Do Exploitive Agents Benefit from Asymmetric Power in International Politics?

Stephen J. Majeski

Department of Political Science

University of Washington

Seattle, WA 98195

206-543-2783

(fax) 206-685-2146

majeski@u.washington.edu

Abstract

Endowing agents that prefer cooperative outcomes with asymmetric power substantially increase the chances that both cooperative agents survive and that cooperative worlds evolve across a variety of structural settings of conflict and cooperation present in international relations; particularly when agents are endowed with the ability to selectively interact with other agents. These results are consistent with the general finding that non-compulsory play consistently helps cooperators. The question addressed in this analysis is whether or not asymmetric power also helps exploitive agents in the same structural settings; a question heretofore not analyzed. Contrary to expectations, the simulation results reported here suggest that exploitive agents benefit from asymmetric power only in very restricted circumstances; ones relativity unlikely to occur in international relations. In effect there is an asymmetry in the benefits of asymmetric power.

Introduction

Whether and how cooperation emerges and can be maintained 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.[1] In particular, international relations scholars have been interested in features of agents (typically nation-states), the relations among agents, and the structural environment agents are embedded in that make cooperation either possible or more likely.[2] Nation-states, of course, are differentiated on many dimensions and it is not surprising that differentiation among agents in terms of power and capability, labeled here asymmetric power, has been of particular interest to international relations scholars because such asymmetries have been an enduring feature of international systems.[3]

In an earlier analysis Majeski[4] constructed a set of agent-based models[5] based upon a group of repeated 2X2 games (Prisoners Dilemma, Chicken, Stag, Assurance, and Deadlock)[6] and examined the effects of the introduction of asymmetric power among agents on the emergence of cooperation.[7] The analysis demonstrated that the introduction of asymmetric power substantially increased the chances that cooperative agents survived and that cooperative worlds evolved; particularly when agents were endowed with the ability to selectively interact with other agents in their world. Selective interaction helps agents that want to cooperate more effectively establish and maintain what could be characterized as cooperative regimes where norms of niceness, retaliation, and limited forgiveness prevail.[8] For those cooperative regimes to be maintained agents must punish both free riders and exploiters.

The reason that selective interaction helps agents that prefer cooperative outcomes also appears to apply to agents that want to exploit other agents. [9] Exploitive agents also do well when they interact with other agents that cooperate and do poorly when they interact with other agents that defect. Exploitive agents that almost exclusively interact with agents that are exploitable (those who unilaterally cooperate and/or do not punish unilateral defection) ought to do well and should be more effective at exploiting and potentially destroying cooperative regimes than exploitive agents that must interact with cooperative agents that do retaliate and punish defection and with other exploitive agents.

Agents that seek to exploit other agents exist in most international contexts. Indeed, the fear that other agents will attempt to exploit them is a central concern for nation-states engaged in are wide variety of relations with other nation-states. For example the possibility of exploitation is central to such phenomena as arms races and arms control arrangements. Economic embargos and sanctions often fail because some states exploit the cooperative arrangement and secretly break the embargo agreement. International agreements such as the Ottawa Treaty to Ban Landmines, Kyoto Accords on Global Warming, and the Nuclear Non-proliferation Treaty are undermined because key states either cheat or never enter into the agreement, and bargaining about trade agreements with threats of sanctions and the possibility of trade wars is typically designed to constrain both potential and actual exploiters. Given, the prevalence of contexts in international politics where exploitation is likely and the fact that the international system is characterized by the presence of exploitive agents, it makes sense to determine whether or not exploitive agents also benefit from asymmetric power.[10]

In this analysis, the impact of providing exploitive agents with asymmetric power is examined across a variety of different settings of conflict and cooperation prevalent in international relations (Chicken, Stag, Assurance, Deadlock, and Prisoners Dilemma). A detailed discussion of the five 2X2 game structure settings, the agent-based model, and the two types of asymmetric power provided to exploitive agents can be found in Majeski.[11]

