Computer Testbeds for Mechanism Design

Jasmina Arifovic and John Ledyard *

April 18, 2005

Abstract

We develop a behavioral model as a a computer testbed we can use to study the probable performance of a wide range of mechanisms prior to testing them in a laboratory or using them in practice. In this paper, we describe an implementation of our model and the computer testbed methodology to Groves-Ledyard (1977) mechanisms for provision of public goods. Previous experimental evidence, and some theory, strongly suggest that the value of a free mechanism parameter is important for the dynamic performance of the mechanism (when it is simulated as a repeated game). In our model messages converge to the Nash Equilibrium for all of the values of the mechanism parameter that we studied. However, the convergence times depend on the value of the parameter. Our analysis suggests there are values of the free parameter that result in the fastest convergence. The range of values is robust with respect to the changes in the behavioral model's parameter values and details of the updating procedures. This prediction is validated with data from experiments with human subjects.

JEL classification: D83, C63, C92, H41
Keywords: computer testbed, learning, experiments with human subjects, public goods

*Simon Fraser University and California Institute of Technology, respectively. We thank Stephanie Chow and Stephen Tse for their very able research assistance. We would also like to thank Peter Bossaerts, Colin Camerer, Tom Palfrey and Charlie Plott for helpful discussions and participants at the seminars at the California Institute of Technology and the University of California at San Diego, and at the ESA meetings, Tucson, September 2000, the NBER Decentralization Conference, White Plains, October 2000 for useful comments. This material is based on work supported by the National Science Foundation under grant #SES-0121478. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the the NSF.

1 Introduction

Mechanism design has become very sophisticated since its introduction by Hurwicz in 1960. It is now a well developed body of theory taking into account informational and incentive constraints. However, its applications remain difficult and not at all straightforward. Three main roadblocks still exist: (1) much of the theory is about one-shot games while many applications involve repeated play against the same, or very similar, opponents[1], (2) there is no generally accepted view as to the *right* model of individual and group behavior and (3) we have not yet incorporated *computational* limitations, of either the mechanism or the agents, into the theory. In this paper we report on the beginnings of our development of a method to provide support for those doing applied design when repeated play, behavior, and computation are important. In this paper, we use the design of a public goods mechanism as an example of what might be accomplished.

The framework of mechanism design is reasonably simple and well-known. A mechanism specifies the rules of the game - who should communicate with whom and how, as well as what actions to take and when. Given a set of individuals, their preferences and their endowments (all part of the environment), the outcome we observe will be the result of both the mechanism rules and the choices made by the agents. A particularly interesting question is whether individuals in a group are better off under one mechanism than another. To answer this we must be able to evaluate the performance of mechanisms. In order to do that we need to be able to predict what outcomes will occur in each environment when that mechanism is used. We need a *model of behavior* - how agents choose their actions given the mechanism and the environment. Unfortunately there is as yet no generally accurate and agreed upon standard game-theoretic model of behavior. Among the candidates are the use of dominant strategies (if they exist), Bayesian equilibrium, and Nash equilibrium.

Absent a unique, compelling model of behavior some economists have turned to the use of the experimental economics laboratory as a *testbed* for new mechanisms in much the same way that early aircraft designers turned to the wind-tunnel to test their designs. In this approach, one picks a mechanism design, picks a few *environments*, puts people in place, and then runs the mechanism. Performance is measured and comparisons between mechanisms are made. But this is expensive and time consuming. It would be extremely helpful if one had a computer testbed one could, at least, use to eliminate bad designs. Better, one would want to be able to identify the few mechanisms and environments that should be tested in the lab. To create a computer based testbed one needs to come up with a model of behavior that is accurate and robust. Accurate means that when a test pair, of environment and mechanism, are chosen and tested in the computer, then the outcome from the behavior modeled in the computer is close to what would be produced in laboratory experiments with

[1]Even when repeated play is acknowledged, one often invokes the Revelation Principle which masks interesting dynamic features of a problem such as an inability to commit.

human subjects. Robust means that the computer modeled behavior has to be accurate over a wide range of environments and mechanisms and that parameters of the models do not have to be re-estimated whenever there are substantial changes in the class of mechanisms or environments.

The key issue in testbed construction is the model of the agents' behavior. Our approach to modeling behavior is based, to some extent, on evolutionary algorithms[2] such as genetic algorithms, classifier systems, genetic programming, evolutionary programming, etc. A large number of applications have focused on models of *social learning* where a population of agents (each agent is represented by a single strategy) evolves over time such that the entire population of agents jointly implements a behavioral algorithm. However, in some applications (e.g. Arifovic, 1994 and Marimon, McGrattan, and Sargent, 1989), these algorithms have been used as models of individual learning, where evolution takes place on a *set of strategies* that belong to an individual agent. We follow the latter approach.

In this paper, to illustrate our approach and to provide some evidence that it can be successful, we consider a specific class of mechanisms for the provision of a public good - Groves-Ledyard mechanisms.[3] Theory establishes for their tax and allocation rules, in a one-shot game, that the mechanisms yield a Nash equilibrium outcome at a Pareto-optimal level of the public good. But the theory is mostly silent on the dynamics of such a mechanism in a repeated play situation. Two exceptions are the papers by Muench and Walker (1983) and by Chen and Tang (1998), both of which suggest that the dynamics might depend on the value of a free parameter, even though that parameter does not affect the Nash equilibrium outcomes. Muench and Walker (1983) relies on Cournot strategies for their analysis. Chen ant Tang (1998) produce results that apply more broadly to any adaptive learning strategies. Although there is no serious game-theoretic reason to assume agents would adopt adaptive strategies, evidence from the experiments with human subjects does support this theoretical insight. In experiments with human subjects (Chen and Plott, 1996; Chen and Tang, 1998), the Groves-Ledyard mechanism was implemented as a repeated stage game. In these experiments, messages did not converge to Nash equilibrium for a low value of the free parameter, but did converge for a high value. (Chen and Tang, 1998).

To study the dynamics of Groves-Ledyard mechanisms in repeated play, we use our testbed to simulate the dynamics in a particular environment for a wide range of values of the free parameter. The class of mechanisms is described in Section 2. The testbed is described in Section 3. The results of the simulations of the mechanisms in the testbed are contained in Section 4. A comparison of the testbed results to experimental data from humans is contained in Section 5.

