# The Emergence of Institutions of Self-Governance on the Commons

**Robert Axtell**†
<raxtell@brook.edu>

Center on Social and Economic Dynamics
Economic Studies Program
Brookings Institution
1775 Massachusetts Avenue, NW
Washington, DC 20036

This version: October 2003



† Prepared for the Conference on Computational Political Economy at the University of Michigan. For helpful conversations and comments I thank Peyton Young and Conference participants in Ann Arbor. Financial support from the MacArthur Foundation the National Science Foundation is gratefully acknowledged.

## ABSTRACT

A model is described in which a society of heterogeneous agents solves resource allocation problems in the presence of decreasing returns. Each agent has preferences for both income and leisure and supplies labor effort. Each agent's income is a fraction of total production in proportion to the quantity of total labor it supplies. Agents adjust the amout of labor they provide to maximize individual welfare. Nash equilibrium effort levels are Pareto-dominated by labor configurations—involving less effort—that are not individually rational. A decentralized mechanism for achieving solutions that can Pareto dominate the Nash equilibrium is described and studied using an agent-based computational model. In particular, when boundedly rational agents contract with other agents to reduce their labor inputs then Pareto improvements can be achieved. Such contracts can be interpreted as joint monitoring relationships. The resulting system of contracts can be viewed as a kind of self-governance network. The parameterization of the model is investigated.



keywords: commons, decreasing returns, bounded rationality, agent-based computational model, decentralized exchange, economic complexity

JEL classification codes: C63, C73, D62, D71

# 1 Introduction

When a resource is held in common by society, and free access to it is permitted, self-interested agents are wont to over-exploit it. This is the so-called "tragedy of the commons," popularized by Hardin [1968] and studied by many.[1] Examples of over-exploited commons abound, and include natural resources (e.g., fisheries) and man-made systems (e.g., the Internet), resources that are fixed spatially (e.g., forests) as well as fugitive ones (e.g., irrigation water), areas small in scale (e.g., pastures) and others vast (e.g., overpopulation of the Earth).

The prisoner's dilemma is often invoked as a formalization of the strategic situation on the commons. The dominant strategy (defect) is equated with over-exploitation, while the Pareto optimal outcome (cooperate-cooperate) is associated with the socially-optimal level of exploitation.

A somewhat richer economic formalization involves self-interested agents harvesting a resource in the presence of decreasing returns to scale. This differs from the simple prisoner's dilemma insofar as no dominant strategy exists (Dasgupta and Heal [1979]). Nash equilibria generally exist in such an environment, but these are not efficient—they represent over-exploitation.

Various alternative solutions to commons problems have been compared by Roemer [1996]. These solutions are motivated by axiomatic considerations. That is, solutions are required to possess certain properties, like private ownership of labor (self-ownership), Pareto efficiency, and so-called 'free access on linear economies'.[2] Four mechanisms satisfy these criteria. They are:

- ❏ Nash dominator solution—a firm operates the common and agents are assigned Walrasian allocations associated with firm ownership shares based on the share of labor each contributed at the Nash equilibrium;
- ❏ equal benefits solution—a firm runs the common and firm ownership shares are distributed equally, resulting in Walrasian allocations of the profits that yield equal gains for all individuals;
- ❏ proportional solution—each agent receives a share of the output that is proportional to their harvest from the commons;
- ❏ constant returns equivalent solution—each agent receives the utility that it would receive in a certain constant returns environment under free access.

Only the first of these solutions Pareto dominates the common ownership solution in general.[3]

---

1 Ostrom [1990: 2-3] gives some intellectual history.

2 This means, in essence, that in a constant returns environment, where no commons problem can exist, the solutions specialize to the Nash equilibrium.

3 This result is conceptually similar to Weitzman [1974].

While these four solutions all have desirable properties, the question arises as to whether any of them can be obtained by decentralized mechanisms? That is, can strategic environments be designed that give these solutions as the outcome of some game? It turns out that the equal benefits and proportional solutions can be implemented in Nash equilibrium, the constant returns equivalent solution can be implemented in sub-game perfect equilibrium in a two-stage game. But for the most desirable of these solutions, the Nash dominator, there is a negative result:

**Theorem** (Roemer [1989]): There is no efficient mechanism that Pareto-dominates the common ownership equilibrium and can be implemented in Nash equilibrium.

Roemer views this as damaging to Nozick's neo-Lockeanism, insofar as it suggests there is no way to privatize the commons. Either someone will be made worse off—failure to achieve an allocation that is a Pareto improvement—or the heavy hand of the state will be needed—a more centralized mechanism than Nash implementation.

In the real-world, commons are managed by institutions. Since the institutions economists know best are the market and the firm, the range of options typically prescribed by economists for managing common property resources is usually limited to privatization in one form or another (Ostrom [1990: 12-13]).

But extant institutions for managing commons resources look less like firms and more like cooperatives or non-profit associations. Ostrom [1990] has analyzed both successful and unsuccessful institutions. Features that are found in flourishing institutions include [Ostrom *et al*., 1994: 301-302]:

- ❑ the presence of boundary rules, stipulating which agents have access to the common;
- ❑ the presence of authority rules relating to allocation;
- ❑ active forms of monitoring and sanctioning;
- ❑ the absence of grim trigger strategies.

This paper describes initial attempts to get institutions for managing common pool resources to emerge as a result of repeated interactions among individual resource appropriators following certain rules. In particular, self-interested agents who provide a variable input ('effort') to production are placed in an environment characterized by decreasing returns to scale. These agents receive income in proportion to the amount of input they supply. The Nash equilibrium associated with this set-up is analyzed and it is demonstrated that optimizing agents can achieve this equilibrium. Then, agents are permitted to make contractual agreements to reduce their effort levels in order to increase their individual utilities. A set of such agreements constitute a mutual monitoring institution of governance.[4] The ability of this decentralized system

[4] Other multi-agent models of governance include Axelrod [1995] and Cederman [1997].

to achieve Pareto improvements is studied as a function of the number and type of agents.

In the next section (§ 2) some analytical results are obtained. Then, in § 3, a computational model is described and its output analyzed. Finally, § 4 summarizes the main findings and draws conclusions.

## 2 A Variable Effort Model of Common Property Resource Allocation

Here an economic model of the 'tragedy of the commons' is described and analyzed. It involves a population of purposively-behaving agents who individually harvest resources from a common pool resource characterized by decreasing returns to scale.