Assessing the Potential Benefits of Asymmetric Power for Exploitive Agents

Do exploitive agents also benefit from asymmetric power? The approach taken here is to assess whether agents having asymmetric power are more able to take over and dominate a world initially populated with cooperative agents than exploitive agents without asymmetric power. To implement this approach a design similar to that used by Axelrod[12] is employed. Groups of agents with various types of exploitive strategies and various forms of asymmetric power are comparatively assessed to see whether they can "invade" a set of agents employing cooperative strategies. A small number of exploitive agents are said to successfully invade a larger number of cooperative agents if they can survive, replicate, and drive the cooperative agents to extinction.

In the five 2X2 game structures, 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). An agent employing the Always Defect (All-D) strategy 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 Always Cooperate (All-C) 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. An agent employing the TFT strategy cooperates the first time in 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 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. In the simulations that follow, exploitive (All-D) agents and cooperative agents employing three types of cooperative strategies (in increasing order of exploitability)— Grim, TFT, and Always Cooperate (All-C)— are randomly distributed on the grid of the agent-based model.

To establish a benchmark and to determine whether the introduction of asymmetric power makes a difference in the ability of exploitive agents to invade a large group of cooperative agents, five simulations of each strategy mix (i.e., 50 Grim and 10 All-D) were run for each of the five different game structures **without** the introduction of asymmetric power.[13] Each simulation was run for 200,000 iterations and these results are reported in Table 1. The actual payoffs the four outcomes (CC, CD, DC, DD) employed in the simulations are for the PD game (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).[14] 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.[15]

The outcome of each simulation run is placed into one of three possible categories; stable cooperation (SC), non-cooperation (NC), and punctuated equilibrium (PE). 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 and is maintained until the end of the simulation run. A non-cooperative outcome occurs when the average cooperation rate declines and stays at less than 5% for the duration of the 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.[16]

The baseline results are reported in Table 1. For the most cooperative oriented Assurance and Stag games, stable cooperation always occurs. For the most conflictual Deadlock game, non-cooperation always occurs. For the PD and Chicken games, whether or not stable cooperation occurs depends upon the type of cooperative strategy employed. Stable cooperation occurs in PD only when the less exploitable Grim and TFT strategies are employed. Stable cooperation occurs in Chicken only when Grim is employed. TFT strategies generate the PE outcomes and All-Cooperate strategies generate non-cooperative outcomes.

The baseline results also establish where there is an opportunity to assess whether and to what extent endowing exploitive agents with asymmetric power increases the likelihood of successful invasion. Such an opportunity arises in any situation where exploitive agents **without** asymmetric power fail to invade 100% of the time. This occurs when the non-cooperative outcome (NC) is achieved less than 100% of the time.[17] An examination of Table 1 reveals that there is an opportunity for "improvement" for All-D exploitive agents in all the 2X2 games (PD, Chicken, Stag, and Assurance) except Deadlock. Specifically, exploitive agents can benefit from asymmetric power in the following game/cooperative agent strategy settings; Stag/All-Cooperate, Assurance/All-Cooperate and for all four games (PD, Chicken, Stag, and Assurance) where the cooperative agents "employ" TFT or Grim strategies; ten in total.

Table 1 about here

Ten simulations were run for each of these ten game/cooperative agent strategy combinations where the exploitive agents were endowed with the ability to selectively interact and another ten for each of the ten game/cooperative agent strategy combinations where the exploitive agent benefited from differential payoffs. In total 200 simulations were run for these twenty-game/cooperative agent strategy type/asymmetric type cases. The results are reported in the third columns (labeled 50 Coop 10 All-D) of Tables 2 and 3. The figures in the two Tables denote the percentage of time cooperative outcomes occur. There is not a single instance where exploitive agents that benefited from selective interaction successfully invade a world of cooperative agents. **Selective interaction is simply of no help to exploitive agents.** There were only two runs where exploitive agents endowed with favorable payoff differentials were able to successfully invade the cooperative agents. The two cases occurred in Assurance games where All-Defect agents were paired against Grim and TFT agents.