[2]For surveys, see Arifovic, 2000 for applications to macroeconomic models, see LeBaron, 1999 for applications in finance, and see Dawid, 1999 for general overview of applications in economics and game theory.

[3]See Groves-Ledyard 1977.

2 The Groves-Ledyard Mechanisms[4]

We restrict our attention to environments in which a public good is produced using a constant returns to scale production function with a per unit cost of production, z. The total cost of production is equal to Xz. There are N agents, $i \in \{1, \ldots, N\}$, each of whom has a quasi-linear, quadratic utility function for the public good

$$V^i(X) = A^i X - B^i X^2 + \alpha^i.$$

A mechanism takes messages from agents and computes a level of public good and a tax payment for each agent. Let M be the set of messages. Each agent i selects an element $m^i \in M$. The total amount of public good produced is:

$$X(m) = \sum_{i=1}^{N} m^i.$$

The message m^i can be thought of as representing agent i's requested addition to the total amount of public good (given the proposed additions of other agents). Agents are free to misrepresent their requests for the public good and, if this were a voluntary mechanism, we would expect them to do so. However, the tax and allocation rules of the mechanism are specifically designed so that in Nash equilibrium it is in each agent's individual self-interest to reveal her true incremental demand for the public good. The GL mechanism uses the following tax computation:

$$T^i(m, \gamma) = (X(m)/N)z + (\gamma/2)\left[\frac{N-1}{N}\left(m^i - \mu_{-i}\right)^2 - \sigma_{-i}^2\right]$$

where T^i is the amount of tax paid by agent i, γ is an arbitrary free parameter greater than 0, $\mu_{-i} = \frac{\sum_{h \neq i} m^h}{N-1}$ is the mean value of messages of all the other agents, and $\sigma_{-i}^2 = \frac{\sum_{h \neq i}(m^h - \mu_{-i})^2}{N-2}$ is the squared deviation from this mean. The reader should notice that different values of γ imply different outcome functions and, therefore, different mechanisms.

The payoff of agent i, if the messages are m, is:

$$U^i(m) = V^i(X(m)) - T^i(m, \gamma)$$

The mechanism, $[X(m), T(m, \gamma)]$ is an incentive compatible mechanism with a balanced budget on and off the equilibrium path. It is well known that, in this environment with quasi-linear preferences, if the agents follow Nash equilibrium behavior, then the Nash equilibrium public good outcome of the one-shot game will be

[4]This section is intended mainly as a reminder to the reader of the structure of the problem. For more details, see Groves and Ledyard (1977) or Chen and Plott (1996).

the unique, Pareto optimal level of public good.[5] So, in particular, in quasi-linear environments, the equilibrium outcome level of the public good is independent of γ.

In a repeated play version of the public good allocation problem, it is assumed that the public good lasts only for 1 period. Further, payoffs are additive over time without discounting. So, at each iteration t, an amount of the public good and taxes are chosen. An agent's payoff from the sequence $(X_1, T_1, ..., X_{t'}, T_{t'})$ is

$$\begin{aligned} U^{*i} &= \sum_{t=1}^{t'} U^i(m_t) \\ &= \sum_{t=1}^{t'} V^i(X_t) - T_t^i. \end{aligned}$$

What can we say about the theory of the Groves-Ledyard mechanism in repeated play? Groves and Ledyard are themselves silent on any aspect of dynamics. But, it can be shown, at least for agents following adaptive strategies, that γ is important for the dynamic performance of the mechanism. Chen and Tang (1998) derive a sufficient condition for the convergence of the mechanism in repeated play in which agents play best responses given the messages of the other agents.[6] If agents use best responses in a sequence of repeated stage GL mechanisms, messages will converge to Nash equilibrium if agents' strategies are strategic complements; i.e., if the stage game

$$\partial^2 U^i / \partial m^i \partial m^j \geq 0.$$

This is true for quadratic preferences iff $\gamma \geq 2NB^i$ for all i . Thus, the strategic complementarity condition is satisfied for a sufficiently high value of γ.[7]

Can experiments shed any light on the properties of the Groves-Ledyard mechanism in repeated play? In experiments with human subjects (Chen and Plott, 1996; Chen and Tang, 1998), messages did not converge to Nash equilibrium for a low value of $\gamma = 1$ in 100 experimental periods, but did converge for a high value of $\gamma = 100$ (Chen and Tang, 1998). It was, however, not obvious what behavioral responses the agents were actually using in these experiments or whether the cut point of $\gamma \geq \max\{2NB^i\}$, predicted by strategic complementarity, was a determining factor in the dynamics.

These interesting, although limited, results leave open questions about the impact of different values of γ on the dynamics of the Groves-Ledyard mechanism in the experimental environment. Most theoretical results, such as the theory of supermodular games, provide predictions about convergence in the limit but do not make predictions about the speed of convergence. We are interested in the more precise

[5]In more general environments, there can be multiple Pareto-optimal allocations but the Nash equilibria of the Groves-Ledyard mechanism will select one of these.

[6]See also, Muench and Walker 1983 for further analysis of these dynamics.

[7]For the set of the parameter values in Chen and Tang 1998, this condition holds for the values of γ greater than 80.

predictions. Does the time that it takes to converge to equilibrium monotonically decrease with an increase in γ? Or is there a value of γ that results in the fastest convergence? Does the mechanism with $\gamma = 1$ ever converge to Nash equilibrium and, if not, is there some other basin of attraction? We will address these questions in the context of our testbed.

3 The Testbed

In a repeated game, in every round, each agent must select a message from M , given the mechanism and given history. Our testbed is created by implementing a particular behavioral model for that selection.[8] In each round, agents will send messages to the mechanism based on random selection from a set; that is, they use a mixed strategy. The mechanism will pick outcomes and inform the agents about them. It will also (as part of the mechanism design) provide other information. The agents will then adjust both the set from which they are selecting and the probability density that determines their selection. The testbed is driven, stochastically, over time by the sequence of mixed strategies and outcomes. Since the updating process takes place for each of the N agents, the mixed strategy of each agent co-evolves, through the mechanism, with the mixed strategies of the other agents.