### 2.1 Set-Up

There is a finite, fixed set of agents, $A$. Each agent supplies some effort level $e_{i\in A} \in [0, 1]$. The total effort in the population is simply

$$E = \sum_{i\in A} e_i . \tag{1}$$

Each agent harvests resources from the common yielding total output, $O$, as a function of $E$, according to

$$O(E) = aE^b \tag{2}$$

with $b \leq 1$. This represents the production function of the commons. The case of $b = 1$ corresponds to constant returns to cooperation, while $b < 1$ amounts to decreasing returns. Decreasing returns means, essentially, that agents working in the same environment produce less than they can as individuals. To see this, consider two agents having effort levels $e_1$ and $e_2$. If each were to work alone on the common the total output would be $O_1 + O_2 = a(e_1^b + e_2^b)$, while working along side of one another they make $a(e_1 + e_2)^b$. Clearly this latter quantity is no larger than the former since $(e_1 + e_2)^b \leq e_1^b + e_2^b$.

The agents in society receive income in proportion to their effort levels. That is, agent $i$ receives a share of total output of $e_i\ O(E)/E$. This corresponds to equal productivity across the population.

Each agent has Cobb-Douglas preferences for income and leisure. All time not spent working is spent in leisure, thus agent $i$'s utility can be written as a function of its effort level, $e_i$, as

$$U^i\left(e_i;\theta_i,E_{\sim i}\right) = \left[\frac{ae_i}{\left(e_i + E_{\sim i}\right)^{1-b}}\right]^{\theta_i}\left(1-e_i\right)^{1-\theta_i}, \tag{3}$$

Note that in this expression the group output, $O$, has been written as a function of $e_i$, with the remainder of the total group effort, $E_{\sim i}$, considered a parameter; $E = e_i + E_{\sim i}$.

## 2.2 Equilibrium

Let us say that each agent knows its preferences, $\theta_i$, and the output of the society, $O$, from which it can determine $E$ and thus $E_{\sim i}$. Individual effort is not observable. Furthermore, each agent, $i$, selects the effort level, $e_i^*$, that maximizes its utility; formally,

$$e_i^* = \arg\max_{e_i}\left[U^i\left(e_i;\theta_i,E_{\sim i}\right)\right]. \tag{4}$$

It is straightforward to show that the solution to equation 4 is given by

$$e_i^*\left(\theta_i,E_{\sim i}\right) = \frac{b\theta_i - E_{\sim i} + \sqrt{\left(E_{\sim i} - b\theta_i\right)^2 + 4\theta_i E_{\sim i}\left(1+\theta_i(b-1)\right)}}{2\left(1+\theta_i(b-1)\right)} \tag{5}$$

For $b = 1$, that is constant returns, $e_i^*(\theta_i, E_{\sim i}) = \theta_i$. Note that (5) depends on $E_{\sim i}$—the amount of effort expended by the other agents—but is independent of $a$. In order to develop some intuition for the general dependence of $e_i^*$ on its parameters, we plot it for $b = 1/4$ in Figure 1 below, as a function of $E_{\sim i}$, for several values of $\theta_i$.

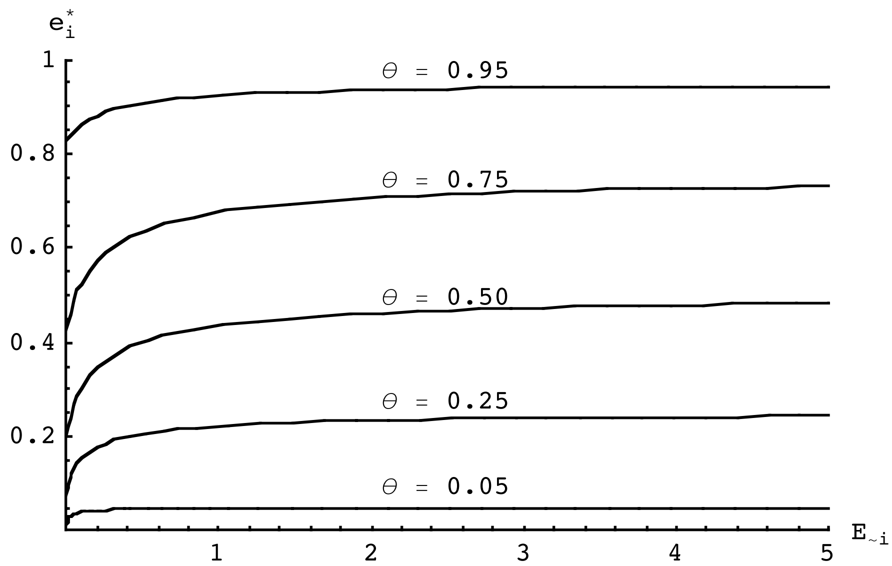


**Figure 1**: Dependence of $e^*$ on $E_{\sim i}$; $b = 0.25$, $\theta_i \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$

Note that the optimal effort level increases monotonically as 'other agent effort,' $E_{\sim i}$, increases. In Figure 2, the dependence of $e_i^*$ on $E_{\sim i}$ as a function of the decreasing returns exponent, $b$, is depicted, for $\theta = 1/2$ and $b \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$.

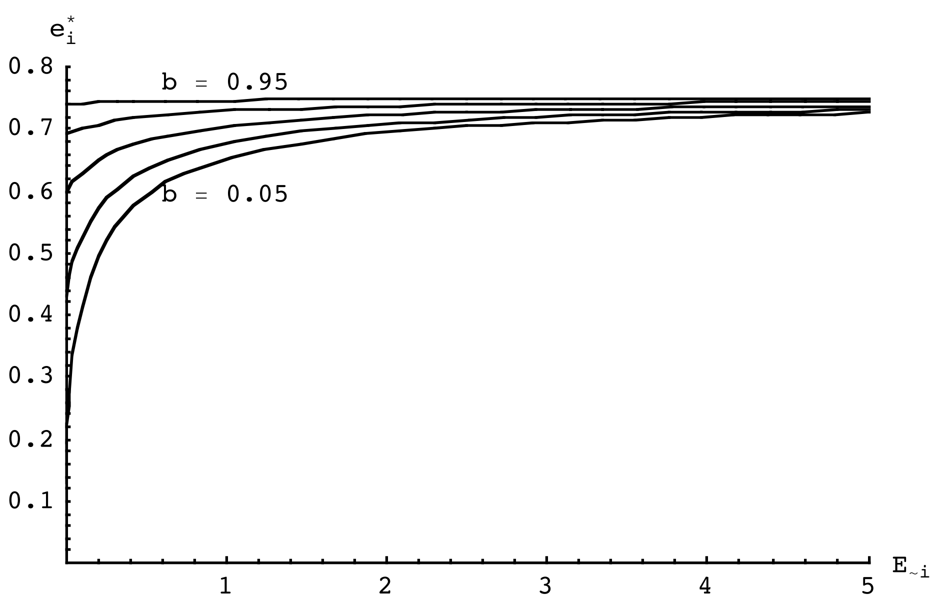


**Figure 2**: Dependence of $e^*$ on $E_{\sim i}$; $\theta = 0.75$, $b \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$

Equilibrium on the commons corresponds to all agents computing their effort levels from equation 5, using $E_{\sim i}^*$ in place of $E_{\sim i}$ such that

$$E_{\sim i}^* = \sum_{i \neq j} e_i^* .$$

From the continuity of the RHS of (5) and the compactness of the space of effort levels, it is clear that solutions to this set of equations exist. Further, given the monotonicity of the RHSs (see figures 1 and 2) the solution is unique. Such an equilibrium configuration is a Nash equilibrium, since once it is established no agent can make itself better off by working at some other effort level.