Table 2 and Table 3 about here

Asymmetric power appears to be of almost no help to exploitive agents. To test the robustness of this result, an additional set of simulations was run with exploitive agents that employ strategies that are more manipulative and more likely to be able exploit cooperative agents than those that employ the All-Defect strategy. There are a number of candidates for exploitive strategies that are more "manipulative" or at least more "devious" than All-Defect. In most games, All-Defect can be an effective strategy against the "naïve" All-cooperate strategy but far less so against provocable strategies such as TFT and Grim. A potentially more effective strategy is mean tit-for-tat (MTFT). An agent employing a MTFT strategy cooperates when it first plays another agent always defects when the other agent defected in the previous iteration, and, most importantly, defects with a probability (1-P) if the opponent cooperated on the last iteration of the game.[18] Because agents employing such a strategy cooperate more than All-Defect strategies, they are likely to have more opportunities to benefit from exploiting their opponent. This may put them in a position to be more likely to succeed in invading a large number of agents employing various types of cooperative strategies.

Simulations were run across the twenty game/cooperative agent strategy/asymmetric type cases (noted above) with MTFT agents replacing All-Defect agents. The simulation results are strikingly similar to those reported for exploitive All-D agents in Tables 2 and 3.[19] Asymmetric power fails to help even what appear to be more "sophisticated" exploitive agents.

The evidence presented to this point indicates that asymmetric power is essentially of no help to exploiting agents in their efforts to invade worlds populated with significant numbers of cooperative agents. So let us ask a different question. Does asymmetric power help exploitive agents *prevent invasion* by various types of cooperative agents? To answer this question simulations were run where 50 All-Defect agents defended against invasion by ten All-cooperate agents for the Assurance and Stag games, ten TFT agents for PD, Chicken, Assurance and Stag games, and ten Grim agents for PD, Chicken, Assurance and Stage games.[20] The results are also reported in Tables 2 and 3.

Asymmetric power does help exploitive agents defend against invasion. All-Defect agents, who benefit from payoff differentials, defend against invasion 100% of the time in seven of the ten-game/cooperative strategy cases (see right most Column of Table 3 where the figures denote the percentage of time cooperative outcomes occur). Only in the most conflictual games (PD and Chicken) do they defend successfully **less** than 100% of the time; 80% against invading TFT in PD, 40% against invading Grim agents in Chicken, and 0% against invading TFT agents in Chicken games.

The ability to selectively interact also significantly increases the chances that exploitive agents can ward off invasion by cooperative agents in all cases (see right most column of Table 2). While exploitive agents endowed with the ability to selectively interact are generally less effective in defending against invasion than those endowed with payoff differential, they are just as effective as payoff differentials in five cases and more effective in one case (Chicken/TFT).

The simulation results suggest that exploitive agents with asymmetric power cannot invade a world dominated numerically by cooperative agents but have considerable success in defending against invasion when they dominate the world numerically. The obvious question is what happens in worlds that are not so dominated initially either by cooperative or exploitive agents. Is there gradual transition toward exploitive agent success as the world become more populated with exploitive agents or is the transition sudden? In order to address this question, simulations were run for the twenty game/cooperative agent strategy/asymmetric type cases for three separate mixes of cooperative and exploitive agents; 40 cooperative and 20 exploitive agents, 30 cooperative and 30 exploitive agents, and 20 cooperative and 40 exploitive agents. The results of the simulation runs are reported in columns 4-6 of Tables 2 and 3. An examination of these results suggests that the transition toward exploitive agent success as the world becomes more populated with exploitive agents is usually quite abrupt.[21] The transitions start to occur only when worlds are dominated numerically (at least a 2 to 1 margin) by exploitive agents. An exception to this pattern occurs in Assurance games where exploitive agents benefit from payoff differences.