3.1 Agent behavior

At the beginning of round $t \in \{1, 2, ..., T_{\max}\}$, each agent $i \in [1, \ldots, N]$ has a collection A_t^i of possible alternative messages at time t. A_t^i consists of J alternatives[9] where $a_{j,t}^i \in M$, for $j \in \{1, \ldots, J\}$. At each t, an agent selects an alternative randomly from A_t^i using a probability density[10] π_t^i on A_t^i. This alternative is her message m_t^i to the mechanism. We construct the initial set A_1^i by randomly selecting, with replacement, J messages from the set of all possible messages. We construct the initial probability π_1^i by letting $\pi_1^i(a_{j,1}^i) = 1/J$.

After receiving m_t^i from each i, the taxes and level of public good are determined by the mechanism using the Groves-Ledyard outcome function, $g(m_t)$. Agents are then informed about X_t and T_t^i. They are also informed about the value of $s_{t+1}^i = (\mu_{-i,t}, \sigma_{-i,t}^2)$ as defined in Section 2. Using the information in s_{t+1}^i, each agent computes a new A_{t+1}^i and π_{t+1}^i. This computation is the heart of our behavioral model and consists of three pieces: foregone utility, experimentation, and replication.

[8]In this section, we describe our "baseline" model. Later, we will consider a number of variations in our baseline model in order to test its robustness.

[9]J is a free parameter of the behavioral model that can be varied in the simulations. It can be loosely thought of as a measure of the processing and/or memory capacity of the agent.

[10]In essence the pair (A_t^i, π_t^i) is a mixed strategy for i at t.

3.1.1 Foregone utility

In updating A_t^i and π_t^i, the first step is to calculate what we call *foregone* utilities for each alternative in the set. This is the (expected) payoff, given the signal s_t^i, that the alternative $a_{j,t}^i$ would have received if it had been actually used, taking the behavior of other agents as given.[11] We use the notation $U^i(a_j^i|s_t^i)$ to represent this utility. Given, s_t^i an agent can compute

$$X(a_j^i, \mu_t^i) = a_j^i + (N-1)\mu_t^i$$

for each alternative $a_j^i \in A_t^i$. Then they can compute

$$\begin{aligned} U^i(a_j^i|s_t^i) &= V_i(X(a_j^i, \mu_t^i)) - T_i(a_j^i, \mu_t^i, \sigma_t^i) \\ &= V_i(X(a_j^i, \mu_t^i)) - X(a_j^i, \mu_t^i)z - (\gamma/2)\left(\left(\frac{N-1}{N}\right)\left(a_j^i - \mu_t^i\right)^2 - \sigma_t^{i2}\right) \end{aligned}$$

This is i's foregone utility for a_j^i.

3.1.2 Updating A_t^i

We modify A_t^i with processes of experimentation and replication.

Experimentation We first modify A_t^i as follows. For each $j = 1, ..., J$, with probability ρ we select one message at random from M and replace $a_{j,t}^i$ with that message. For our baseline simulations we use a normal density for this experimentation and a rate of experimentation $\rho = 0.033$. For each j, the mean value of the distribution is set equal to the value of the alternative, $a_{j,t}^i$ that is to be replaced by a 'new' idea. The standard deviation is set to 1.

This, apparently random, experimentation introduces new alternatives that otherwise might not ever have a chance to be tried. This insures that a certain amount of diversity is maintained. But, this experimentation is not as random as it looks. While it is true that an alternative is selected at random from M, we will see that the alternative selected must have a reasonably high foregone utility relative to the last period or future periods to have any chance of ever being used. A newly generated alternative has to increase in frequency in order to increase its selection probability. This can happen only if it proves successful over several periods.

There are at least two possible interpretations of our experimentation process. One is that it is a *trembling hand mistake* and the other is that it is *purposeful experimentation* intended to improve an agent's payoff. We feel the latter interpretation is most appropriate because a choice generated through experimentation is implemented

[11] This is an entirely retrospective and myopic view of the situation an agent faces. At this stage in the development of our testbed, we have decided to refrain from including expectations formation, complex intertemporal strategies (e.g., grim triggers, tit-for-tat, etc.), and other complexities introduced by repeated play. We intend to address these issues in our future research.

only if it demonstrates a potential for bringing a higher payoff. Thus, we call this method *directed experimentation* since only those newly generated alternatives that appear promising are actually tried out.[12]

Replication After experimentation, using the foregone utility computations, we construct A^i_{t+1} to reinforce messages that would have been good choices in previous rounds. We allow potentially better paying alternatives (using their foregone payoffs at t) to replace those that might pay less. For $j = 1, \ldots, J$, we let $a^i_{j,t+1}$ be chosen as follows. Pick two members of A^i_t randomly (with uniform probability) with replacement. Let these be $a^i_{k,t}$ and $a^i_{l,t}$. Then

$$a^i_{j,t+1} = \left\{ \begin{array}{c} a^i_{k,t} \\ a^i_{l,t} \end{array} \right\} if \left\{ \begin{array}{c} U(a^i_{k,t}|s_t) \geq U(a^i_{l,t}|s_t) \\ U(a^i_{k,t}|s_t) < U(a^i_{l,t}|s_t) \end{array} \right\}$$

Replication for $t+1$ favors alternatives with a lot of replicates at t and alternatives that would have paid well at t if they had been used. So it is a process with a form of averaging over past periods - if the actual messages of others have provided a favorable situation for an alternative $a^i_{j,t}$ on average then that alternative will tend to accumulate replicates in A^i_t, (it is fondly remembered), and thus will be more likely to be actually used in the mechanism. Over time, the sets A^i_t become more homogeneous as most alternatives become replicates of the best performing alternative.

3.1.3 Updating π^i_t

Given A^i_{t+1}, we now update the selection probabilities. Let

$$\pi^i_{k,t+1} = \frac{U^i(a^i_{k,t+1}|s_t) + \varepsilon^i_{t+1}}{\sum_{j=1}^J (U^i(a^i_{j,t+1}|s_t) + \varepsilon^i_{t+1})}$$

for all $i \in \{1, \ldots, N\}$ and $k \in \{1, ..., J\}$ and where

$$\varepsilon^i_{t+1} = a^\epsilon \in A^i_{t+1} min\{0, U(a^\epsilon|s_t)\}.$$

This implies that if there are negative foregone payoffs in a set, payoffs are normalized by adding a constant to each payoff that is, in absolute value, equal to the lowest payoff in the set, $U(a^\epsilon|s_t)$.

