specifying how the agent behaves when it interacts with other agents. Agent strategies are restricted to those employing the previous interaction with other agent(s) to determine current choices.[9] Strategies

---

[4] See Snyder and Diesing (1977) for a discussion of PD, Chicken, and Deadlock games and their application to international conflicts; Jervis (1978) for a discussion of Stag, PD, and Chicken and their application to international conflicts; Aggarwal (1996) for a discussion of PD and Chicken and their application to international debt rescheduling; Taylor (1987) for a discussion of Chicken and Assurance games and their application to public goods and collective action; Evangelista (1990) for a discussion of PD and arms control; Martin (1992) for a discussion of PD as coadjustment games in multilateral economic sanctions; Zagare (1990) for a discussion of Chicken games and deterrence; and Stein (1990) for the application of PD, Chicken, and Deadlock to international conflict and Stag (he labels the game Assurance) to international collaboration.

[5] There are, of course, a number of agent-based models where agents differ on a variety of features (see Epstein and Axtell, 1996; Axelrod, 1997b; Lustick, 2000) but these models are typically not based upon the RPD.

[6] A number of simulation and agent-based models have been developed, in part to test realist theory, that have focused on power and differences or asymmetries in power among agents (Bremer and Mihalka, 1977; Cusack and Stoll, 1990; and Cederman, 1997).

[7] The agent-based model introduced here was developed and analyzed via simulation by Majeski et al. (1999). In that analysis only the RPD was examined and agents were undifferentiated. In addition, agent mobility was a key feature of the model. Since nation-states do not move, that feature has been removed from the model for this analysis.

[8] As with most agent-based modeling approaches (Axelrod, 1997a), we begin with an explicit set of assumptions about some phenomena and use them to generate simulated data. We search for patterns in the simulated data, particularly the large-scale effects from the interactions of locally interacting agents or what are often referred to as "emergent properties" of the system (see Epstein and Axtell, 1996).

[9] Agent strategies are based upon only the previous interaction. However, some agents will be endowed with the ability to choose whether or not to interact with other agents. This decision (not the actual probabilities of the strategy) is based on interactions with the other agent that can go back as far as six prior interactions.

are probabilistic, defined by the conditional probabilities to cooperate (p_1, p_2, p_3, p_4) given that the outcome of the previous interaction was (CC, CD, DC, or DD), respectively. In addition, a strategy stipulates whether the agent cooperates or defects when it interacts for the first time with another agent. For example, an agent employing the familiar "tit for tat" (TFT) strategy cooperates the first time it interacts with another agent. However, it cooperates with 100% probability the next time only if the other agent cooperated the last time (following a CC or DC outcome). It will defect with 100% probability the next time if the other agent defected during the prior interaction (following a CD or DD outcome). Thus, the conditional probabilities to cooperate for the TFT strategy are [1.0, 0.0, 1.0, 0.0].

Several features are added to the basic repeated (2 × 2) game to produce the agent-based model. First, in most social contexts agents are located at or occupy some place or position in their world at any given moment in time; as a result, most interactions among social units are dictated by spatial proximity.[10] Therefore, an explicit spatial dimension is introduced by constructing a toroidal world or environment (a 20 × 20 grid of cells) consisting initially of sixty agents randomly assigned locations on the grid.[11] Each cell contains at most one agent. For each round of the simulation, agents interact with all agents who occupy the four non-diagonal cells that immediately surround the agent, a von Neumann neighborhood.

Second, most social agents—be they firms, tribes, individuals, or families—can and do move. Agents typically move when they find themselves in an unprofitable and undesirable situation or location. While there is usually a cost to the agent associated with moving, the benefits from freeing itself from the negative consequences of a particular location can be sufficient to warrant relocation. The problem is that nation-states in the modern international system do not move. States may expand, collapse, or disappear; however, the state, in some sense defined as the control of a particular physical territory, does not move. Therefore, in this analysis agents cannot move. Once they are located on a cell on the spatial grid, they remain there until they run out of energy or die of old age and at that point are removed from the grid.[12]

Third, all social agents consume various resources to sustain themselves and all ecologies (environments) can support a finite number of agents. As more agents compete, the economic and environmental costs of available resources increase. Therefore, an environmental carrying capacity is incorporated into the agent-based model by introducing a cost of survival for agents. As the artificial world or environment becomes more populated, the cost of living and ultimately surviving for agents increases.[13] The number of agents is restricted to a fixed range by applying a cost of surviving to every agent, and each iteration is dependent on population size. The formula for the cost of surviving (α) is

$$\alpha = k + 4 * (DC + CC) * N/(X * Y)$$

where k is a constant, (DC) and (CC) are the RPD payoffs, N is the number of agents in the world, and X is the width and Y is the height of the world grid. The cost of surviving indirectly allows the simulation to select the percentage of the population with the highest energy levels for reproduction and the lowest percentage for elimination. The change in energy for an agent at each iteration

---

[10] See Nowak, Latane, and Lewenstein (1994) for an argument about the importance of considering social dilemmas as occurring in social space and a discussion of various geometries to represent social space.

[11] Dugatkin and Wilson (1991), Nowak and May (1992), Oliphant (1994), Lindgren (1991), Lindgren and Nordahl (1995), and Lomberg (1996) have introduced a spatial component into the RPD or PD environment.

[12] This means that an agent with no neighbors does not interact. It cannot move, so interaction can only occur if a new agent (via replication) is located in its neighborhood.

[13] This perspective on the relevance of an environmental carrying capacity is based upon the work of Hardin (1977, 1991) and Clayton and Radcliffe (1996) who argue that human populations do have cultural carrying capacities. Note as well that the cost-of-survival mechanism is similar to Epstein's (1998) global metabolic rate.

(ΔE), where ΔE equals the energy level of the agent at iteration (t) minus the energy level at iteration (t − 1), is the sum of all interactions minus the cost of surviving

$$\Delta E = \sum_{i=1}^{4} A_i - \alpha$$

where A_i is the payoff from the interaction in the i^{th} direction, and (α) is the cost of surviving. An agent is eliminated from the simulation when its energy falls below zero.

Fourth, because all individuals die and all social units, including nation-states, eventually fall apart, disband, go bankrupt, are taken over or are overrun, agents are assumed to have a limited existence or life span. An agent has a probability (Γ) of being eliminated for all iterations of the repeated game,

$$\Gamma = (A - T)/M$$

where A is the age of the agent, T is a constant for the minimum life span, and M is a constant where T + M is the maximum life span.[14] Once an agent reaches the minimum life span, then it has a nonzero and increasing probability of elimination until it reaches the maximum life span, and then it is eliminated with a probability of 1.0.[15] Of course, agents can be eliminated at any time if their energy level falls below zero.