A comparison of the full set of results reported in Tables 2 and 3 lend further support indicating that payoff differentials are more helpful to exploitive agents than selective interaction. Second, the results also indicate that payoff differentials help exploitive agents a great deal in more cooperative settings such as Assurance and Stag and increasingly less so as the settings become more competitive (PD and Chicken). Third, not surprisingly, exploitive agents with either form of asymmetric power are more successful when competing against cooperative agents that employ more exploitable strategies (All-C) than less the forgiving TFT and Grim strategies.

The two forms of asymmetric help exploitive agents to succeed in quite different ways. Exploitive agents endowed with the ability to selectively interact succeed in defending against invasion because they induce a highly non-interactive world. They do not interact with each other and the few cooperative agents initially present in the world usually are isolated from each other thus have great difficulty obtaining any gains from cooperation. The fact that more sophisticated MTFT exploitive agents, that are more likely to interact more with other exploitive agents and cooperative agents, fail to do any better than All-Defect agents suggests that non-interaction and isolation of cooperative agents from each other is the real key to exploitive agent success. Exploitive agents endowed with payoff differentials succeed in defending against invasion by taking advantage of more forgiving cooperative agents (except in Assurance situations) and by benefiting from interactions with cooperative agents in game situations that are more cooperative.

Discussion

There is clearly an asymmetry to the provision of asymmetric power. Asymmetric power helps cooperative agents invade an exploitive, nasty, and conflict ridden world and build stable cooperation because it provides cooperative agents with a mechanism to build a network of cooperators that generates prosperity and effectively isolates exploiters and eventually eliminates them.[22] Asymmetric power fails to help exploitive agents invade and tear down cooperative worlds because it does not enhance the ability of exploiters to survive in the early stages of the simulation since there are very few other exploiters to avoid and networks of exploiters (whether they have asymmetric power or not) do not form to compete with networks of cooperators since there are no benefits derived from networks of exploiters. This is good news for international relations cooperation theorists. Small groups of exploitive agents, even when privileged, have little success in taking advantage of much less destroying communities of cooperative agents.

However, asymmetric power does help exploitive agents defend against invasion by cooperative agents particularly when they benefit from favorable payoff differentials. But, exploitive agents with either form of asymmetric power only succeed in maintaining dominance when they hold a distinct numerical edge. Indeed, exploitive agents usually must hold at least a 2 to 1 numerical edge in order to be able to fend off "invasion" by cooperative agents. Again, this is good news for international relations cooperation theorists. Most international systems are not composed of a high proportion of exploitive and essentially uncooperative agents.

Exploitive agents that benefit from selective interaction are more successful at fending off invasion by cooperative agents than those without that benefit. Their success is, however, due less to their ability to increase interactions with cooperative agents whom they can exploit but rather to the fact that they can reduce interactions with other exploitive agents. The world becomes almost totally devoid of interactions among agents and over time the initial numerical superiority of the exploitive agents allows them to prevail. Also, exploitive agents benefiting from selective interaction fare less well when cooperative agents employ less exploitive strategies such as Grim and TFT. This suggests that the circumstances under which selective interaction would benefit exploitive agents in international relations is quite small because most agents in international politics do not employ highly exploitable strategies (most agents do not consistently allow exploiters to go unpunished) and, as noted above, most international systems are not populated mostly with exploitive agents.

Exploitive agents that benefit from payoff differentials are much more successful in fending off invasion by cooperative agents than those without that benefit. Their relative success is based on three factors. First, they benefit more when they can exploit cooperative agents. Second, while they still suffer when they interact with other exploitive or unforgiving cooperative agents, the costs they incur from those interactions (typically mutual defection) are reduced. Third, as was the case for exploitive agents with selective interaction, they benefit by having cooperative agents isolated from each other. The first two factors help to account for why exploitive agents that benefit from payoff differentials fare uniformly better in the most cooperative game setting (Assurance), less so in more conflictual settings (Stag and PD), and realize almost no benefit in the most conflictual setting (Chicken). In cooperative games the costs from mutual defection are sufficiently lower than in conflictual games so that the payoff differential benefit afforded exploitive agents makes a significant difference.