We now have completed our model. We have A^i_{t+1}, and π^i_{t+1}.

3.2 Some remarks

3.2.1 Free parameters

Our model is not entirely determined a priori. There remain several *free parameters.* These are: (1) J - the size of A (a measure of cognitive capacity) and its initial seeding,

[12]We discuss this in much greater detail below in 3.2.3

(2) ρ - the rate of experimentation and the random process that is used to generate new values during experimentation, (3) the way replication is performed, and (4) the way in which the selection probabilities are calculated as well as their initial values. We will vary all of them in order to examine the robustness of our results to changes in the model's specification. We defer further discussion until Section 4.3.3.

3.2.2 Large strategy sets

Our model shares some common features with other learning models in the literature. For example, like the Experience Weighted Attraction Model, EWA, (Camerer and Ho, 1999), the probabilities that particular messages will be actually selected are based on their hypothetical (foregone) payoffs. Also, the choice of a player's actual message is probabilistic. However, there are important differences. For example, unlike other models studied in the literature, our model is well suited to handle large strategy spaces. For example, the Groves-Ledyard mechanism which uses a continuum of possible messages for each agent. In order to apply other models of learning, the continuum must be discretized as, for example, in Chen and Tang (1998). However, discretization causes problems when there are very fine differences in equilibrium values between different mechanisms. Our model handles that problem well. While it does start out with randomly chosen sets of alternatives for each agent, due to directed experimentation, there is a sufficiently high probability that any important omitted messages, such as the Nash Equilibrium messages, will be added to the set.

3.2.3 The Role of Directed Experimentation

Experimentation, and the subsequent replication, plays a very important role in our model. Because of its role, it is worthwhile to discuss the way in which experimentation affects the outcomes in greater detail.

Consider an example where, at time t, a set of an agent's alternatives is homogeneous, and represented by a single value, a^*, which will also be that agent's actual message. Thus, $a^i_j(t) = a^*$ for all j. Suppose that at the beginning of $t+1$, a new value is generated, via experimentation. Denote this value by $a^\rho_i(t+1) = b \neq a^*$. Next, its foregone payoff is evaluated. Suppose that given $\mu_{-i}(t)$ and $\sigma^2_{-i}(t)$, $U(b) > U(a^*)$. That is, b results in a higher foregone payoff relative to all the other alternatives that are represented by a^*. However, at that point, the probability of message b of actually being selected, $\pi^\rho_{i,j}(t+1)$, is quite small since the probability mass of the remaining alternatives in the set is relatively large and is given by $(J-1)U(a^*)/[U(b) + (J-1)U(a^*)]$.

In order for there to be a reasonable probability that b will be played it has to receive more replicates for the several subsequent periods. This will eventually happen if it continues to receive high foregone payoffs. Through updating of frequencies, it will then receive more and more replicates and increase its probability mass in the set. This actually can happen fairly fast. A rough, approximate calculation using expected values, based on the assumptions that $\rho = .03$ (and is uniform) and that there are about 1% of the messages of M that are better (in the sense of foregone utility) than a^*, suggests it will only take about 7 rounds before a^* is entirely replaced by better messages. So if μ_{-i} and σ_{-i} remain constant the better alternative will replace the previous best fairly rapidly. If μ_{-i} and σ_{-i} change and b no longer yields higher forgone utility, it can disappear from A_t^i before it is ever used. Before a message drawn through experimentation stands a chance of becoming an actual message, that alternative has to prove successful over a number of periods. Thus when a new alternative (idea) occurs to an agent, she evaluates it over a time span of several periods. Only if it proves successful and increases its frequency in the set, does its probability of being chosen increase.

In general, there may be periods when agents experiment a lot with their actual messages, and those when they just adhere to their choices from previous period(s). What happens depends on the payoff landscape that a player is facing which is determined by the exogenous parameters of the mechanism as well as by the actions of other players. If there is room for improvement, given the existing choices, experimentation will help in finding alternatives that result in the improvement of agent's performance and the agent will experiment more with her actual messages. On the other hand, if the payoff landscape is such that there is not much room for further improvement then there will be less experimentation with actual choices.

Experimentation of one agent might lead to more experimentation in actual choices of the others if a different response is called for. Even if there is not much room for improvement given the payoff landscape, agents might experiment more frequently if the payoff structure is such that the punishment for deviation from 'good' messages is not too severe.

4 Testbed results

In this section we report performance results from our testbed for a wide range of values of the mechanism free parameter γ. The range we consider is $\gamma \in \{1, 100\}$ in increments of 1, and $\gamma \in \{120, 1000\}$ in increments of 20. We focus on two measures of performance: time of convergence to Nash Equilibrium and the stability of convergence. In addition, we report the results of sensitivity analyses we used to examine the robustness of our testbed to changes in values of its free parameters.

4.1 The testbed parameters

The parameter values, related to the agents' utility functions and to the cost of producing a unit of public good, are precisely those used by Chen and Tang.

Parameters of the testbed environment

agent	A_i	B_i	α_i
1	26	1	200
2	104	8	10
3	38	2	160
4	82	6	40
5	60	4	100
$\sum$	310	21	500

The cost, z, of producing a unit of the public good is set to 100.

For each value of γ (and a given set of parameter values), we simulate a series of repeated one-shot games of the Groves-Ledyard mechanism. We call this simulation, a *run*, $r \in \{1, \ldots, R\}$, where $R = 10,000$. Each run is terminated 100 periods after a *convergence criterion* is fulfilled. The convergence criterion is defined in terms of how close all agents' messages are to the equilibrium messages. Our convergence criterion is fulfilled when the difference between the equilibrium value and the value of the selected message of each agent is less than or equal, in absolute terms, to 0.1; i.e., when $|m_i^a - m_i^e| \leq 0.1$ for all i.[13]

4.2 Performance measures

There are many possible measures of performance but the most important for this paper are: (1) the time of first passage through equilibrium and (2) an index of equilibrium stability.