Fifth, agents reproduce in the sense that they create a replication of themselves. Since it is the nation-state and not the individual that is modeled, it makes sense to think of reproduction as asexual. The "genetic material" (here the strategy of the agent) comes from one "parent" and there is no "crossover" of genetic material. The replicated agent ("child") has the same strategy as its "parent" unless mutation (discussed next) occurs. Replication requires a certain level of energy (maturity, size, power, wealth) and it costs the agent a significant amount of energy. Agents must reach a fixed level of energy (ρ) before they can replicate. Once an agent replicates, the energy of the agent and the replicated agent are both set to (ρ/2). In the simulations analyzed here, (ρ) is set to 1000 energy units. Following replication the "parent" has 500 energy units and the "child" has 500 energy units. The replicated agents are placed in a randomly selected open cell on the grid. This approach to setting reproduction thresholds and determining the relative fitness of the members of the population has the advantage of performing the reproduction/elimination calculation at every iteration instead of making periodic sweeps through the population and more gracefully modifies the population of the simulation.

Sixth, a mechanism to vary agents' strategies and introduce new strategies must be incorporated to give the agent-based model a dynamic component. Without the introduction of such a mechanism, the model would be static and devoid of change. In this analysis an evolutionary approach is taken and the introduction of new strategies is introduced generationally via strategy mutation. Modelski (1996) suggests that an evolutionary approach to global politics is useful when the focus of understanding, as it is in this analysis, concerns international institutions (the emergence of networks of cooperative agents in this analysis) and their transitions,

---

[14] In the simulation results reported later, the minimum life span of an agent (T) is 2048 iterations and the maximum life span (T + M) is 6144 iterations.

[15] In an earlier analysis (Majeski et al., 1999), simulations were run with varying life spans of agents (longer and shorter). While minor variations in the timing and likelihood of cooperation emerging occur, the model is quite robust to these minor variations. However, if agents do not die of old age eventually (they can and do still die because their energy level goes to zero), the simulation results are quite different. Cooperative outcomes are far less likely to occur, and change in the mixtures of types of agents occurs far more slowly. Introducing a lifespan to agents gives the simulations more dynamics—changes are more likely to occur and to happen more quickly.

where the perspective is long-term and where choice processes are a function of trial-and-error search and selection.[16]

There are a number of other ways to vary an agent's strategy or introduce new strategies. Agent strategies can change via imitation, learning, or innovation. While a case can certainly be made that nation-states do attempt to adapt strategies via imitation and/or learning, various features of international politics make change in strategies in this fashion problematic. Imitation is difficult for agents because it is often not in the interest of agents to reveal their strategies to other agents. Agents may observe the behavior of other successful agents but not be able to induce and thus replicate the strategies that the successful agents employ. Learning is often difficult because the international environment is usually noisy (agents sometimes incorrectly implement their strategy choice and agents may not know or may incorrectly interpret the moves of other agents) making it hard for agents to correctly interpret the actions of other agents and thus the strategies employed by those agents that generate those actions.[17]

The introduction of new strategies occurs when agents replicate. There is a fixed (20%) chance that a strategy will mutate during the replication process. When a mutation occurs the agent's strategy is modified by changing each of the p_i in the strategy of the parent by $[-\delta, \delta]$ where $0<\delta<1$. Specifically, if δ is set to (.1), as is the case for the simulation results reported later, then the actual value to change the p_i is randomly selected from a uniform distribution over the interval $(-.1)$ to (.1). For example, an agent with the strategy [.5, .5, .5, .5] reaches the 1000 energy units required to replicate. When it does so, there is a 20% chance that the strategy of the replicated agent will be mutated. Suppose that this occurs. Next, a value δ is selected randomly from the uniform distribution over the interval $(-.1)$ to (.1) for each of the four p_i. Suppose that the values selected are $-.08$, .04, .06, and .03. Then the strategy for the new replicated agent is [.42, .54, .56, .53].

Asymmetric Power

Two forms of asymmetric power are introduced into the agent-based models. First, asymmetric power is introduced by differentially rewarding agents for the joint outcomes of the various games. Those agents having asymmetric power receive uniformly higher payoffs across all joint outcomes of the relevant game than the remaining agents. For example, for the Prisoner's Dilemma (PD) game the payoffs for the four outcomes (CC, CD, DC, DD) are $[1, -3, 3, 1]$, respectively, for those agents without asymmetric power and $[2, -2, 4, 2]$ for those having asymmetric power.

Second, asymmetric power is introduced by giving some agents the ability to selectively interact with other agents while making interaction mandatory for all other agents.[18] Selective interaction can help all agents, whether they want to cooperate or not, because they benefit when they interact with other agents who cooperate and suffer when they interact with agents that defect. Selective interaction is introduced into the agent-based model in the following fashion. When an agent interacts with another agent, it develops a history of play with that specific agent. The agent keeps track of how many times the other agent(s) defects. If the other agent defects (n) times in (m) prior interactions, where $0<n<m$, then

---

[16] For the use of an evolutionary approach to explain aspects of international relations see Florini (1996) and Gilpin (1996).

[17] These features of the international system also make changes in strategy due to innovation difficult. Also, when innovation is characterized as a process whereby a very small percentage of agents randomly vary their current strategy (e.g., Lomberg, 1996) as in a trial-and-error search mechanism, then it has similar properties to generational change via mutation.

[18] Selective or non-compulsory play has been shown to increase the likelihood of cooperation (Kitcher, 1993; Orbell and Dawes, 1993; Batali and Kitcher, 1994; Stanley, Ashlock, and Tesfatsion, 1994).