However, this Nash equilibrium in effort levels is not efficient. In general, there exists a continuous set of agent effort levels that Pareto dominate the Nash equilibrium, as well as a subset—also having cardinality of the continuum—that are Pareto optimal. These solutions all (a) involve smaller amounts of effort than the Nash equilibrium, and (b) are not individually rational. To see (a) note that

$$dU^i\left(e_i^*;\theta_i,E_{\sim i}^*\right) = \frac{\partial U^i}{\partial e_i} de_i + \frac{\partial U^i}{\partial E_{\sim i}} dE_{\sim i} < 0$$

since the first term on the RHS vanishes at the Nash equilibrium and

$$\frac{\partial U^i}{\partial E_{\sim i}} = -\frac{(1-b)\theta_i\left[ae_i\left(e_i + E_{\sim i}\right)^{b-1}\right]^{\theta_i}\left(1-e_i\right)^{1-\theta_i}}{e_i + E_{\sim i}} < 0.$$

Part (b) is true as a result of the fact that each agent's utility is monotone increasing on the interval [0, $e_i^*$), and monotone decreasing on ($e_i^*$, 1]. Therefore,

$$\frac{\partial U^i}{\partial e_i} > 0 \forall e_i < e_i^*, E_{\sim i} < E_{\sim i}^*.$$

This region of effort levels that Pareto dominate the Nash equilibrium is the space of solutions to the 'tragedy of the commons' that we wish to explore.

*Example:* Graphical depiction of the solution space

Consider two agents having identical preferences. Effort level deviations from the Nash equilibrium by either agent alone are Pareto dominated by the Nash equilibrium. If both agents increase their effort levels the utility of each falls, while joint decreases in effort are welfare-improving for both. There exists a symmetric Pareto optimal solution in the case of identical agents. However, no solutions involving less effort than the Nash equilibrium arenot incentive compatible: from any of these solutions each agent gains utility by increasing its effort.

All of this is depicted in Figure 2, below. It is a plot of iso-utility contours as a function of effort levels for two identical agents having $\theta = 0.5$. The lines that are upside down 'U' shaped with respect to the page refer to the first agent, with utility increasing to the bottom. The backward 'C' shaped curves correspond to the second agent, with utility increasing to the left. The point labeled 'N' is the Nash equilibrium. The 'core' shaped region extending below and to the left of 'N' is the set of effort levels that Pareto dominate the Nash equilibrium. The set of effort levels on the curve from 'P' to 'P' are Pareto optimal.

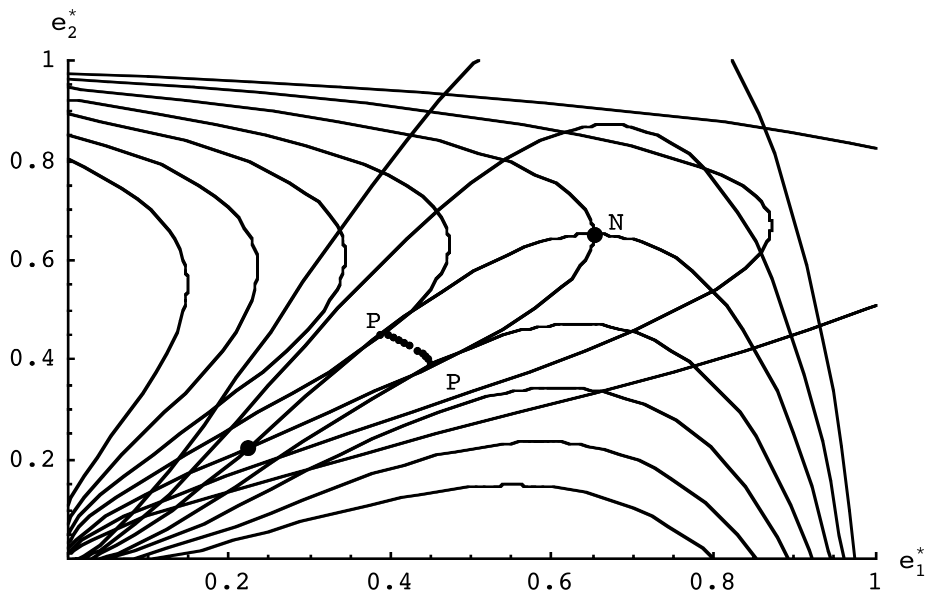


**Figure 3**: Effort level space for two agents each having $\theta = 0.5$; lines are iso-utlity contours, 'N' corresponds to the Nash equilibrium, and the heavy line from P-P corresponds to Pareto optimal solutions

For two agents with distinct preferences for income the qualitative structure of the solution space shown in figure 2 is preserved, but the symmetry is lost. Essentially, decreasing returns insures the existence of solutions that Pareto dominate the Nash equilibrium.

For more than two agents the Nash equilibrium and Pareto optimal solutions continue to be distinct. For $N = 3$, figure 2 can be thought of as the $e_3 = 0$ solution space. Then, because $e^*$ is increasing in $E_{\sim i}$, for $e_3 > 0$ the effort levels of agents 1 and 2 that correspond to the Nash and Pareto optimal solutions are greater than in the $e_3 = 0$ case.

## *Identical Agents*

It is informative to consider a population composed of agents of a single type, that is, all having the same preference for income. In a homogeneous group each agent works with the same effort in equilibrium. This can be determined from (5) above, by substituting $(N\text{-}1)e_i^*$ for $E_{\sim i}$, and solving for $e_i^*$. Doing this yields the following expression

$$e_i^* = \frac{(N + b - 1)\theta_i}{N + (b - 1)\theta_i}. \tag{6}$$

## 2.3 Stability of Equilibrium

While a unique Nash equilibrium always exists in this model, it remains to determine its stability. To investigate stability, the eigenvalues of the Jacobian matrix associated with the set of equations governing equilibrium effort levels must be investigated. This amounts to differentiating (5) with respect to each agent's effort. Doing this yields, for the case of $i \neq j$

$$J_{ij} \equiv \frac{\partial e_i}{\partial e_j} = \frac{-1 + \dfrac{E_{\sim i} + \theta_i\left(2 - 2\theta_i + b\left(2\theta_i - 1\right)\right)}{\sqrt{4\theta_i E_{\sim i}\left(1 + (b-1)\theta_i\right) + \left(E_{\sim i} - b\theta_i\right)^2}}}{2\left(1 + (b-1)\theta_i\right)},$$

while $J_{ii} = 0$. Unlike the case of self-interested agents in an increasing returns environment, where the Nash equilibria are unstable for sufficiently large group size (Axtell [1999]), here it can be shown that the dominant eigenvalue of the Jacobian has modulus strictly less than unity, and thus the equilibrium is always stable.