Does convergence to the predicted Nash Equilibrium messages occur? If so, how fast? The period when the convergence criterion is fulfilled is *the time of the first passage through equilibrium*, $T_c^{\gamma,r}$ for run r and given γ. The average time of the first passage through equilibrium for R runs, $\bar{T}_c^{\gamma}$ is given by:

$$\bar{T}_c^{\gamma} = \frac{\sum_{r=1}^{R} T_c^{\gamma,r}}{R}.$$

We denote the standard deviation from this value, across the R runs, by $\sigma_{T_c^{\gamma}}$.

What happens to the sets of alternatives, A_t^i, once the first passage through equilibrium has been recorded? Tne time of first passage would not be very interesting if the agents just rushed on by and the outputs cycled around, occasionally coming back near to the equilibrium. So we want to know how *stable* is the system after the first passage? In order to answer these questions, we have created a measure called the *index of equilibrium stability* E_s^{γ}. It measures the frequencies with which equilibrium values of messages are represented in the entire sets of agents' alternatives during 100 periods after the first passage through equilibrium. It is given by:

[13]The maximum number of periods for each run was set at $t_{max} = 20,000$. If the convergence criterion is not fulfilled by that time, a run is terminated.

$$E_s^{\gamma,r} = \frac{\sum_{t=T_c^{\gamma,r}+1}^{T_c^{\gamma,r}+100} \sum_{i=1}^{N} \sum_{j=1}^{J} a_{j,t}^i E_{j,t}^i}{NJ}$$

where $E_{j,t}^i$ is an index variable such that $E_{i,j}(t) = 1$ if $a_{j,t}^i = m_i^e$ and $E_{i,j}(t) = 0$ if $a_{j,t}^i \neq m_i^e$. It can take values between 0 and 100 and represents stability of equilibrium in percentage terms. We denote the average of this index across the R runs by $\bar{E}_s^\gamma$. We denote the standard deviation from this value, across the R runs, by $\sigma_{E_S^\gamma}$. Notice that this measure takes into account the values of the messages in the whole set of each agent, not just those that are actually selected to be sent. For that reason it may underestimate the percentage of messages sent after T_c^γ that are equilibrium messages.

4.3 The testbed findings

There are three main groups of findings. (1) There is convergence to Nash equilibrium messages for all of the values of γ within the range that we simulated. Convergence is relatively fast for a much larger set of the values of γ then that predicted by the strategic complementarity condition. And, $\bar{T}_c^\gamma$, the time to first convergence, is U-shaped in γ. (2) The Nash equilibria of the model are stable in the sense that once the model first passes through the equilibrium, it remains in its neighborhood. (3) The above features characterize the dynamics of all of the different versions of the model that we examined.

4.3.1 Convergence to Nash Equilibrium

Table 2 contains detailed data for the average time of first passage through equilibrium, for $10,000$ runs.

[Table 2 about here]

The first column gives the value of γ, the second indicates the total number of runs ($10,000$ for each value of γ), the third column presents the average values of times of first passage through equilibrium (averaged over $10,000$ runs), $\bar{T}_c^\gamma$, and the values of standard deviations, $\sigma_{T_c^\gamma}$, in the parenthesis. In Figure 1, we present, graphically, the results for $\gamma \in \{1, \dots 100\}$, and in figure 2 the data for $\gamma \in \{10, 100\}$. Figure 3 shows the data for γ between 120 and 1000 in the increments of 20.

The first thing to note is that our simulations converge on average to Nash equilibrium for all the values of γ, although there is a difference in the average time of the first passage across values of γ. $\bar{T}_c^\gamma$ does not decrease monotonically with increases in γ. It has a very high value for $\gamma = 1$, and then drops to a much lower value even for $\gamma = 2$. It reaches values below 100 once $\gamma = 10$. The average times to convergence are not much different for the values of γ between 10 and 100. For example, $\bar{T}_{20}^c$ for $\gamma = 20$ is within a minus one standard deviation of $\bar{T}_{50}^c$ for $\gamma = 50$ (and also, $\bar{T}_{50}^c$ is within a plus one of its own standard deviation of $\bar{T}^c_{20}$). Similar relationships can be observed for all the values of γ up to 100. A decreasing pattern is observed until γ reaches the values in the range $\{40, \ldots, 60\}$. After this, $\bar{T}^{\gamma}_c$ starts increasing, and its behavior results in a U-shaped curve. It appears that the curve in Figure 1, plotting the behavior of $\bar{T}^{\gamma}_c$, just levels off and stays flat. However, when the scale changes, as in figure 2 where we plot the same data starting with $\gamma = 10$, a U -shaped curve emerges, showing that the minimum values of $\bar{T}^{\gamma}_c$ are reached for the range of values of γ between 40 and 60. After that, for $\gamma > 60$, $\bar{T}^{\gamma}_c$ starts increasing although, even for $\gamma = 100$, $\bar{T}^{\gamma}_c$ is below 100. As γ increases further, the average time to convergence $\bar{T}^{\gamma}_c$ starts increasing monotonically for all the values of γ up to 1000 (See Figure 3).

The standard deviation of $\bar{T}^{\gamma}_c$ decreases with the increases in γ, drops to the lowest values for $\gamma \in \{40, 60\}$, but then slightly increases for the values of $\gamma > 60$. As can be seen from Table 2, runs for $\gamma = 50$ have a lower standard deviation and less variance than the runs for $\gamma = 80$ or $\gamma = 100$.

4.3.2 Stability of equilibrium

The fourth column in table 2 gives the values of the measure that we use to study the stability of equilibrium after the first passage. This is of interest particularly because of the random, but directed, experimentation of our testbed agents. It is possible that the experimentation could create instability and drive things very far from equilibrium for very long times. But that doesn't happen. The values of this measure are above 85% for all of the values of $\gamma \in \{1, \ldots, 100\}$. It is above 90% for $\gamma \geq 10$ and equal or above 95% for $\gamma \geq 20$. This reflects a high degree of homogeneity in the sets of alternatives A^i_t after convergence, despite the fact that experimentation is present at the same exogenously given rate. Standard deviations from $\bar{E}^{\gamma}_s$ that are reported in the parentheses in the fourth column of Table 2 are generally low, but are also decreasing as γ increases. Thus our measure of the *stability* of equilibria shows system robustness despite the continuing exogenous shocks due to experimentation.