TABLE 1. Properties of the (2 × 2) Games

	Preference Ordering	*Nash Equilibria*	*Key Difference actual payoffs used*
Prisoners Dilemma	DC > CC > DD > CD CC = l, CD = −3, DC = 3, DD = −1	DD	But CC jointly preferred to DD
Chicken	DC > CC > CD > DD CC = 1, CD = −1, DC = 3, DD = −3	DC and CD	Equilibria are unstable
Slag	CC > DC > DD > CD CC = l, CD = −3, DC = 0, DD = −1	CC and DD	CC jointly preferred but DC > DD
Assurance	CC > DD > DC > CD CC = l, CD = −3, DC = −2, DD = 0	CC and DD	CC jointly preferred but DD > DC
Deadlock	DC > DD > CC > CD CC = −l, CD = −3, DC = 3, DD = 1	DD	DD jointly preferred to CC

the agent will not interact with that agent again with the following proviso:[19] an agent can "choose" not to interact with the other agent only if it is more powerful than the other agent. This constraint on the ability to "choose" to selectively interact is designed to capture the notion that the ability to have the choice to interact or not is often not universally shared by agents in many social contexts. In many contexts in international relations, the choice to interact or not is often dictated by an agent's relative capability. Weak states lack choices that more powerful states have. In the agent-based model the power to choose to interact is measured by the agent's energy level, a surrogate for material wealth and/or power.

(2 × 2) Games Representing a Variety of Conflict and Cooperation Settings

Many of the different kinds of structural situations of conflict and cooperation that nation-states find themselves in can be captured by the following set of (2 × 2) games: PD, Chicken, Stag, Assurance, Deadlock. The preference orderings over the four outcomes of each game, the Nash equilibria for the five games, key differences in the preference orderings over outcomes, and actual payoffs used in the simulations analyses are presented in Table 1.

Agents face different incentives in each of the five games elaborated in the table. In some games the incentives that agents face make the decision to cooperate relatively easy and free from risk and punishment, while in other games choosing to cooperate requires assuming substantial risk and cost. The ability to achieve cooperation among agents varies across these games and the meanings attributed to cooperation across the games also vary. Cooperation, as it is used in this analysis, simply means that both agents choose to cooperate. That cooperation or cooperative outcomes are obtained in various simulations simply means that agents choose to cooperate at very high rates (over 95% of the time).

In the much-studied Prisoner's Dilemma game, agents find themselves in a situation where they can both benefit by cooperating. But when agents cannot form binding agreements and can only engage in "cheap" talk, they are tempted to exploit the other agent and both fear that they will be exploited.[20] While mutual

---

[19] In the simulations results reported, m = 6 and n = 3. Other values of m and n were examined as well with only minor alterations in the results.

[20] Cheap talk is a form of communication that economists and others define as being costless, non-binding, and having no bearing on the agent's payoffs (Schelling, 1960; Kreps, 1990; Crawford, 1990; and Johnson, 1993).

cooperation (CC) is a preferred and more readily achieved outcome in a RPD setting, it remains difficult for agents to achieve, as there are considerable advantages to be gained from unilateral defection and disadvantages to cooperating when the other agent defects. The PD is a situation where agents' decisions are motivated by fear, greed, and a lack of trust, motives commonly attributed to groups acting on behalf of nation-states by scholars of international relations. RPD games have been applied to such phenomena as arms races and arms control arrangements.

Chicken games capture high-risk situations. Agents want to "win" by forcing the opponent to back down, give in, or acquiesce. The problem is that while both agents want to win (DC), they both want to avoid the very high costs should they both fail to cooperate (DD) and the humiliating result of unilateral cooperation (CD). As a result, agents typically engage in threats and efforts to make credible commitments in an effort to "win." But, given the structure of the game, agents that can "collude" and agree to mutually cooperate (CC) create a better option for single and repeated play situations. In repeated play the best option is for agents to "collude" and agree to trade off (CD) and (DC) outcomes. International crises are often characterized as Chicken games, though they are typically single shot affairs. Bargaining about trade agreements with threats of sanctions and the possibility of trade wars are often characterized as repeated play games of Chicken.

The Stag Hunt game captures another type of mixed motive setting. Stag games are similar to PD games but differ in one crucial respect. In Stag games agents prefer the mutual cooperation outcome (CC) to the unilateral defection (DC) outcome. Groups of states that impose economic embargoes or sanctions find themselves in something like a Stag situation. If all adhere to the sanctions, then the chances of making the state or states facing the sanctions act in prescribed ways are high. But some may find it in their interest to break the economic embargo and trade with the embargoed state while all others do not. Nonetheless, unlike in PD, mutual cooperation is a Nash equilibrium for the Stag game. This suggests that agents should be more likely to cooperate in Stag games than in PD games. The problem is that if an agent does not think that the other agent will cooperate, then the agent should defect. Thus, Stag is a game of trust and fear, though greed has seemingly been removed as a motive to defect. Jervis (1978) demonstrated that Stag games capture a number of international security dilemmas and Stein (1990) suggested that the international extradition of criminals that started in the early 19th century is an example of a Stag game (what he labeled Assurance).

Agents face a different problem in the Assurance game. Franzen (1995) used the example of a relay race to capture the essence of Assurance games. Each member of the team must do her part for the team to reach the finish line. However, if one runner shirks rather than giving her maximum effort, then no one has a chance of winning the race. Since the runners are assumed to want to win the race, then cooperation is in the self-interest of all players, cooperation is a dominant strategy, and mutual cooperation is a Nash equilibrium. But if one player expects any of the others to shirk and not cooperate, then defection becomes the best option because the goal of winning cannot be achieved and it is better to minimize one's own individual effort or contribution.[21] Whereas in a game of Stag the collective good can be forsaken for a smaller personal payoff (cheating on a trade embargo), in the Assurance game no other positive payoff besides the collective good is possible. The best outcome for each player is for all involved to cooperate and reap the rewards of the mutually won prize. However, unlike in the Stag game, the second most preferred outcome for an agent in an Assurance game is for all to defect because no

---

[21] The Assurance game depicted here is based on the game elaborated by Taylor (1987:18–19, 38–39) and is consistent with its use by Franzen (1995). The structure of the Stag game presented here is consistent with what Lichbach (1996:47) and Stein (1990) call Assurance games.

agent will incur the costs (effort) of cooperating and thus none will have a reason for anger with shirkers. Assurance games are, like Stag and PD, based on trust and fear, but the incentive to unilaterally defect is considerably less than in either Stag or PD and so mutual cooperation is more likely to be achieved. International agreements such as the Convention on Rights of the Child or treaties such as the Ottawa Treaty to Ban Landmines, the Kyoto Accords on Global Warming, and the Nuclear Non-Proliferation Treaty appear to be structured like Assurance games.[22]

Deadlock games capture structural settings where agents are simply not prepared to compromise to achieve their desired outcome. Agents would rather fight than compromise or "lose." In Deadlock an agent most prefers the outcome in which it stands firm while the opponent capitulates, the (DC) outcome. Also, agents would rather have overt conflict—the mutual defection (DD) outcome—in defense of these aims than give in to the other's demands. What makes Deadlock situations so intractable is that both agents prefer the mutual defection outcome to the mutual cooperation outcome—each agent believes that the situation will be resolved only by overt conflict or capitulation by the other agent. Uncertainty in Deadlock situations arises when one agent perceives that the other's actions are contingent upon its own. The agent may believe that the other will give in if pushed long or hard enough. Before the outbreak of the Gulf War in 1990, the United States was bent on protecting its oil supplies, forcing Iraq from Kuwait, and maintaining the balance of power in the Middle East. On the other hand, Iraq was apparently determined to maintain its occupation of Kuwait even in the face of what seemed to be a likely military defeat. Both sides preferred to "win," but both sides also preferred armed conflict to negotiation or settlement.