Mutual defection in Chicken games is the least preferred outcome for all agents and so, even if there is a payoff advantage afforded to exploitive agents, it is not surprising that agents that prefer cooperation and obtain the most preferred mutual cooperation outcome dominate. Repeated Chicken structural settings such as bargaining about trade agreements with threats of sanctions and the possibility of trade wars usually end in cooperative arrangements. In Stag and PD games that capture phenomena such as security dilemmas, arms race contexts and arms control processes, obtaining mutual defection, the second least preferred outcome, is not that helpful to exploitive agents even when they benefit from payoff differentials. The higher rate of success that exploitive agents in the Stag and PD contexts enjoy compared to Chicken settings is more likely attributed to the fact that unilateral defection is a substantially more valued outcome. In Assurance game settings, while agents most prefer mutual cooperation, their second most preferred outcome is mutual defection. Assurance game contexts are those where it only makes sense for all agents to either cooperate or defect. Therefore, if exploitive agents prefer defection and do comparatively better than the cooperative agents when mutual defection occurs, it is not surprising that exploitive agents prevail even when they begin at a numerical disadvantage.

Conclusions

Does asymmetric power, in the form of either differences in capabilities agents are endowed with or in payoffs received from interactions, help exploitive agents to the same extent that it has been shown to benefit cooperative agents in a variety of settings of conflict and cooperation prevalent in international politics? To address this question a set of agent-based models based upon a group of repeated 2X2 games (Prisoners Dilemma, Chicken, Stag, Assurance, and Deadlock) was developed and a series of simulations were run comparing the ability of exploitive agents with and without asymmetric power to invade worlds dominated by cooperative agents. Contrary to expectations, the simulation results indicate that exploitive agents are not as fortunate as cooperative agents. Providing exploitive agents with two different forms of asymmetric power, selective interaction or payoff differentials, fails to enhance their ability to invade groups of cooperative agents. Indeed, exploitive agents with asymmetric power fail to "invade" even when the world is initially populated with an equal number of exploitive and cooperative agents.[23]

Exploitive agents, whether they are endowed with asymmetric power or not, appear to find themselves in one of two situations in most structural settings of conflict and cooperation prevalent in international politics. Either exploiters lead a lonely (few interaction opportunities with other agents) and numerically insignificant existence in a world dominated by a cooperative regime (most effectively populated by cooperative agents that do not unilaterally defect but do retaliate), or they dominate a world where there is essentially no cooperation, little to no interaction among agents, and very little prosperity (low payoffs). At best, asymmetric power helps exploitive agents maintain this bleak, low interaction Hobbesian-type world.

Table 1

Exploitive Agent Benchmark Simulation Results

	Overall	Grim	TFT	All-C
Assurance	SC 100% NC 0%	SC 100% NC 0%	SC 100% NC 0%	SC 100% NC 0%
Stag	SC 100% NC 0%	SC 100% NC 0%	SC 100% NC 0%	SC 100% NC 0%
Prisoners Dilemma	SC 67% PE 33% NC 0%	SC 100% PE 0% NC 0%	SC 100% PE 0% NC 0%	SC 0% PE 0% NC 100%
Chicken	SC 33% PE 33% NC 33%	SC 100% PE 0% NC 0%	SC 0% PE 100% NC 0%	SC 0% PE 0% NC 100%
Deadlock	SC 0% NC 100%	SC 0% NC 100%	SC 0% NC 100%	SC 0% NC 100%
Total	SC 60% PE 13% NC 27% N=75	SC 80% PE 0% NC 20% N=25	SC 60% PE 20% NC 20% N=25	SC 40% PE 20% NC 40% N=25