4.3.3 Sensitivity analysis

We ultimately would like to be able to use our testbed as a reliable, first-cut substitute for expensive, experimental analyses of mechanisms. In order to do so, we need the testbed not only to be accurate but also to be robust to changes in the free parameters. In this section, we examine how the performance of the mechanism changes as we change these testbed parameters. In Table 3, we list the different values of the testbed parameters that we use. Unless otherwise specified below, we keep the other baseline parameters fixed. A detailed description of our sensitivity analysis are reported in Arifovic and Ledyard (2005). Here we provide a brief overview of our robustness analysis.

[Table 3 about here]

Increases in the **size of** A^i_t, J, sped up the rate of convergence, at least for the range we have considered. For $J = 200$, we observed, on average, 1/3 increase in the speed of convergence. However, further increases, to $J = 500$, and $J = 1,000$ did not bring further significant decreases.

The **experimentation process** does have some effect on the speed of convergence. Experimentation using the uniform distribution results in higher values of $\bar{T}_c^\gamma$. The time to convergence under the historically independent experimentation is roughly twice that for the historically dependent normal experimentation. For a given distribution, increases in the rate of experimentation seem to increase the time of convergence although the effect seems small.[14]

Regarding the **replication process**, in addition to tournament replication, we tried replicating the alternatives via proportional replication. However, this method resulted in a consistently large increases in the times of convergence.

Selection We will consider two types of probabilistic selection of messages from the set A_t^i: proportional and exponential selection. Our baseline uses the proportional approach as described in Section 3.1.3. For exponential selection, one computes selection probabilities, using exponentialized payoffs,

$$\pi_{k,t+1}^i = \frac{e^{\lambda U(a_{k,t+1}^i | s_{t+1}^i)}}{\sum_{j=1}^J e^{\lambda U(a_{j,t+1}^i | s_{t+1}^i)}}.$$

for every i and j, where λ is an exogenously given parameter. A number of models of individual learning use this method to compute the probabilities because it directly maps negative foregone payoffs into positive probabilities. However, this method, does introduce another free parameter, λ. So we need to see whether changes in λ have any significant effect. We report, in Table 4, the convergence times for 4 different values of λ, for $\gamma = 1, 50$ and 100 using uniform experimentation with the rate $\rho_u = 0.033$.

[Table 4 around here]

Surprisingly, at least to us, exponentialized payoffs as a basis for selection resulted in essentially the same behavior as proportional selection with no significant differences in the values $\bar{T}_c^\gamma$. In fact, even more surprisingly, the value of the parameter λ seems to have no effect on the times of the first passage through equilibrium, or on the dynamics in general. We conducted simulations for a number of values that ranged from very small to relatively large and found no significant differences.[15] This is interesting since the performance of other learning models that use this approach to update the payoffs is very sensitive to the value of this parameter and rather different dynamics are generated as this value varies. See for example the effect on quantal response equilibria (McKelvey and Palfrey, 1995), experience weighted attraction learning (Camerer and Ho, 1999), and reinforcement learning (Roth and Erev, 1995).

[14]The results of comparison between simulation and experimental results favor experimentation from the normal distribution.

[15]For the analysis of why λ does not affect the dynamics of the system, see Arifovic, Ledyard and Tse, 2002.

5 Comparison with experimental data

The testbed would not be very interesting or useful if it did not, at least approximately, correspond to actual behavior. In this section, we compare the behavior of our model to the behavior of human subjects in several experiments using two sets of data. One set is described in Chen and Tang (1998). They conducted 7 experiment sessions each with $\gamma = 1$ and $\gamma = 100$. The second set was generated by us at the California Institute of Technology in April and May 2002. We conducted 4 experiment sessions each with $\gamma = 50$ and $\gamma = 150$.[16] Data from these, compared to our baseline data, can be found in Table 8.

[Table 8 around here]

A quick look at the two key statistics, time to equilibrium and stability, reveals there is a reasonable correspondence between our testbed and the experiments.[17]

The first statistic of interest is the average time of the first passage through equilibrium ($\bar{T}_c^\gamma$). For $\gamma = 1$, the experiments run by Chen and Tang did not converge to the equilibrium within 100 experimental periods for any of the 7 sessions that were conducted. For $\gamma = 1$, the baseline simulations from our model produced really high average convergence times that are close to or over $1,000$ periods regardless of the model's exact specification. Thus the experimental and simulation evidence are consistent with each other. For $\gamma > 1$, convergence occurred, according to the 0.1 criterion, in all of the experiment sessions except one with $\gamma = 50$. In this session, one of the subjects kept her message consistently 0.2 units below the equilibrium value of 1 for the entire session.[18] We do not include the data from this session in our subsequent analysis. In the other experimental sessions, convergence is relatively fast for all of the three values of γ. The time to convergence is less than 20 periods on average which is even faster than the baseline simulations.

The second statistic of interest is the measure of stability (E_s^γ).[19] The stability measures for the experimental sessions are all very high, averaging over 0.9 with small

[16] Our experimental design is very similar to Chen's and Tang's. However, we introduced two modifications. First, in Chen and Tang, subjects could make only integer number choices. We let the subjects make real number choices, with a two decimal points restriction. Second, we added a calculator to the windows interface that allowed the subjects to calculate their payoffs varying the size of μ_{-i} and σ_{-i}.

[17] We note that we did not make any special effort to calibrate the free parameter values of our model to this particular situation.

[18] There might be two reasons why this would happen. The subject might be trying to manipulate the others in the manner of a Stackelberg equilibrium. Or, the subject is just lazy or confused. If the first were true then the subject should be making a higher payoff than subjects equally situated in other sessions. Those subjects made 229.44 and 230.41 experimental francs on average in each of their periods. The subject who stayed .2 units away from the equilibrium made only 225.27 experimental francs on average in their periods. So we infer from this that if they were trying to be strategic in some unknown way it didn't work. But then they would have ultimately given up and adjusted. Since they didn't we will not use this session in our data analysis.

[19] Because A_t^i is not observed in our experiments, we need to use a slightly different measure than

standard deviations. These numbers correspond to the numbers we reported in Table 2.

But we need more than a quick look at the statistics. We should really subject them to a more rigorous test. We turn to that now.