Given the differences in preference orderings over outcomes and variations in Nash equilibria of the five (2 × 2) games in Table 1, the following are the anticipated variations in the likelihood that stable cooperation—defined as a situation when a high level of cooperation (the average cooperation rate among agents is over 95%) is achieved and maintained.[23] First, it is expected that stable cooperative outcomes are most likely to evolve in Stag and Assurance games because mutual cooperation is the most preferred outcome and is a Nash equilibrium for each game. In Chicken and Prisoner's Dilemma games mutual cooperation is not a Nash equilibrium and in both games is the second most preferred outcome. Nonetheless, stable cooperative outcomes are expected to evolve frequently in these two games but at a lower rate than in Stag and Assurance games.

Second, cooperative outcomes in PD and Chicken games are more likely to be less stable and collapse than in Stag and Assurance games because there are greater advantages to exploitation (DC is the most preferred outcome) in these games.[24] Third, in repeated Chicken games where payoffs from (DC + CD) are greater than (2 × CC), a switching equilibrium—where agents swap (DC) and (CD) payoffs and where cooperation rates are approximately .5—is likely to occur. Finally, given that mutual defection is the second most preferred outcome and is a Nash equilibrium, non-cooperative outcomes should be the norm for Deadlock games.

---

[22] The assumption is that benefits are achieved when all cooperate and that none see unilateral defection as a positive outcome. However, in each case, a few nation-states (some quite important) appear to have preference orderings inconsistent with those of Assurance games.

[23] As with most collective action type situations, as the number of agents increases the likelihood of free riding increases. In addition, the probability that some agent may have a different preference ordering across outcomes also increases. Also, when the number of agents involved is quite large, if a very small number choose to defect the rest may still prefer mutual cooperation. In the bilateral (dyadic) setting of the agent-based model, free riding is not a problem, but any defection leads to the breakdown of mutual cooperation.

[24] There is also a tendency for agent strategies to evolve toward strategies that are susceptible to exploitation over time in these cooperative eras.

Simulation Analyses

Assessing Levels of Cooperation Across the Various (2 × 2) Games

To determine whether expected variations in cooperation for the five (2 × 2) game structures noted earlier occur, thirty simulations of the agent-based model were run for each of the five game structures.[25] Outcomes of the simulation runs fall into one of four possible categories: stable cooperation (SC), no cooperation (NC), punctuated equilibrium (PE), and stable cooperation at the .5 level. As noted earlier, a stable cooperative outcome occurs when a high level of cooperation (the average cooperation rate among agents is over 95%) is achieved at some point in the simulation run and is maintained until the end of the simulation. A non-cooperative outcome occurs when the average cooperation rate quickly declines and stays at less than 5% for the entire simulation run. A simulation run is categorized as an instance of punctuated equilibrium if, after stable cooperation is achieved, it is followed by periodic (one or more) massive dips to near universal defection, a pattern similar to that found by Nowak and Sigmund (1993). A simulation run is categorized as a case of stable cooperation at the .5 level when the average cooperation among agents reaches 50% and is maintained for the duration of the simulation run.

The simulation results reported in Table 2, not surprisingly, generally conform to expectations.[26] First, stable cooperation fails to evolve in Deadlock. Second, stable cooperation is significantly more likely to evolve in Assurance and Stag games than in PD or Chicken games. Third, while high levels of cooperation are achieved with considerable frequency in PD games (combining the stable cooperation and punctuated equilibrium categories) and Chicken games (combining the stable cooperation, punctuated equilibrium and cooperation at the .5 level categories), they are far more fragile than in Stag and Assurance games and usually collapse. Lastly, the switching (DC and CD) .5 cooperation level equilibrium does occur with considerable frequency (53%) in Chicken games.

Most scholars of international relations would not be surprised by these long-term simulation results. In Stag and Assurance games, where the incentives to cooperate are high and the incentives to defect are relatively low, it is not surprising that networks of cooperative agents emerge about 90% of the time. However, it is a bit surprising that these cooperative "regimes" are so stable since agents have some incentives to defect. Once cooperation is achieved it does not collapse, a result that most scholars would be skeptical of even in settings where incentives to cooperate are very high. In more high-risk contexts such as PD and Chicken, where the incentives to defect are high, cooperation is still achieved at fairly high rates for

---

[25] As noted in Table 1, the actual payoffs employed in the simulations for the four outcomes (CC, CD, DC, DD) are for PD [1, −3,3, −1], Chicken [1, −1,3, −3], Assurance [1, −3, −2, 0], Stag [1, −3, 0, −1], and Deadlock [–1, −3, 3, 1]. These payoffs conform to the preference orderings over the outcomes for the various games. The specific payoff values were selected to make the payoffs across the five games as comparable as possible. In addition, the range of the payoffs across the games must be consistent so that the function that affects the environmental carrying capacity and the cost of survival operates consistently across the five games. Each simulation was run for two million iterations, a period of time sufficiently long for the simulations to stabilize and either reach a highly cooperative state or remain in a conflictual state and also to observe the collapse of high levels of cooperation. The initial mix of agent strategies is a set of identical strategies, all with the following set of conditional probabilities [.5, .5, .5, .5]. Also, half the agents cooperate the first time they interact with another agent and the other half defect the first time they interact with another agent. The purpose of choosing this "random strategy mix" is to avoid an initial selection of strategy mixes that is biased toward generating either a cooperative or non-cooperative world. Following Nowak and Sigmund (1993), the combination of seeding the initial strategy pool with one strategy and the application of a genetic algorithm allow for complex emergent behavior.