# 3 Computational Implementation Using Agents

The motivation for a computational version of the above model is simple. Can we come up with a decentralized procedure that collectively moves agents to effort configurations that Pareto dominate the Nash equilibrium? It turns out that the answer to this can be positive. § 3.1 first describes the structure of the agent-based computational model. Then, § 3.2 describes its actual implementation. Finally, typical dynamics of the model are presented and analyzed in § 3.3.

## 3.1 Set-Up of the Computational Model

The agents are as described above, with preferences distrbuted variously (described below). Agents are activated at random, that is, each agent has a Poisson clock that wakes it up periodically. When an agent is active it adjusts its effort levels so as to maximize its utilty, according to (5).[5] If this is all there were to the model then the agents would arrive at the Nash equilibrium, since it is stable, starting from any initial configuration of effort levels.

[5] In order for the computational model to be applicable to agents having arbitrary preferences, each agent performs a line search over the feasible range of efforts in lieu of using the expression given by (5) above.

Therefore, instead of permitting agents to autonomously adjust their efforts, we permit pairs of agents to mutually adjust their effort levels in ways that are advantageous to both. That is, the agents engage in mutual monitoring in order to improve utility.[6] Since agents are not evaluating all possible trades, this model is intrinsically one of bounded rationality.

The initial configuration of the agent population has each agent putting in no effort. Agents 'buy' effort levels for themselves over time, selling others the right to work. The process proceeds until no further welfare-improving exchanges are possible.[7] Decentralized exchange processes of this type lead to allocations that are not generally in the core (from the initial endowment).[8] The resulting history of distributed exchange interactions represents a network of agents who must monitor one another to insure that the Nash equilibrium does not arise as a result of agent cheating.

The essential feature of this model is that it is described at the level of individual agents. Thus it is common to call such models 'agent-based' or 'individual-based'. The only equations present in the model are those governing individual agent decision-making. No attempt has been made to mathematically aggregate the agents' behaviors, and therefore there are no equations governing agent-agent interactions. Rather, "solving" an agent-based model amounts merely to iterating it forward in time and observing the evolution of the agent population, both at the individual and aggregate levels. In this, agent-based computational modeling is similar in spirit to traditional OR simulation. However, in most ways that agent-based computational models have been used to date this methodology is quite unlike conventional simulation.[9]

## 3.2 Object-Oriented Implementation

There are many ways to computationally implement the model just described. This can be done more or less easily in any modern programming language, as well as with any number of mathematical or simulation software packages. However, since the model is stated in terms of individual agents, it turns out that there is one idea from modern computer science that renders the

---

6 General aspects of bilateral exchange models are described in Axtell and Epstein [1997].

7 Technically, there exists a Lyapunov function for this process, so there are no problems with cycling and so on.

8 Conditions under which bilateral exchange equilibria are Pareto optimal are due to Rader [1966] and Feldman [1973].

9 For more on the distinction between simulation and agent-based computational models see Axtell [1997].

implementation both transparent and efficacious. This is the notion of object-oriented programming.

Objects are contiguous blocks of memory that contain both data—so-called instance variables—as well as functions for modifying this data—the so-called methods. This ability of objects to hold both data and functions is called *encapsulation*. Agent-based models are very naturally implemented using objects by interpreting an object's data as an agent's state information, while the object's functions become the agent's rules of behavior.[10] A population of agents that have the same behavioral repertoire but local state information is then conveniently implemented as multiple instantiations of a single agent object type or class.[11]

The model described above has been implemented using object-oriented programming. Individual agents are objects, as is the population of agents as a whole.

The agent object has a variety of state variables and behavioral methods. The state information each agent has includes its preferences and its current effort level. It is also useful to keep track of each agent's income and utility associated with production from the previous period, as well as a running total of its wealth. All of this information is stored locally in the agent object as real numbers. Each agent also keeps track of some number of other agents that it has traded with. This data is maintained in an array of pointers to other agent objects. The behavioral abilities that each agent possesses in the present model include the facility to compute its utility (given some income and effort level) and the capability of determining how much effort level to exchange with other agents. These are the agent object's methods. This agent object specification is summarized in pseudo-code block 1.

```
OBJECT agent;
   preferences;
   effort_level;
   last_income;
   last_utility;
   wealth;
   next_agent_in_agent_list;
   FUNCTION initialize;
   FUNCTION compute_utility;
   FUNCTION draw.
```

**Pseudo-code block 1**: Agent object

---

10 Other features of the object model, including *inheritance* and *polymorphism*, seem to be less relevant to agent-based computational models than encapsulation.

11 For a discussion on the distinction between object and agent, see Jennings *et al.* [1998].

In practice it makes sense to implement as private some of these data and methods, while others are public, although this is not essential.[12]

The agent population is also conveniently implemented as an object. The details are less important here and so pseudo-code will not be given. The main data of the population object is the data structure—commonly either an array or linked list—that holds either the individual agents themselves, or reference to them. The agent population object's methods include agent access routines as well as a host of routines for computing various statistical measures of the agent population, such as the average effort level, average income and so on. This object is summarized as pseudo-code block 2.

```
OBJECT society;
   agent_list;
   size;
   last_output;
   FUNCTION initialize;
   FUNCTION compute_total_effort;
   FUNCTION compute_average_effort;
   FUNCTION compute_output;
   FUNCTION allocate_income_to_agents;
   FUNCTION draw_edge_between_agents(agent1, agent2);
   FUNCTION dispose.
```

**Pseudo-code block 2**: Society object

As with the agent object, in practice it is usually useful to make some of these fields private.

Putting all of this together the computational model simply amounts to (1) initializing all agents, then repeatedly: (2) pairing agents at random and letting them engage in bilateral exchange of effort levels; (3) computing total output; (4) periodically gathering statistics on the populations of agents and firms. This is summarized in pseudo-code block 3.

[12] Private data and methods are accessible only by the agent instance to whom they belong, unless other objects are given special access priviledges.

```
PROGRAM commons;
   initialize agents;
   initialize society;
   repeat:
      select 2 agents at random;
      for each agent selected:
         compute dU/dE~i;
         propose new effort levels;
         check if these are welfare improving:
            if yes, then adopt them;
         draw edge between agents;
      for society:
         compute output;
         allocate income;
         compute welfare;
      compute statistics;
      check for user input;
   until user terminates.
```

**Pseudo-code block 3**: Pseudo-code for the model overall

The object model is largely responsible for the relatively short description of this code.[13] In the next section we describe typical realizations of this model.