5.1 The statistical analysis

The question we address in this section is, could the laboratory observations from the experiments have come from a population generated by our testbed? In particular, we ask whether the mean and standard deviation of $\bar{T}_c^\gamma$, the average time to convergence, from the experiments could have been generated by a particular testbed with high probability. This is a standard statistical hypothesis test using, respectively, a Student's t-test for the mean and a Chi-squared test for the variance. We provide two caveats up front. First, our experimental sample sizes are really small so the confidence intervals are pretty large. Second, the population distributions (those generated from the testbed) are not symmetric. Together these suggest one should be very careful in dealing with our statistical evidence. Nevertheless, we believe these statistics do provide important information.

We proceed as follows. We first select values for the two basic free variables for our testbed: J, the number of strategies tracked by an agent, and ρ, the rate of experimentation by an agent. Then for some γ, we complete 1000 runs of the simulation and, from those we compute the mean and variance of the average time to completion. We assume these are the true mean and variance for the population generated by that mechanism in that model. We can then think of an experimental trial as a draw from some population and ask whether the means and variances of the sample experimental data could have been generated by the population of our. We can compute confidence intervals into which the experimentally generated means and variances should fall with some probability if that data came from the model population. For the means this is a two-tailed t-test and for the variances this is a two-tailed Chi-squared test. Given (J, ρ), if the mean and variance of the sample values lie outside these intervals for some γ, then we can say, with 95% confidence that the behavior exhibited in the experiment is not the same as the behavior predicted by (J, ρ) or, put another way, there is a 95% chance that the model (J, ρ) is not consistent with the experimental evidence.

We did not run through all possible values of (J, ρ) to find what is best. We did look at the parameter values of $J = 50, 100, 200$, $\rho = 0.033, 0.25$, and $\gamma = 50, 100$. We also considered experimentation using a uniform density and a normal density. For every variation we considered, there is at least one mechanism γ for which at least one of the statistics from the experiments lies outside its 95% confidence region.[20] A

we do in the testbed simulations. For the experimental data, we take the remaining number of periods of a particular session once the first passage through equilibrium is achieved, about 80 periods on average, and compute the the percentage of actual messages that are equilibrium messages.

[20]If we raise the bar for rejection to 99%, then there is one set of parameter values for which no

summary of those facts is given in Table 9. If an entry reads "too high" that means that the sample statistic, based on the experiments, lies below the 95% confidence interval based on that model. So if the model mean is "too high", that means that the simulations do not converge as fast as the experiments. And if the model variance is "too low", that means that the simulations do not show as much variation in "time to equilibrium" as do the experiments.[21]

[Table 9 around here]

If we look first at the mean "time to equilibrium", we see that there is only one model that is not rejected for at least one of the γ. This is $\rho = 0.033$, normal density, and $J = 200$. We also see that when a rejection of a model occurs it is always because that model is too slow. There is a natural explanation for this.[22] Humans subjects know more at the beginning of an experiment than do our models. In particular, humans know their payoff function.and can do some preliminary culling of bad strategies and identification of good strategies before even playing - a type of one round elimination of dominated strategies. But our model initially seeds the A^0 sets randomly with J draws. That is, all strategies are treated equally whether dominated or not. This probably gives humans a headstart over our model, but that can presumably be changed with a slightly different initialization process for A^0.

Looking at the variance of the "time to equilibrium", we can see a significant difference between the models that use a normal density for experimentation and those that use the uniform density. Those that use the normal are almost never rejected on the grounds of the variance. However, those that use the uniform are almost always rejected. It is possible that a new initialization process could lower rejections of the uniform since it would lower any variance due to the initial seeding of A^0.

Clearly, more data and a better initialization procedure are both called for in future research.

6 Conclusion

We are ultimately interested in constructing a model that can capture the significant qualitative features of experimental behavior, and can also be used to make predictions for a wide variety of mechanisms. If the model calibrates well with experimental data, it can then be used to predict the outcomes in the absence of experimental data. If the model is robust it can do this for many different mechanisms. Measurements from the testbed can then be used as a guide to what type of experiments should be conducted. This is desirable because conducting experiments is much costlier in terms

rejection occurs for either $\gamma = 50$ or 100. That is the case in which $J = 100$, $\rho = .25$, and the normal density is used for experimentation.

[21] All of the numbers for these various means, variances, and confidence intervals can be found in Arifovic and Ledyard (2005).

[22] The following remarks have significantly benefitted from the report of a referee.

of time and monetary resources required for their implementation. In this paper we have seen that our testbed is both accurate and robust for the mechanisms studied.

Our testbed appears to be accurate at a minimum in its predictions of average times of first passage to Nash Equilibrium. In the context of the Groves-Ledyard mechanism, we have seen that, in our testbed, messages converge to Nash Equilibrium for all values of the free parameter γ between 1 and 100 and, as well, for all values between 120 and 1000 in increments of 20. Very low values, between 1 and 10, and very high values between 200 and 1000, take very long time to converge. Any value of γ greater than 10 and less than 200 results in relatively fast convergence. These predictions of our model have been confirmed in experimental sessions with $\gamma = 50$, 100 and 150. In addition, the experimental data indicate that the mechanism $\gamma = 50$ results in faster convergence than the mechanism $\gamma = 100$ or 150 which is also a prediction of our model.

One interesting implication of these findings is that strategic complementarity is not a necessary condition for convergence in the lab or in our testbed. (Remember for the utility functions we are using, strategic complementarity obtains for all $\gamma \geq 80$.) Further, strategic complementarity can not be used as a guide to the rate of convergence. This is a bit surprising since our behavioral model has a significant element of best reply in its formulation.

The performance of our testbed is not entirely independent of variations in its "free parameters". Given the experimental data, it appears the testbed model with tournament replication, historically dependent normal experimentation and proportional selection is the right one. There does seem to be sensitivity to the size of the memory, J, and the rate of experimentation, ρ. Higher J and higher ρ both appear to increase the average time to first passage.[23] Our testbed is, however, statistically consistent with the experimental data for both $(J, \rho) = (50, 0.25)$ and for $(J, \rho) = (100, 0.033)$. We intend to explore this relationship in our future work.

Directed experimentation is very important for our model's dynamics. It is worthwhile to point out that this is different from *experimentation* or *mutation* traditionally discussed in the literature on learning or evolutionary game theory. Directed experimentation gives our model the ability to quickly adjust to changes in the environment if changes in their actions are called for. This will happen even after the sets have converged to a single (equilibrium) value and remained there for a long time. This ability to adjust to changes in the environment (shifts in regime) has not been demonstrated in other models of individual behavior studied in the literature. This is an important issue that has not been given much attention in the studies of models of learning. However, it should be addressed in greater detail in future research.