[26] The agent-based model described earlier has a number of parameters that must be fixed at some value. A large number of additional simulations were run to assess the robustness of the simulation results to variations in the values of payoffs, mutation rates, mutation magnitudes, reproduction thresholds, and lifespan lengths. The sensitivity analysis indicated that the general results are robust to variations in these important parameters.

TABLE 2. Long Term Simulation Results

	PD	*Chicken*	*Assurance*	*Stag*	*Deadlock*
Stable Cooperation	30%	7%	90%	90%	0%
No Cooperation	13%	10%	7%	10%	100%
Punctuated Equilibrium	57%	30%	0%	0%	0%
Cooperation at .5	0%	53%	3%	0%	0%
N =	30	30	30	30	30

both PD (87%) and Chicken (90%). However, it is very fragile and collapses 57% of the time in PD settings and 83% of the time in Chicken settings.[27]

Two trends observed in an earlier RPD game analysis (Majeski et al., 1999) continue for the Stag, Chicken, and Assurance games. First, when there is a transition from a non-cooperative to a highly cooperative world, a version of the Grim strategy (a pure form of the Grim strategy is [1.0, 0.0, 0.0, 0.0]) dominates numerically and appears to be essential for this transition.[28] Second, the transition from non-cooperation to cooperation is characterized by the formation of large clusters of Grim-like agents.[29]

Assessing the Effects of the Introduction of Asymmetric Power

There are a number of ways to assess the effects of the introduction of various forms of asymmetric power. The approach taken here is to assess whether agents having asymmetric power are more able to generate and maintain stable cooperation than those that do not. To implement this approach, a design similar to that used by Axelrod (1984) is employed.[30] Groups of agents with various types of cooperative strategies and various forms of asymmetric power are comparatively assessed to see whether they can "invade" a set of agents that are employing an exploitive strategy. A small number of cooperative agents are said to successfully invade a larger number of exploitive agents if they can survive, replicate, and drive the exploitive agents to extinction.

In the simulations, fifty exploitive All-Defect (All-D) agents and ten cooperative agents employing three types of cooperative strategies—TFT [1, 0, 1, 0], Grim [1, 0, 0, 0], and All-Cooperate (All-C) [1, 1, 1, 1]—are randomly distributed on the grid of the agent-based model.[31] An agent employing the All-D [0, 0, 0, 0] strategy

---

[27] The .5 switching equilibrium for the Chicken game can be seen as a pareto preferred equilibrium to the mutual cooperation equilibrium since long-term payoffs are higher to agents. Seen in this light, some form of stable cooperation is achieved in 60% of the Chicken game simulation runs.

[28] When there is a transition from a non-cooperative to a highly cooperative world, a version of the Grim strategy dominates numerically and appears to be essential for this transition. The emergence of small, relatively stable nodes or networks of agents characterizes every transition from uncooperative to cooperative worlds. In a study involving iterated Prisoner's Dilemma games among groups of human subjects, Majeski (2002) found that groups (n = 37) over the course of ten iteration games where the ending was unknown, appeared, on average, to play a noisy version of the Grim strategy. The average probability (across all groups and all iterations of the games) that groups cooperate given the four possible joint prior outcomes can be represented as an average group strategy which turns out to be [.83, .28, .28, .34]. This strategy mix is quite similar to the Grim-like strategies that come to dominate numerically and lead to cooperation in the agent-based models developed here. The lone exception to this pattern occurs in Assurance games. Almost all the time cooperation is Grim led. However, in two instances cooperation is generated by the clustering of Pavlov-like [.8, .2, .2, .8] agents.

[29] See Majeski and Sylvan (2000) for a detailed discussion of the mechanisms at work in this transition from non-cooperation to cooperation.

[30] Using a model of territoriality, Axelrod (1984) showed that when agents are restricted to interacting with their four neighbors and where strategies are "spread" by imitation or "colonization," cooperative TFT strategies could successfully invade a world of All-D agents as long as TFT strategies are introduced in a spatial cluster and not one at a time.

[31] The set of fifty All-D agents and ten agents with some cooperative strategy mix (Grim, TFT, All-C) is used because it creates differentiation. If a more cooperative mix (e.g., forty All-D agents and twenty cooperative strategy agents) is employed, cooperation almost always emerges under all types of strategies. The lone exception to this is All-C.

Table 3. Cooperative Agent Benchmark Simulation Results

	Overall		*Grim*		*TFT*		*All-C*	
Prisoners Dilemma	SC	67%	SC	100%	SC	100%	SC	0%
	NC	33%	NC	0%	NC	0%	NC	100%
Chicken	SC	33%	SC	100%	SC	0%	SC	0%
	PE	33%	PE	0%	PE	100%	PE	0%
	NC	33%	NC	0%	NC	0%	NC	100%
Assurance	SC	100%	SC	100%	SC	100%	SC	100%
	NC	0%	NC	0%	NC	0%	NC	0%
Stag	SC	80%	SC	100%	SC	80%	SC	60%
	NC	20%	NC	0%	NC	20%	NC	40%
Deadlock	SC	0%	SC	0%	SC	0%	SC	0%
	NC	100%	NC	100%	NC	100%	NC	100%
Total	SC	56%	SC	80%	SC	56%	SC	32%
	PE	7%	PE	0%	PE	20%	PE	0%
	NC	37%	NC	20%	NC	24%	NC	68%
	N = 75		N = 25		N = 25		N = 25	

defects whenever it plays another agent for the first time. If it interacted with an agent in the previous round, it always defects the next time regardless of the prior joint outcome. An agent employing the All-C [1, 1, 1, 1] strategy cooperates the first time it interacts with another agent and always cooperates the next time it interacts with that agent regardless of the prior joint outcome. As noted earlier, an agent employing the TFT [1, 0, 1, 0] strategy cooperates the first time it interacts with another agent. However, it cooperates the next time only if the other agent cooperated the last time [following a (CC) or (DC) outcome]. It will defect the next time if the other agent defected during the prior interaction [following a (CD) or (DD) outcome]. An agent employing the Grim [1, 0, 0, 0] strategy cooperates the first time it interacts with another agent. It cooperates for all subsequent interactions with that agent as long as the agent cooperates. Once the other agent defects just once, it will always defect whenever it interacts with that agent again.