Table 2 Exploitive Agents with Selective Interaction						
Game Structure	Cooperative Strategy	50 Coop 10 All-D	40 Coop 20 All-D	30 Coop 30 All-D	20 Coop 40 All-D	10 Coop 50 All-D
Assurance	Grim	100%	100%	100%	100%	80%
Assurance	TFT	100%	100%	100%	100%	40%
Assurance	All-C	100%	100%	100%	100%	0%
Stag	Grim	100%	100%	100%	80%	40%
Stag	TFT	100%	100%	100%	100%	40%
Stag	All-C	100%	100%	100%	0%	0%
PD	Grim	100%	100%	100%	100%	0%
PD	TFT	100%	100%	100%	100%	20%
Chicken	Grim	100%	100%	100%	80%	60%
Chicken	TFT	100% *	100% *	100% *	100% *	40% *
* Denotes instances of Punctuated Equilibrium and settling to .5 cooperation level						

Table 3 Exploitive Agents with Payoff Differentials						
Game Structure	Cooperative Strategy	50 Coop 10 All-D	40 Coop 20 All-D	30 Coop 30 All-D	20 Coop 40 All-D	10 Coop 50 All-D
Assurance	Grim	90%	0%	0%	0%	0%
Assurance	TFT	90%	0%	0%	0%	0%
Assurance	All-C	100%	80%	0%	0%	0%
Stag	Grim	100%	100%	100%	60%	0%
Stag	TFT	100%	100%	100%	40%	0%
Stag	All-C	100%	100%	80%	20%	0%
PD	Grim	100%	100%	100%	80%	0%
PD	TFT	100%	100%	100%	40%	20%
Chicken	Grim	100%	100%	100%	100%	60%
Chicken	TFT	100% *	100% *	100% *	100% *	100% *
* Denotes instances of Punctuated Equilibrium and settling to .5 cooperation level						

Endnotes

[1] See Arthur Stein, 'Coordination and Collaboration: Regimes in an Anarchic World'; *International Organization,* 36 (1982), 299-324; Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984); Robert Keohane, *After Hegemony* (Princeton, NJ: Princeton University Press, 1984); Duncan Snidal, 'Coordination Versus Prisoner's Dilemma: Implications for International Cooperation and Regimes'; *American Political Science Review*, 79 (1985), 23-42; and Kenneth Oye, ed, *Cooperation Under Anarchy* (Princeton, NJ: Princeton University Press, 1986).

[2] As James Fearon, 'Bargaining, Enforcement, and International Cooperation'; *International Organization*, 52 (1998), 269-305 suggests, international relations cooperation theorists have tended to focus on various strategic structures that they assert fit different international issues.

[3] Stuart Bremer and Michael Mihalka, 'Machiavelli in Machina: Or Politics Among Hexagons', in Karl Deutsch, ed, *Problems of World Modeling* (Boston: Ballinger, 1977), pp. 303-338; Thomas Cusack, and Richard Stoll, *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation* (Boulder, CO: Lynne Rienner, 1990), and Lars-Erik Cederman, *Emergent Actors in World Politics: How States and Nations Develop and Dissolve* (Princeton: Princeton University Press, 1997) developed simulation and agent-based models that focused on power and differences or asymmetries in power among agents.

[4] See Stephen Majeski, 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations'; *International Studies Quarterly*, 48 (2004), 455-470.