## 3.3 Results

In this section some results from a relatively small sample of configurations and realizations of the model are described. Two distinct population sizes are employed, and both homogeneous and heterogeneous populations are analyzed.

### *Configuration #1: Ten Identical Agents*

Consider 10 agents each of whom has $\theta_i = 0.5$. The Nash equilibrium effort level can be computed directly from (6) above and is 17/35. There exists a Pareto optimal allocation of effort that is the same for all agents and this can be readily computed. These are shown in Figure 4, below.

[13] The actual source code runs to some several thousand lines and compiles in the CodeWarrior environment for the Macintosh. A small fraction of the source is Mac-specific. A Java implementation is underway.

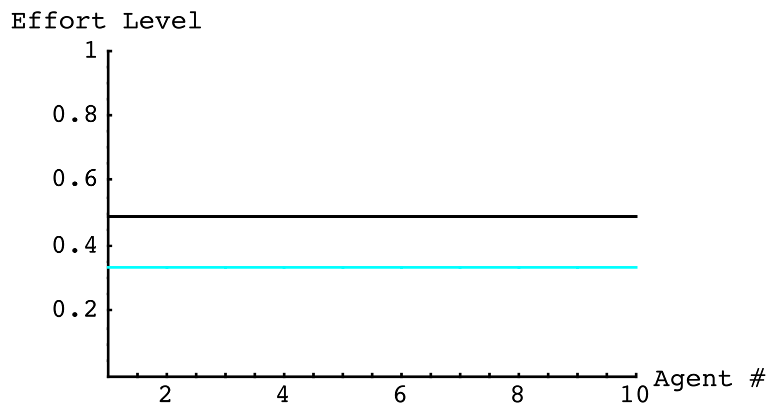


**Figure 4**: Effort levels in a population of 10 agents having $\theta_i = 0.5$, both at Nash equilibrium (top line) and at the symmetric Pareto optimal level (bottom line)

Note that agents put in more effort in Nash equilibrium, thus the over-exploitation of the commons. The utility levels for these effort inputs are shown in Figure 5.

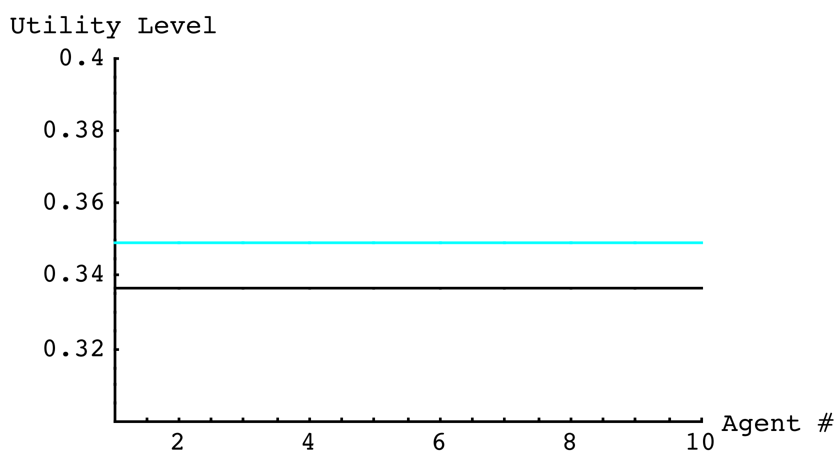


**Figure 5**: Utility levels in a population of 10 agents having $\theta_i = 0.5$, both at Nash equilibrium (bottom line) and at the symmetric Pareto optimal level (top line)

Clearly, the lower effort level Pareto dominates the Nash equilibrium.

Now, starting from all agents at zero effort, we permit the agents to trade effort levels among themselves pairwise, to improve their welfare. Thie process proceeds until no further trade is possible. The result of a particular realization, involving some 23 exchanges, is shown in Figure 6, below.

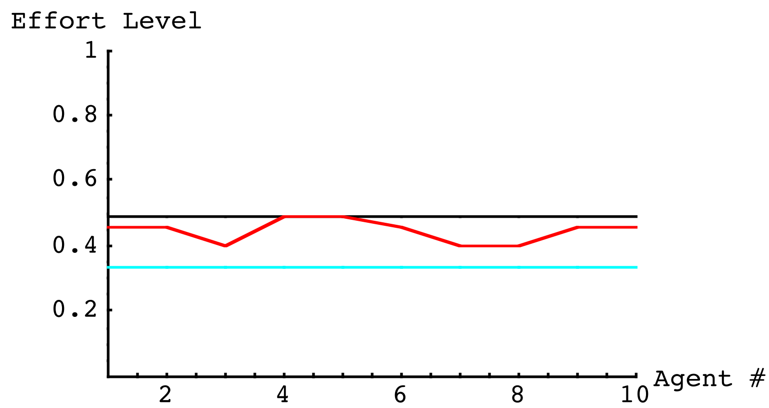


**Figure 6**: Effort levels in a population of 10 agents having $\theta_i = 0.5$, at Nash equilibrium (top line), at the symmetric Pareto optimal level (bottom line), and in bilateral exchange equilibrium (middle line)

Note that the bilateral exchange effort levels are bounded from above by the Nash levels, and from below by the symmetric Pareto ones. Figure 7 gives a plot of the corresponding utility levels.

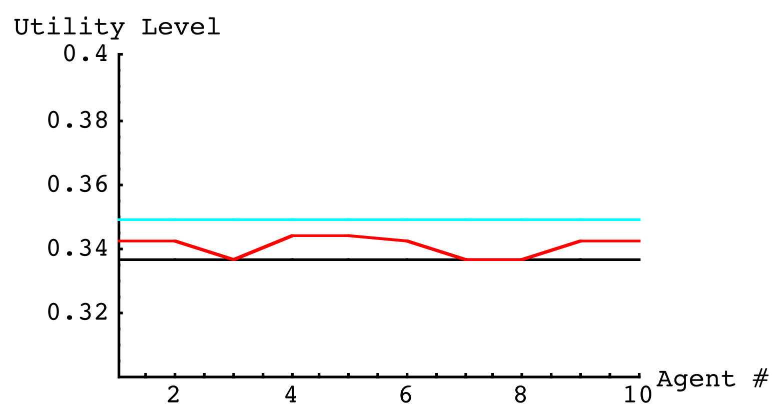


**Figure 7**: Utility levels in a population of 10 agents having $\theta_i = 0.5$, at Nash equilibrium (bottom line), at the symmetric Pareto optimal level (top line), and in bilateral exchange equilibrum (middle line)

The bilateral exchange solution Pareto dominates the Nash (tragedy of the commons) equilibrium, but is dominated by the symmetric Pareto optimal one.