To establish a benchmark and to determine whether the introduction of asymmetric power makes a difference in the ability of cooperative agents to invade a large group of exploitive agents, five simulations of each strategy mix (i.e., fifty All-D and ten TFT) were run for each of the five different game structures *without* the introduction of asymmetric power. Each simulation was run for 200,000 iterations and these results are reported in Table 3. The outcome of each simulation run is placed into one of three possible categories: SC, NC, PE. These outcomes are defined in the same fashion as earlier.

Overall, across the five games and three types of cooperative invaders (All-C, TFT, and Grim) there is a 63% chance that cooperative agents were able to successfully invade and essentially eliminate all exploitive All-D agents. Simulation outcomes categorized as either SC or PE indicate that high levels of cooperation were achieved and that cooperative agents successfully invaded. Cooperative agents successfully invaded 100% of the time in Assurance games, 80% in Stag games, 67% in both PD and Chicken, and 0% in Deadlock.[32] Given the payoff structure of Deadlock, it is hardly surprising that cooperative agents (Grim, TFT, and All-C) have no success in invading All-D and creating cooperation. Nor is it surprising that invasion is more successful when agents employ Grim and TFT strategies because, unlike the All-C strategy, these strategies punish defection.

---

[32] These results conform precisely to the expected ordering of the likelihood of cooperative outcomes across these five games.

TABLE 4. Asymmetric Power Results For All-Cooperate Strategy Cases

	Prisoners Dilemma		*Chicken*		*Stag*		*Total Sel Int and Payoff Diff*
	Selective Interaction	*Payoff Difference*	*Selective Interaction*	*Payoff Difference*	*Selective Interaction*	*Payoff Difference*	
Stable Cooperation	20%	0%	20%	0%	90%	60%	32% N = 19
Punctuated Equilibrium	50%	20%	80%	70%	0%	0%	36% N = 22
No Cooperation	30%	80%	0%	30%	10%	40%	32% N = 19
N =	10	10	10	10	10	10	N = 60

These baseline results establish where there is an opportunity to assess whether and to what extent endowing cooperative agents with asymmetric power increases the likelihood of successful invasion. This opportunity arises in all situations where cooperative agents without asymmetric power failed to invade 100% of the time. This occurs in eight situations: All-C agents in PD, Chicken, Stag, and Deadlock games; TFT and Grim agents in Deadlock games; and TFT agents in Stag and Chicken games.

It should come as no surprise that providing agents with asymmetric power is of absolutely no help in the Deadlock games. Even under the most favorable circumstances (i.e., fifty Grim agents and ten All-D agents with either version of asymmetric power) cooperative agents *cannot* successfully invade All-D agents and generate stable cooperation. Generating cooperation is simply hopeless for Deadlock games. Thus our focus shifts to the five remaining situations: All-C agents in PD, Chicken, and Stag games and TFT agents in Stag and Chicken games. Both forms of asymmetric power introduced earlier are examined in these five situations. A total of one hundred simulations were run for 200,000 iterations, twenty runs for each five game/cooperative agent combination (PD/All-C, Chicken/All-C, Stag/All-C, Chicken/TFT, and Stag/TFT). Of the twenty runs for each game structure/cooperative agent combination, ten were run where cooperative agents were endowed with selective interaction and ten where payoff differentials were provided to cooperative agents. The results are reported in Tables 4 and 5.

First, the simulation results reported in Table 4 indicate that providing All-C agents with asymmetric power significantly increases the likelihood that they can successfully invade (invasion combines both SC and PE outcomes) a world dominated by non-cooperative agents. For the three games PD, Chicken, and Stag, All-C agents without asymmetric power successfully invaded and generated cooperation 20% of the time (see Table 3) whereas All-C agents with asymmetric power (both types combined) successfully invaded and generated cooperation 68% of the time.

Providing TFT cooperative agents with asymmetric power also increases the likelihood that they can successfully invade a non-cooperative world of All-D agents. For the two games Chicken and Stag, TFT agents without asymmetric power successfully invaded and generated cooperation 90% of the time (see Table 3). As reported in Table 5, TFT agents with asymmetric power (both types combined) successfully invaded and generated cooperation 100% of the time. TFT agents endowed with either type of asymmetric power generated stable cooperation 100% of the time for Stag games (up from 80% without asymmetric power). For Chicken games, TFT agents without asymmetric power were able to successfully invade 100% of the time but in all instances cooperation collapsed (PE).

Second, there is a significant difference in the effectiveness of the two forms of asymmetric power. For All-C agents, the introduction of selective interaction is

TABLE 5. Asymmetric Power Results For TFT Strategy Cases

	Chicken		*Stag*		
	Selective Interaction	*Payoff Difference*	*Selective Interaction*	*Payoff Difference*	*Total Sel Int and Payoff Diff*
Stable Cooperation	100%	0%	100%	100%	75% N = 30
Punctuated Equilibrium	0%	100%	0%	0%	25% N = 10
No Cooperation	0%	0%	0%	0%	0% N = 0
N =	10	10	10	10	N = 40

substantially more effective than the introduction of payoff differentials, generating cooperation 87% of the time vs. 50%. For TFT cooperative agents, the introduction of selective interaction not only led to successful invasion 100% of the time but to stable cooperation in 100% of the cases. The introduction of payoff differentials led to successful invasion 100% of the time but failed to improve the likelihood that cooperation would be stable.

Why is selective interaction more helpful than favorable payoff differences for cooperative agents? In a world where there is a cost of survival, where agents survive by maintaining positive energy levels, and where they reproduce by growing to a relatively high level of energy (1000 energy units in the simulation), avoiding long strings of mutual defection outcomes via selective interaction is more evolutionarily helpful than a payoff advantage gained for each of the four payoff outcomes by the introduction of payoff differentials. Selective interaction helps, as we would expect, because it allows cooperative agents to refuse to interact with the agents that defect and seek to exploit them. The cooperative agents survive long enough so that they replicate themselves and the new agents can be located next to other cooperative agents and benefit from cooperation. Over time, cooperative agents out-replicate non-cooperative agents.

Summary

The question addressed in this analysis is whether endowing agents with various forms of asymmetric power makes cooperation more likely across a variety of structural settings of conflict and cooperation present in international relations. To answer this question, an agent-based model incorporating asymmetric power among agents in a set of (2 × 2) games that represent different forms of conflict and cooperation prevalent in international relations (Chicken, Stag, Assurance, Deadlock, and Prisoner's Dilemma) was developed and analyzed via simulation. The findings reported here indicating that cooperation is achieved with varying levels of difficulty across the five games' structures are consistent with game-theoretic predictions and the fact that cooperation among nation-states in international affairs does vary substantially across different substantive contexts.