[5] An agent-based approach has been used to model a variety of political phenomena (see Joshua Epstein and Robert Axtell, *Growing Artificial Societies: Social Science from the Bottom Up* (Washington, D.C.: The Brookings Institution, 1996); Robert Axelrod, *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*, (Princeton: Princeton University Press, 1997); Robert Axelrod, 'The Dissemination of Culture: A Model of Local Convergence and Global Polarization'; *Journal of Conflict Resolution,* 41 (1997), 203-226; Cederman, *Emergent Actors in World Politics*; Joshua Epstein, 'Zones of Cooperation in Demographic Prisoner's Dilemma'; *Complexity,* 4 (1998), 33-48; Stephen Majeski, Greg Linden, Corina Linden, and Aaron Spitzer, 'Agent Mobility and the Evolution of Cooperative Communities'; *Complexity,* 5 (1999), 16-24; Ian Lustick, 'Agent-based Modeling of Collective Identity: Testing Constructivist Theory'; *Journal of Artificial Societies and Social Simulation,* 3:1 (2000), http://www.socsurrey.ac.uk/JASSS/3/1/1.html; Ravi Bhavnani and David Backer, 'Localized Ethnic Conflict and Genocide: Accounting for Difference in Rwanda and Burundi'; *Journal of Conflict Resolution,* 44 (2000), 283-307; Ian Lustick and Dan Miodownik, 'Deliberative Democracy and Public Discourse: The Agent Based Argument Repertoire Model'; *Complexity,* 5 (2000), 13-30; Lars-Erik Cederman, 'Modeling the Democratic Peace as a Kantian Selection Process'; *Journal of Conflict Resolution*, 45:4 (2001) 470-502; Christoph Zott, 'When Adaptation Fails: An Agent-Based Explanation of Inefficient Bargaining under Private Information', *Journal of Conflict Resolution*. 46

(2002), 727-753; Lars-Erik Cederman, 'Endogenizing Geopolitical Boundaries with Agent-based Modeling'; *Proceedings of the National Academy of Sciences of the United States*, 99 (2002), 7296-7304; David Sylvan and Stephen Majeski "An Agent-based Model of Acquisition of U.S. Client States", presented at the 44th Annual Convention of the International Studies Association, Portland, February 25-March 1, 2003)

[6] Agents face different incentives in each of the five 2X2 games. 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. See Oye, *Cooperation Under Anarchy* for a description of these and other games in an international relations context as well as an assessment of the relatively difficulty of achieving cooperation across the various game structures. See Majeski, 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations' for a discussion of the preference orderings over the four outcomes of each game, the Nash equilibria for the five games, and key differences in the preference orderings over outcomes that affect the likelihood of cooperation.

[7] See Glenn Snyder and Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises* (Princeton: Princeton University Press, 1977) for a discussion of PD, Chicken and Deadlock games and their application to international conflicts; Robert Axelrod and Robert Keohane, 'Achieving Cooperation Under Anarchy: Strategies and Institutions', in Kenneth Oye, ed, *Cooperation Under Anarchy*, (Princeton: Princeton University Press, 1986), pp. 226-254 for a discussion of

PD and the prevalence of Stag, Chicken and Deadlock situations in international relations; Robert Jervis, ‘Cooperation Under the Security Dilemma’; *World Politics,* 30 (1978), 167-214 for a discussion of Stag, PD, and Chicken and their application to international conflicts; Vinod Aggarwal, *Debt Games: Strategic Interaction in International Debt Restructuring* (Cambridge: Cambridge University Press, 1996) for a discussion of PD and Chicken and their application to international debt rescheduling; John Conybeare, *Trade Wars* (New York: Columbia University Press, 1987) for a discussion of PD and tariff and non-tariff barrier policies in trade; Michael Taylor, *The Possibility of Cooperation* (Cambridge: Cambridge University Press, 1987) for a discussion of Chicken and Assurance games and their application to public goods and collective action; Matthew Evangelista, ‘Cooperation Theory and Disarmament Negotiations in the 1950s’; *World Politics*, 42 (1990), 502-528 for a discussion of PD and arms control; George Downs, David Rocke, and Randolph Siverson ‘Arms Races and Cooperation’, in Oye, *Cooperation Under Anarchy*, 118-146 and George Downs and David Rocke, *Tacit Bargaining, Arms Races and Arms Control* (Ann Arbor: University of Michigan Press, 1990) for an application of PD to arms racing; Glenn Snyder, ‘The Security Dilemma in Alliance Politics’; *World Politics,* 36 (1984), 461-495 for a discussion of PD and competitive alliance formation; Lisa Martin, *Coercive Cooperation: Examining Multilateral Economic Sanctions* (Princeton: Princeton University Press, 1992) for a discussion of PD as co adjustment games in multilateral economic sanctions; Frank Zagare, ‘Rationality and Deterrence’; *World Politics,* 42 (1990), 238-260 for a discussion of Chicken games and deterrence; and Arthur Stein, *Why Nations Cooperate* (Ithaca: Cornell University Press, 1990) for the application of