[23] It is our current conjecture that both J and ρ can be large enough to cause problems for convergence. But, we do not yet have the simulations to support this.

References

Arifovic, J. (1994) Genetic Algorithm Learning and the Cobweb Model. *Journal of Economic Dynamics and Control* 18: 3-28.

Arifovic, J. (2000) Evolutionary Algorithms in Macroeconomic Models. *Macroeconomic Dynamics* 4: 373-414.

Arifovic, J. and J. Ledyard (2004), Scaling Up Learning Models in Public Good Games, Journal of Public Economic Theory ,6,2004, 205-238.

Arifovic, J. and J. Ledyard (2005) Computer Testbeds: The Dynamics of Groves-Ledyard Mechanisms, Working Paper, California Institute of Technology.

Camerer, C. and T. Ho (1999) Experience Weighted Attraction Learning in Normal Form Games. *Econometrica* 67: 827-873.

Chen, Y. and C. Plott (1996) The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design. *Journal of Public Economics* 59: 335-364.

Chen, Y. and F. Tang (1998) Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study. *Journal of Political Economy* 106: 633-662.

Chen, Y. and R. Gazzale (2004)When Does Learning in Games Generate Convergence to Nash Equilibria? The Role of Supermodularity in an Experimental Setting, American Economic Review, 94, 1505-1535.

Dawid, H. (1999) *Adaptive Learning by Genetic Algorithms, Analytical Results and Applications to Economic Models*, 2nd ed. Berlin: Springer-Verlag.

Groves, T. and J. Ledyard (1977) Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem. *Econometrica*, 45: 783-809.

Hurwicz, L. (1960) Optimality and Efficiency in Resource Allocation Processes. In K.J. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods in the Social Sciences*, pp. 27-48. Stanford, CA: Stanford University Press.

Ledyard, J. (1995) Public Goods: A Survey of Experimental Research. In J. Kagel and A. Roth (eds.) *Handbook of Experimental Economics*, pp. 111-194. Princeton: Princeton University Press.

LeBaron, B. (2000) Agent Based Computational Finance: Suggested Readings and Early Research. *Journal of Economic Dynamics and Control* 24: 679-702.

Marimon, R., E. McGrattan and T.J. Sargent (1989) Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents. *Journal of Economic Dynamics and Control* 14: 329-373.

McKelvey R.D, and T.R. Palfrey (1995) Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10: 6-38.

Muench, T. and M. Walker (1983) Are Groves-Ledyard Equilibria Attainable? *The Review of Economic Studies,* 50: 393-396.

Roth, A. and I. Erev (1995) Learning in Extensive Games: Experimental Data and Simple Dynamic Model in the Intermediate Term *Games and Economic Behavior* 8: 164-212.

Table 2 - A set of baseline runs

γ	R	$\bar{T}^c_\gamma$ ($\sigma_{T^c_\gamma}$)	E^γ_s ($\sigma_{E^\gamma_s}$)
1	10000	903.38 (273.97)	85.13 (8.41)
10	10000	34.81 (15.98)	94.12 (2.51)
20	10000	17.18 (7.61)	94.95 (1.59)
30	10000	14.61 (6.17)	95.21 (1.35)
40	10000	13.67 (5.65)	95.29 (1.22)
50	10000	13.48 (5.76)	95.31 (1.21)
60	10000	13.79 (5.98)	95.28 (1.29)
70	10000	14.48 (6.4)	95.24 (1.33)
80	10000	15.50 (7.15)	95.18 (1.43)
90	10000	17.42 (8.68)	95.12 (1.50)
100	10000	19.65 (10.52)	95.00 (1.74)

Table 3

Parameters of the behavioral model

γ	J	replication
$\{1, 100\}$	50	tournament
$\{120, 1000\}$	100	proportional
	200	
	1,000	

experimentation	selection of $m^a_i(t)$	initial A^i_t, π^i_t
$\rho_u = 0.033$	proportional	random
$\rho_u = 0.25$	exponential	from prior equilibrium
$\rho_n = 0.033$		

Table 7

First passage through equilibrium for different values of λ
with $\rho_u = .033$.

λ $\quad\gamma$	1	50	100
0.0001	2924.57 (1065.42)	27.07 (22.83)	59.23 (40.07)
0.006	2621.51 (991.68)	29.44 (25.15)	43.40 (31.62)
1	3363.33 (1344.17)	28.00 (21.14)	38.27 (31.03)
3	2775.08 (1086.73)	30.60 (22.16)	52.01 (36.10)
proportional	2556.53 (999.18)	28.59 (23.56)	46.28 (35.39)

Table 8

Convergence Times and Stability of Equilibria
Comparison of experimental and simulated data

	experimental		simulated	
γ	$\bar{T}^c_\gamma$ $(\sigma_{T^c_\gamma})$	E^γ_s $(\sigma_{E^\gamma_s})$	$\bar{T}^c_\gamma$ $(\sigma_{T^c_\gamma})$	E^γ_s $(\sigma_{E^\gamma_s})$
50	5.75 (4.42)	98.00 (1.00)	13.48 (5.76)	95.31 (1.21)
100	18.86 (12.034)	99.00 (1.00)	19.65 (10.52)	95.00 (1.74)
150	20.00 (17.20)	92.00 (4.00)	38.89 (22.81)	94.72 (2.09)

Table 9

Statistical Test of the Models Compared to Experimental Data

Model Parameters	Mechanism	Model Mean	Model Variance
$J = 50, \rho = .033, normal$			
	$\gamma = 50$	too high	
	$\gamma = 100$		
$J = 100, \rho = .033, normal$			
	$\gamma = 50$	too high	
	$\gamma = 100$		
$J = 200, \rho = .033, normal$			
	$\gamma = 50$		
	$\gamma = 100$		too low
$J = 100, \rho = .25, normal$			
	$\gamma = 50$	too high	
	$\gamma = 100$		
$J = 100, \rho = .033, uniform$			
	$\gamma = 50$	too high	
	$\gamma = 100$	too high	too high
$J = 100, \rho = .25, uniform$			
	$\gamma = 50$	too high	
	$\gamma = 100$	too high	too high