The simulation results indicate that endowing cooperative agents with asymmetric power substantially increases the chances that those agents survive and that cooperative worlds evolve particularly when agents have the ability to selectively interact with other agents in their world. This holds for PD, Chicken, Stag, and Assurance games. Not surprisingly cooperative agents, whether they have asymmetric power or not, do not survive nor do cooperative worlds evolve in Deadlock games. The simulation results of the agent-based model are consistent with prior game-theoretic analyses regarding non-compulsory play and standard notions that nation-states with differential advantages in capabilities are likely to fare better in international affairs. Even when agents do not have the benefit of asymmetric power, it is highly likely that cooperation evolves and is maintained in the more "cooperative" Stag and Assurance game settings and frequent but difficult to maintain in the more "competitive" and "conflictual" PD and Chicken game contexts.

Providing cooperative agents with asymmetric power raises the likelihood that cooperation evolves in all four game structures and increases the likelihood, particularly in the more conflictual PD and Chicken game settings, that cooperation is maintained. Indeed, selective interaction even allows exploitable agents that do not punish defectors such as All-C to generate stable cooperation even in the relatively harsh environments of RPD and Chicken games. Exploitable agents that can selectively interact succeed in generating and maintaining cooperation nearly as often as cooperative agents that employ punishing (Grim and TFT) strategies without bearing the costs of "sanctioning" that those strategies often entail. That exploitable agents with the ability to selectively interact are quite successful even in conflictual settings such as Chicken and PD is a bit unexpected and it suggests that obtaining cooperation in some international contexts does not require uniform sanctioning regimes where all agents must employ punishing strategies.

Selective interaction helps cooperative agents establish networks of stable cooperative relations because they can avoid interacting with agents who seek to exploit them. Evidence from the simulation results suggests that selective interaction, by providing a form of non-compulsory play or "exit," provides a mechanism for cooperative agents to *survive* in an exploitive world and eventually to locate other cooperative agents and begin to prosper. It accounts for how cooperative agents establish a presence in a hostile world. But this is only part of the story. Cooperative agents come to *dominate* and produce stable cooperative worlds, even in risky environments such as PD and Chicken, because cooperative agents with the ability to selectively interact can effectively isolate exploitive agents and drive them into extinction. Cooperation emerges because the world gets divided into two "camps" of interacting agents. In one camp are cooperative agents generating wealth and power through cooperative interactions and in the other camp are exploitive agents who punish each other by inflicting high costs on each other which eventually leads to their extinction.[33]

A simple individual mechanism of selective interaction is sufficient to isolate and drive out non-cooperators. Non-cooperators do not need to be "tagged" or "branded."[34] Cooperators do not need to share information about non-cooperators with other cooperators nor do they need to "punish" or sanction exploiters by not cooperating. Cooperators can successfully invade and come to dominate worlds originally dominated by exploiters simply by interacting only with other cooperators.

References

AGGARWAL, V. K. (1996) *Debt Games: Strategic Interaction in International Debt Restructuring*. Cambridge: Cambridge University Press.

AXELROD, R. (1980a) Effective Choice in Prisoner's Dilemma. *Journal of Conflict Resolution* **24**:3–25.

---

[33] Cederman (2001) connects the clustering results obtained in his study and generated in other spatial agent-based modeling work with the finding that democratic security communities also tend to be clustered and geographically concentrated (Gaubatz, 1995; Gleditsch and Ward, 2000).

[34] Tags are actor-specific characteristics that are observable to other agents and are used to help other agents predict the behavior of agents. Tags have been shown to have a significant positive impact on cooperation (Holland, 1995; Epstein and Axtell, 1996). Cederman (2001) recently used tags to discern democratic states from non-democratic states in an agent-based model of the democratic peace.

AXELROD, R. (1980b) More Effective Choice in Prisoner's Dilemma. *Journal of Conflict Resolution* **24**:379–403.

AXELROD, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.

AXELROD, R. (1997a) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.

AXELROD, R. (1997b) The Dissemination of Culture: A Model of Local Convergence and Global Polarization. *Journal of Conflict Resolution* **41**:203–226.

AXELROD, R., AND R. O. KEOHANE (1985) Achieving Cooperation Under Anarchy: Strategies and Institutions. *World Politics* **38**:226–254.

AXELROD, R., AND R. O. KEOHANE (1986) "Achieving Cooperation Under Anarchy: Strategies and Institutions." In *Cooperation Under Anarchy*, edited by K. Oye, pp. 226–254. Princeton, NJ: Princeton University Press.

BATALI, J., AND P. KITCHER (1994) "Evolutionary Dynamics of Altruistic Behavior in Optional and Compulsory Versions of the Iterated Prisoner's Dilemma Game." In *Artificial Life IV: Proceedings of the Fourth International Workshop on Synthesis and Simulation of Living Systems*, edited by R. Brook and P. Maes, pp. 343–348. Cambridge, MA: MIT Press.

BENDOR, J. (1993) Uncertainty and the Evolution of Cooperation. *Journal of Conflict Resolution* **37**:709–734.

BENDOR, J., R. M. KRAMER, AND S. STOUT (1991) When in Doubt ...: Cooperation in a Noisy Prisoner's Dilemma. *Journal of Conflict Resolution* **35**:691–719.

BREMER, S., AND M. MIHALKA (1977) "Machiavelli in Machina: Or Politics Among Hexagons." In *Problems of World Modeling*, edited by K. W. Deutsch, pp. 303–338. Boston: Ballinger.

BUSCH, M. L., AND E. R. REINHARDT (1993) Nice Strategies in a World of Relative Gains: The Problem of Cooperation Under Anarchy. *Journal of Conflict Resolution* **37**:427–445.

CEDERMAN, L. (1997) *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton, NJ: Princeton University Press.

CEDERMAN, L. (2001) Modeling the Democratic Peace as a Kantian Selection Process. *Journal of Conflict Resolution* **45**:470–502.

CLAYTON, A., AND N. RADCLIFFE (1996) *Sustainability: A Systems Approach*. London: Earthscan Publications.

CRAWFORD, V. (1990) Explicit Communication and Bargaining Outcomes. *American Economic Review* **80**:213–219.

CUSACK, T., AND R. STOLL (1990) *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation*. Boulder, CO: Lynne Rienner.

DACEY, R., AND N. PENDEGRAFT (1988) The Optimality of Tit-For-Tat. *International Interactions* **15**: 45–64.