PD, Chicken, and Deadlock to international conflict and Stag (he labels the game Assurance) to international collaboration.

[8] It is not surprising that the introduction of selective interaction increases the likelihood that cooperative agents survive and generate high levels of cooperation in an initially hostile environment. Selective interaction is a form of non-compulsory play and, like providing agents with the option to exit, withdraw, or move away from undesirable locations, has been shown to increase the likelihood of cooperation. Selective interaction helps agents that want to cooperate because they can choose to interact only with those agents that also cooperate and avoid interacting with those agents that defect.

[9] An exploitive strategy is one that intentionally or unprovokedly defects while the opponent cooperates. Some typical exploitive strategies are All-Defect (always defect regardless of the prior outcome), cheating tit-for-tat (defect after the opponent cooperates for n times in a row since either player last defected), and random tit-for-tat (play TFT except that following cooperation by the cooperate with probability P and defect with probability 1-P. Mean tit-for-tat, discussed and also analyzed in this analysis, is another exploitive strategy.

[10] Surprisingly, the effects of providing exploitive agents with asymmetric power have not been systematically examined.

[11] See Majeski, 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations', pp. 456-462

[12] See Axelrod, *The Evolution of Cooperation*

[13] The model is used to carry out an experiment via simulation. A number of simulations are run with differing mixes of strategies where exploitive agents do not have asymmetric

power. Then the same set of simulations is run with only one change; the introduction one type of asymmetric power and the simulation results are compared to assess the impact of this "manipulation."

[14] The Assurance game depicted here is based upon the game elaborated by Taylor, *The Possibility of Cooperation*, p. 18-19, 38-39 and is consistent with its use by Axel Franzen, 'Group Size and One-Shot Collective Action'; *Rationality and Society,* 7 (1995): 183-200. The structure of the Stag game presented here is consistent with what Mark Lichbach, *The Cooperator's Dilemma* (Ann Arbor: University of Michigan Press, 1996) and Stein, *Why Nations Cooperate* call Assurance games.

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

[16] Martin Nowak and Karl Sigmund, 'A Strategy of Win-Stay, Lose-Shift That Outperforms TFT in the Prisoner's Dilemma Game'; *Nature*, 364 (1993), 56-8 labeled these types of patterns punctuated equilibria and I use the term here for the sake of consistency.

[17] If exploitive agents invade successfully 100% of the time without asymmetric power, then there is no way to determine whether providing asymmetric power to exploitive agents in these contexts is beneficial.

[18] In the simulations, agents employ the following MFTF strategy [0.3, 0.0, 0.3, 0.0] where the first number in the [] denotes the probability of cooperating following a CC

outcome, the second following a CD outcome, the third following a DC outcome,and the fourth following a DD outcome.

[19] While there are minor variations the results are similar enough to make their presentation quite repetitive and unnecessary.

[20] These ten game/strategy cases are the only situations where fifty exploitive agents without asymmetric power could not defend against invasion by various types of cooperative agents 100% of the time (see Majeski, 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations').

[21] The main exceptions occur in the highly conflictual Chicken game setting where whatever transitions occur are gradual.

[22] See Majeski, 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations' for a detailed discussion of how asymmetric power helps cooperative agents.

[23] The exception occurs in Assurance game settings when exploitive agents benefit from payoff differentials.