By keeping track of which agents bargain with others—that is, the history of successful agent contracts—we have in essence a mutual monitoring network. For when agents *i* and *j* permit one another to harvest certain amounts from the commons, this is a meaningful constraint on the agents only if each enforces it. It is revealing to plot all contracts, by drawing a line between agents who have entered in to agreements with one another. For the realization described in effort and utility terms above, such a plot is displayed in Figure 8, with the agents arranged in a circle, starting with the first agent at the top and proceeding clockwise.

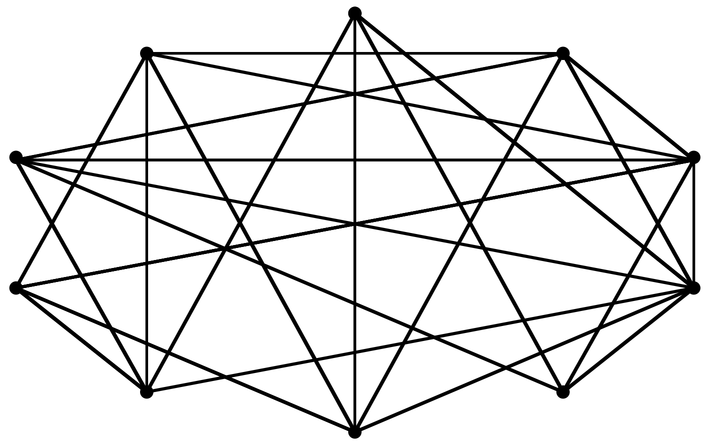

**Figure 8**: Monitoring network arising in a population of 10 agents having $\theta_i = 0.5$, at bilateral exchange equilibrium

In this case, each agent is engaged in at least 4 contracts, while one agent is involved in 7; the average is 4.5 contracts per agent. Thus, each agent would be monitoring, on average, between 4 and 5 other agents.[14]

The particular bilateral exchange solution depicted above is not the only kind of outcome that can happen. Figures 9 - 11 give the result of a different history of agent pairings.

[14] It is a feature of bilateral exchange models that at the population rises the number of bilateral exchanges per agent—here, the amount of monitoring per agent—is constant. That is, the monitoring duties of any particular agent does not increase as the population increases. For more on this, see Axtell and Epstein [1997].

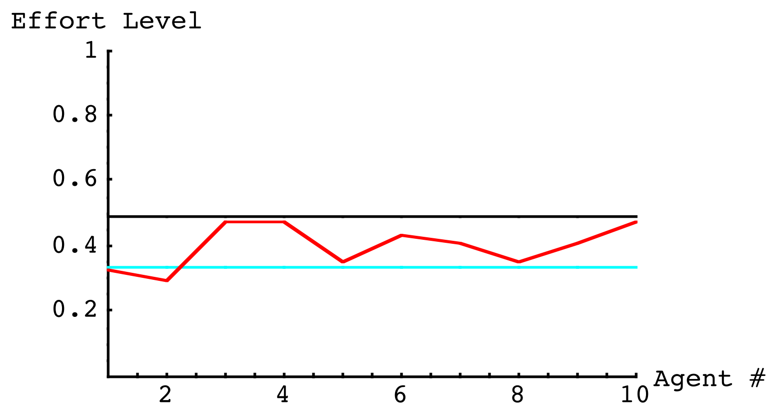


**Figure 9**: Effort levels in a population of 10 agents having $\theta_i = 0.5$, at Nash equilibrium (top line), at the symmetric Pareto optimal level (bottom line), and in bilateral exchange equilibrium (jagged line)

Here, some 27 exchanges took place. In bilateral exchange equilibrium some agents put in more effort in than in Nash equilibrium, while others put in less. The bilateral exchange effort levels of Figure 9 scales to yield Figure 10's utility levels.

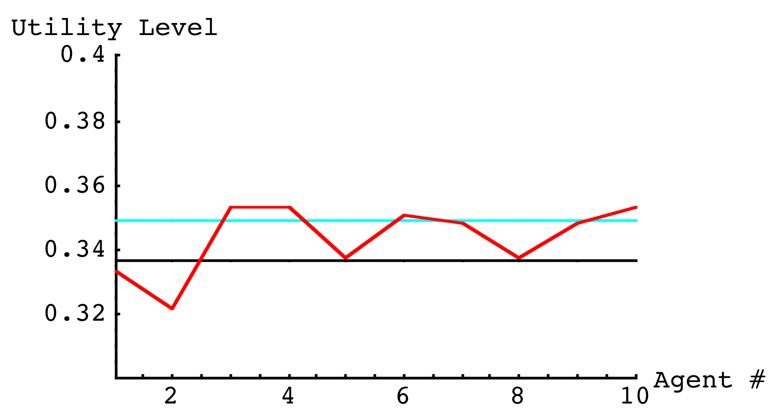


**Figure 10**: Utility levels in a population of 10 agents having $\theta_i = 0.5$, at Nash equilibrium (bottom line), at the symmetric Pareto optimal level (top line), and in bilateral exchange equilibrium (jagged line)

Notice here that the bilateral exchange equilibrium neither dominates the Nash equilibrium nor is it dominated by the symmetric Pareto optimal outcome, even though the symmetric solution dominates Nash.

The contract or mutual monitoring network that has formed is shown in Figure 11. It looks similar to before.

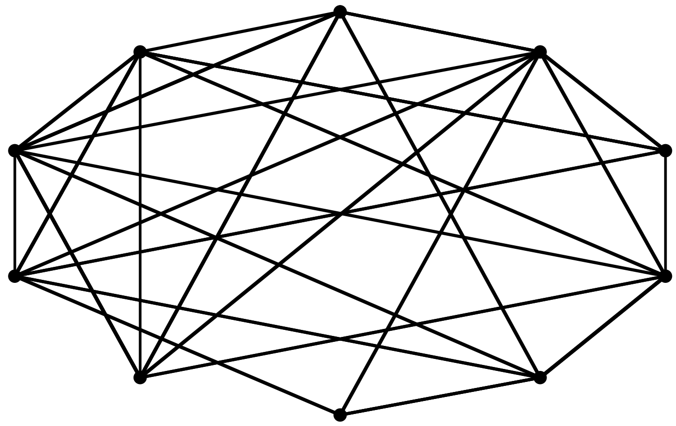

**Figure 11**: Monitoring network arising in a population of 10 agents having $\theta_i = 0.5$, at bilateral exchange equilibrium

Here each agent has 5.4 contracts on average, with the range going from 3 to 7.

## *Configuration #2: Ten Heterogeneous Agents*

We can repeat this analysis for a population of heterogeneous agents, with $\theta \sim U[0, 1]$. First the Nash equilibrium effort and utility levels are found, then the agents are permitted to engage in bilateral exchange. The results are shown in Figures 12 to 14. The agents are ordered from lowest to highest preference for income, i.e., agent 1's $\theta_i$ is lower than agent 2's, and so on.