DUGATKIN, L. A., AND D. S. WILSON (1991) Rover: A Strategy for Exploiting Cooperators in a Patchy Environment. *American Naturalist* **138**:687–700.

EPSTEIN, J. M. (1998) Zones of Cooperation in Demographic Prisoner's Dilemma. *Complexity* **4**:33–48.

EPSTEIN, J. M., AND R. AXTELL (1996) *Growing Artificial Societies: Social Science from the Bottom Up*. Washington: Brookings Institution.

EVANGELISTA, M. (1990) Cooperation Theory and Disarmament Negotiations in the 1950s. *World Politics* **42**:502–528.

FLORINI, A. (1996) The Evolution of International Norms. *International Studies Quarterly* **40**: 363–389.

FRANZEN, A. (1995) Group Size and One-Shot Collective Action. *Rationality and Society* **7**:183–200.

GAUBATZ, K. (1995) "Still Hazy After All These Years: Kant's Secret Plan of Nature and the Expansion of Democratic States in the International System." Typescript, Stanford University.

GILPIN, R. (1996) Economic Evolution of National Systems. *International Studies Quarterly* **40**:411–431.

GLEDITSCH, K., AND M. D. WARD (2000) War and Peace in Space and Time: The Role of Democratization. *International Studies Quarterly* **44**:1–29.

HARDIN, G. (1977) "Ethical Implications of Carrying Capacity." In *Managing the Commons*, edited by G. Hardin and J. Baden, pp. 112–125. San Francisco: W. H. Freeman.

HARDIN, G. (1991) "Paramount Positions in Ecological Economics." In *Ecological Economics: The Science and Management of Sustainability*, edited by R. Costanza, pp. 47–57. New York: Columbia University Press.

HOLLAND, J. (1995) *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Addison-Wesley.

JERVIS, R. (1978) Cooperation Under the Security Dilemma. *World Politics* **30**:167–214.

JOHNSON, J. (1993) Is Talk Cheap? Promoting Conversation Between Critical Theory and Rational Choice. *American Political Science Review* **87**:74–86.

KITCHER, P. (1993) The Evolution of Human Altruism. *Journal of Philosophy*, (October): 497–516.

KREPS, D. (1990) *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.

LICHBACH, M. (1996) *The Cooperator's Dilemma*. Ann Arbor: University of Michigan Press.

LINDGREN, K. (1991) "Evolutionary Phenomena in Simple Dynamics." In *Artificial Life II*, edited by C. G. Langton, J. D. Farmer, S. Rasmussen and C. Taylor, pp. 295–312. Redwood City, CA: Addison-Wesley.

LINDGREN, K., AND M. NORDAHL (1995) "Cooperation and Community Structure in Artificial Ecosystems." In *Artificial Life: An Overview*, edited by C. G. Langton, pp. 15–37. Cambridge, MA: MIT Press.

LOMBERG, B. (1996) Nucleus and Shield: The Evolution of Social Structure in the Iterated Prisoner's Dilemma. *American Sociological Review* **61**:278–307.

LUSTICK, I. (2000) Agent-based Modelling of Collective Identity: Testing Constructivist Theory. *Journal of Artificial Societies and Social Simulation* **3** 1 web page ⟨http://jasss.soc.surrey.ac.uk/3/1/contents.html⟩.

MAJESKI, S. J. (2002) Small Group Dynamics and Foreign Policymaking: Evidence from Iterated Prisoners Dilemma Game Experiments. Typescript.

MAJESKI, S. J., G. LINDEN, C. LINDEN, AND A. SPITZER (1999) Agent Mobility and the Evolution of Cooperative Communities. *Complexity* **5**:16–24.

MAJESKI, S. J., AND D. J. SYLVAN (2000) "An Argument for Non-causal Necessity: Two Approaches to Modeling Essences." Typescript.

MARTIN, L. L. (1992) *Coercive Cooperation: Examining Multilateral Economic Sanctions*. Princeton, NJ: Princeton University Press.

MODELSKI, G. (1996) Evolutionary Paradigm for Global Politics. *International Studies Quarterly* **40**: 321–342.

NOWAK, A., B. LATANE, AND M. LEWENSTEIN (1994) "Social Dilemmas Exist in Space." In *Social Dilemmas and Cooperation*, edited by U. Schulz, W. Albers and U. Mueller, pp. 269–289. Berlin: Springer-Verlag.

NOWAK, M., AND R. MAY (1992) Evolutionary Games and Spatial Chaos. *Nature* **354**:826–829.

NOWAK, M., AND K. SIGMUND (1993) A Strategy of Win-Stay, Lose-Shift That Outperforms TFT in the Prisoner's Dilemma Game. *Nature* **364**:56–58.

OLIPHANT, M. (1994) "Evolving Cooperation in Non-Iterated Prisoner's Dilemma: The Importance of Spatial Organization." In *Artificial Life IV: Proceedings of the Fourth International Workshop on Synthesis and Simulation of Living Systems*, edited by R. Brook and P. Maes, pp. 349–352. Cambridge, MA: MIT Press.

ORBELL, J. M., AND R. M. DAWES (1993) Social Welfare, Cooperators Advantage, and the Option of Not Playing the Game. *American Sociological Review* **58**:787–800.

SCHELLING, T. A. (1960) *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

SIGNORINO, C. S. (1996) International Cooperation Under Uncertainty: The Effects of Symmetric and Asymmetric Noise. *Journal of Conflict Resolution* **40**:152–205.

SMITH, J. M. (1982) *Evolution and the Theory of Games*. Cambridge and New York: Cambridge University Press.

SNYDER, G. H., AND P. DIESING (1977) *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.

STANLEY, E. A., D. ASHLOCK, AND L. TESFATSION (1994) "Iterated Prisoner's Dilemma with Choice and Refusal of Partners." In *Artificial Life III: Proceedings Volume 17, SFI Studies in the Sciences of Complexity*, edited by C. Langton, pp. 131–175. Reading, MA: Addison-Wesley.

STEIN, A. (1990) *Why Nations Cooperate*. Ithaca, NY: Cornell University Press.

TAYLOR, M. (1976) *Anarchy and Cooperation*. London: Wiley.

TAYLOR, M. (1987) *The Possibility of Cooperation*. Cambridge and New York: Cambridge University Press.

ZAGARE, F. C. (1990) Rationality and Deterrence. *World Politics* **42**:238–260.