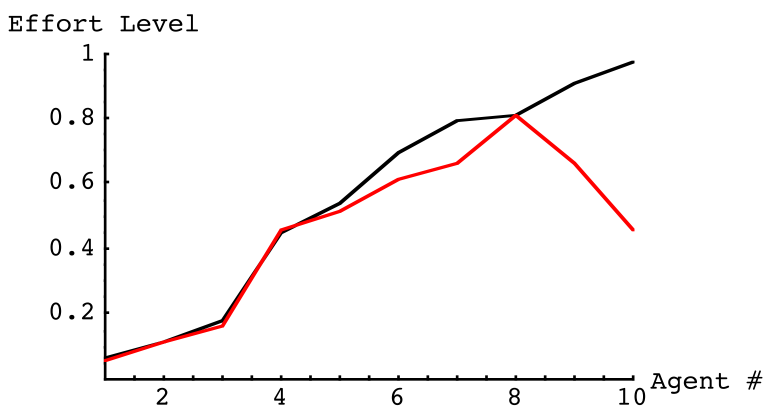


**Figure 12**: Effort levels in a population of 10 heterogeneous agents at Nash equilibrium (top line) and in bilateral exchange equilibrium (bottom line)

Effort levels at the bilateral exchange equilibrium are everywhere below the Nash equilibrium, sometimes by a lot (agent #10) and sometimes by little (agents #1 through #4). The total effort level is approximately 90% of the Nash level.

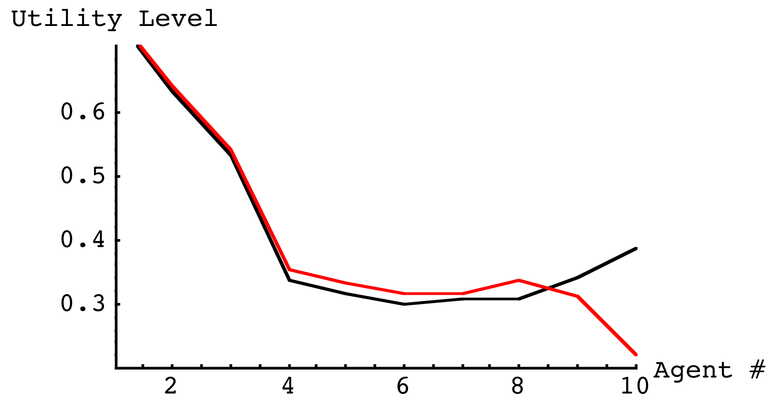


**Figure 13**: Utility levels in a population of 10 heterogeneous agents at Nash equilibrium (bottom line) and in bilateral exchange equilibrium (top line)

The bilateral exchange equilibrum fails to dominate the Nash equilibrium here. Agents #8 and #9 do better at Nash. Failure of bilateral exchange equilibria to completely dominate the Nash outcome becomes increasingly common with large populations, as we shall see below.

The monitoring network for this realization is much less regular than before. Two agents have made but a single contract, while single agents have entered into two and three contracts.

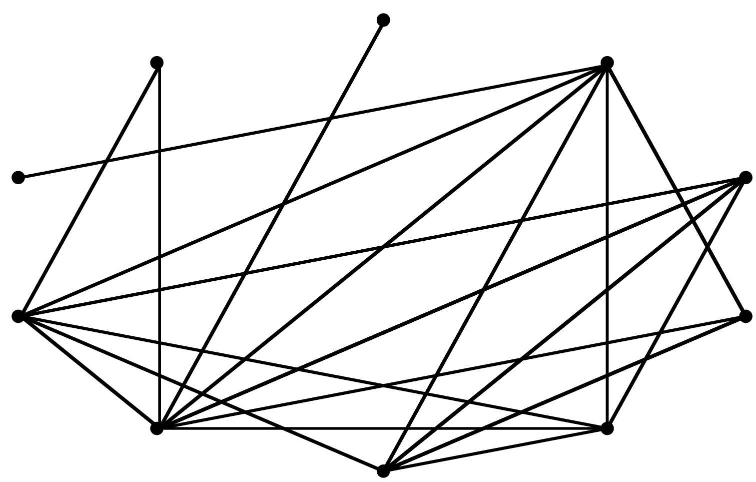

**Figure 14**: Monitoring network arising in a population of 10 heterogeneous agents at bilateral exchange equilibrium

In sum, 20 contracts were agreed upon and each agent monitors 4 of its peers, on average.

## *Configuration #3: Fifty Heterogeneous Agents*

When the population is increased from 10 to 50 agents, the results become more robust, less subject to variation from realization-to-realization. Shown in Figure 15 are the effort levels both at the Nash equilibrium as well as in bilateral exchange equilibrium. Once again, the agents are arranged from lowest preference for income to the highest.

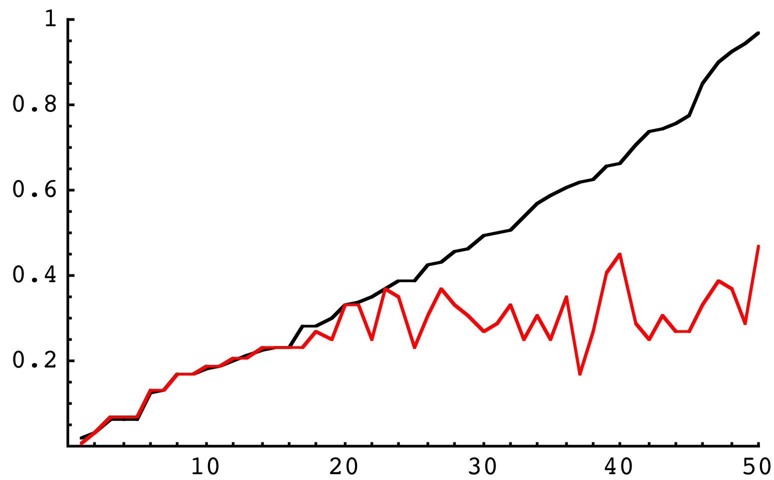


**Figure 15**: Effort levels in a population of 50 heterogeneous agents at Nash equilibrium (top line) and in bilateral exchange equilibrium (bottom line)

As we have now come to expect, the bilateral exchange equilibrium yields significantly lower total effort, approximately 75% of the Nash total in this case. Thus the commons is less exploited in bilateral exchange equilibrium. However, notice that some agents seem to work slightly more than in Nash equilibrium, #6 for instance. The bilateral exchange equilibrium generally dominates the Nash one, but this is not strictly true, as made clear in Figure 16.


**Figure 16**: Utility levels in a population of 50 heterogeneous agents at Nash equilibrium (bottom line) and in bilateral exchange equilibrium (top line)

It is primarily the agents with high preference for income that have lower utility as a result of bilateral exchange (#37 and #42 through #50, for instance). Repeated realizations of this model reveal that strict Pareto improvments are relatively rare, as a few agents are usually worse off in bilateral exchange equilibrium.

The monitoring network for this realization is shown in Figure 17. While this figure is much more dense with lines, the number of contracts per agent is approximately the same as it was in the previous cases.

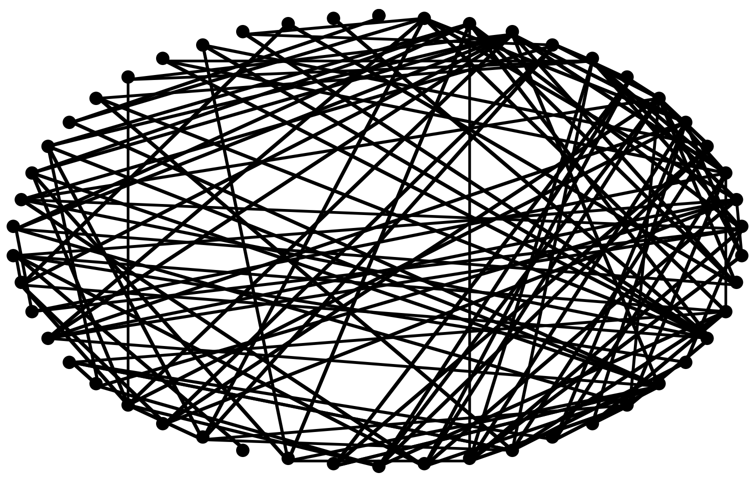

**Figure 17**: Monitoring network arising in a population of 50 heterogeneous agents at bilateral exchange equilibrium

Here, some 224 contracts were written and each agent is involved with 4.5 on average.

## 4 Conclusions

An agent-based model of the 'tragedy of the commons' has been described analytically and explored computationally. Stable Nash equilibrium effort levels exist in this model. These involve agents supplying more than the socially optimal amount of effort. Effort levels that are Pareto superior to the Nash levels exist but are not individually rational.

When agents are permitted to enter into mutually beneficial contracts with one another to reduce their effort levels then tragedy can be partially averted. Sets of such contracts are usefully interpreted as mutual monitoring systems, as are observed in the practice of managing common pool resources. Such monitoring systems are a simple form of appropriator self-governance.

This research represents a first step toward modeling the emergence of institutions of self-governance. Much work remains to be done, particularly exploring the sensitivity of these preliminary results to the parameteriazation of the model. It is expected that a systematic exploration of the parameter space will reveal *quantitative* differences in comparison to the results reported above, but that the *qualitative* nature of the results will remain the same.

## References

Arthur, W.B., J.H. Holland, B. LeBaron, R. Palmer and P. Tayler. 1996. "Asset Pricing Under Endogenous Expectations in an Artificial Stock Market." Santa Fe Institute. Working paper.

Axelrod, R. 1995. "A Model of the Emergence of New Political Actors." In Hillebrand and Stender, eds.

Axtell, R.L. 1999. "Why Agents?" Working paper. Brookings Institution: Washington, D.C.

———. 1999. "The Emergence of Firms Within a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions." Working paper. Brookings Institution: Washington, D.C.

Axtell, R.L. 2003. "The Complexity of Exchange." Working paper. Brookings Institution: Washington, D.C.

Axtell, R.L. and J.M. Epstein. 1994. "Agent-Based Models: Understanding Our Creations." *Bulletin of the Santa Fe Institute*.

Cederman, L-E. 1997. *Emergent Actors in World Politics*. Princeton University Press: Princeton, New Jersey.

Dasgupta, P.S. and G.M. Heal. 1979. *Economic Theory and Exhaustible Resources*. Cambridge University Press: New York.

Dow, G. 1990. "The Organization as an Adaptive Network." *Journal of Economic Behavior and Organization*, *14*: 159-185.

Epstein, J.M. and R. Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press/MIT Press: Washington, D.C./Cambridge, Massachusetts.

Hart, O. 1995. *Firms, Contracts and Financial Structure*. Oxford University Press: New York.

Hart, S. and M. Kurz. 1983. "Endogenous Formation of Coalitions." *Econometrica*, *51*: 1047-1064.

Hillebrand, E. and J. Stender, eds. 1994. *Many-Agent Simulation and Artificial Life*. IOS Press: Washington, D.C.

Holler, M.J., ed. 1984. *Coalitions and Collective Action*. Physica-Verlag: Wuerzburg, Vienna, Austria.

Jennings, N.R., K. Sycara, and M. Wooldridge. 1998. "A Roadmap of Agent Research and Development." *Autonomous Agents and Multi-Agent Systems*, volume 1, number 1: 7-38.

Lane, D. 1993. "Artificial Worlds and Economics." Parts 1 and 2. *Journal of Evolutionary Economics*, *3*: 89-107, 177-197.

Lewin, D.I. 1997. "Growing an a Tree of Knowledge in an Object-Oriented Eden." *Computers in Physics, 11* (1): 6-7.

Luenberger, D.G. 1979. *Introduction to Dynamic Systems*. Wiley: New York.

Mandelbrot, B. 1997. *Fractals and Scaling in Finance*. Springer-Verlag: New York.

Olson, M. 1965. *The Logic of Collective Action*. Harvard University Press: Cambridge, Massachusetts.

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press: New York.

Ostrom, E., R. Gardner and J. Walker. 1994. *Rules, Games, and Common Pool Resources*. University of Michigan Press: Ann Arbor, Michigan.

Prietula, M., K. Carley and L. Gasser, eds. 1998. *Simulating Organizations*. MIT Press: Cambridge, Massachusetts.

Roemer, J.E. 1989. "A Public Ownership Resolution of the Tragedy of the Commons." *Social Philosophy and Public Affairs*, *6*: 74-92.

———. 1996. *Theories of Distributive Justice*. Harvard University Press: Cambridge, Massachusetts.

Roth, A. 1984. "Stable Coalition Formation: Aspects of a Dynamic Theory." In Holler [1984].

Sandholm, T., K. Larsen, M. Andersson, O. Shehory and F. Tohmé. 1998. "Anytime Coalition Structure Generation with Worst Case Guarantees." *Proceedings of the American Association for Artificial Intelligence*.

Simon, H.A. 1997. *The Sciences of the Artificial*. Third edition. MIT Press: Cambridge, Massachusetts.

Weitzman, M. 1974. "Free Access versus Private Ownership of the Commons." *Journal of Economic Theory*.

Young, H.P. 1998. *Individual Strategy and Social Structure*. Princeton University Press: Princeton, New Jersey.

rla/msw