A Behavioral Model for Mechanism Design: IEL

Jasmina Arifovic and John Ledyard*

November 25, 2007

Abstract

We describe a new and different behavioral model for individuals playing in a repeated situation. It is based on a flexible learning process and called Individual Evolutionary Learning. IEL does not require calibration and can be used as a computer testbed to study the probable performance of a wide range of mechanisms over a wide range of environments prior to testing them in a laboratory or using them in practice. We illustrate the utility of the testbed approach by analyzing an open question in mechanism design - the dynamics of Groves-Ledyard mechanisms. Contrary to standard theories, the prediction from the IEL behavioral model is that the average time to convergence varies smoothly and is U-shaped in the mechanism's free parameter. We validate the results from the testbed with data from economic experiments with humans.

*Simon Fraser University and California Institute of Technology, respectively. We thank Stephanie Chow, Paul Healy, and Stephen Tse for their very able research assistance. We thank Peter Bossaerts, Colin Camerer, Yan Chen, Tom Palfrey and Charlie Plott for helpful discussions. We thank participants at 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 ITR/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 information and incentive constraints. 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.

However, mechanism design applications remain difficult and not at all straightforward. An important question in applied work 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 alternative mechanisms. To do that one must 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 from the theoretical literature are the use of dominant strategies (if they exist), Bayesian equilibrium, and Nash equilibrium. Among the candidates from the behavioral literature are reinforcement learning, belief-based learning, and their hybrid, experience weighted attraction. There are many others.

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 use to identify the few mechanisms and environments that should be tested in the lab. The challenge is to provide computer models of humans with all their cognitive capabilities. These must be equal to, not better than, humans.

A natural way to start would be to take an existing model of learning in experiments, such as Reinforcement Learning (Erev and Roth) or EWA (Camerer and Ho), and use that as a basis for the testbed. Unfortunately there are at least two problems with such an approach: (1) these models must be tuned or calibrated; that is, parameters must be set differently for different games or market conditions in a way that is not necessarily known a priori, and , as we show in Appendix I, (2) these models do not scale; that is, when there are a large number of strategies available to each agent, the models behave less and less like human subjects do.

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, 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 human subjects. Robust means that the computer modeled behavior has to be accurate without re-estimation of the parameters.

In this paper, we propose a new model of learning, called Individual Evolutionary Learning (IEL), that is based, to some extent, on evolutionary algorithms such as genetic algorithms, classifier systems, genetic programming, evolutionary programming, etc. We then build a computer testbed based on IEL and use that to analyze an standing problem in mechanism design.

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.[1] We choose this example for three reasons. First, the class of mechanisms is well-known and reasonably straight-forward to describe. Second, there is an accessible, if incomplete, body of theory about this class with respect to both the static and dynamic properties of the mechanism. Third, there is a body of experimental work to draw on.

Theory establishes for the Groves-Ledyard 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 rely on Cournot, or best reply, strategies for their analysis. Chen and Tang 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 experiments with human subjects does seem to 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).

When we used our IEL testbed to analyze the Groves-Ledyard mechanisms, we found convergence for all values of the parameter tested. We also found that the time to converge is smooth and U-shaped in the parameter; that is, there is a value of the parameter that minimizes the time to converge and, also, the efficiency of the allocations chosen. This is significantly different from predictions based on strategic complements or dominant diagonal. Based on our testbed analysis, we ran several experiments for 6 different values of the mechanism parameter. Remarkably, those data confirm the U-shape and the parameter that minimizes the convergence time. But we also found that human subjects are a bit faster, a little less stable, and a little less efficient than the IEL simulations would suggest, although those differences are smallest at the value of the mechanism parameter that minimizes average time to converge.

We turn now to filling in the details. We begin in Section 2 by describing the GL mechanisms and the questions we are interested in. In Section 3 we describe IEL and the computer testbed. In Section 4, we detail the results of our simulations with the GL mechanisms in that testbed. We compare the testbed results to experimental data in Section 5. We look at variations in our model and testbed to determine the robustness of our

[1] See Groves-Ledyard 1977.

approach in Section 6. In Section 7 we finish off with some final thoughts.

2 A Mechanism Design Question

The Groves-Ledyard mechanisms were designed to achieve an optimal allocation and financing of public goods in one-shot situations. We are interested in how these mechanisms perform when used repeatedly, in a sequence of one-shot games.

2.1 Environments and Mechanisms[2]

Public Good Environments We restrict our attention to environments in which a public good is produced using a constant returns to scale production function and in which agents have quadratic, quasi-linear utility functions.[3] There are N agents, $i \in \{1, \ldots, N\}$. Let $z = (X, y^1, ..., y^N)$ where X is the amount of the public good and y^i is i's net consumption of the private good. Each i begins with an initial endowment of the private good α^i. Agent i's consumption of the private good is $\alpha^i + y^i$. The per unit cost of production of the public good is c. The set of feasible allocations is

$$Z = \left\{ z = (X, y) | X \geq 0, cX + \sum_{i=1}^{N} y^i = 0 \right\}$$

Each consumer i has a utility function: $V^i(X, \alpha^i + y^i) = A^i X - B^i X^2 + \alpha^i + y^i$.

In the repeated version of the public good allocation problem, payoffs are additive over time without discounting. At each iteration $t = 1, ..., t'$, amounts of the public good and net consumptions $(X_t, y_t) \in Z$ are chosen. An agent's payoff from the sequence $(X_1, y_1, ..., X_{t'}, y_{t'})$ is $\sum_{t=1}^{t'} V^i(X_t, \alpha^i + y_t^i)$.

We use the notation (N, Z, V) to denote an environment.

The Groves-Ledyard Mechanisms The Groves-Ledyard mechanisms are defined as follows for a one-shot game. Begin with a language, a space of messages, $M = (-\infty, +\infty)$. $m^i \in M$ can be thought of as i's incremental demand for the public good.

Given a vector of messages $m = (m^1, ..., m^N)$, the public good produced is

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

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

[3] The Groves-Ledyard mechanisms also work in more general environments including those with income effects. We restrict our attention in this paper to quadratic, quasi-linear environments to allow comparisons with existing experimental results.

The tax to be paid by i is:

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

where 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 the messages of the other agents, and $\sigma^2_{-i} = \frac{\sum_{h\neq i}(m^h - \mu_{-i})^2}{N-2}$ is the squared deviation from this mean.

The outcome rule of the mechanism is $g(m,\gamma) = (X(m), y^1(m,\gamma), ...y^N(m,\gamma)) = (X(m), -T^1(m,\gamma)..., -$ Let $r^i = r^i(m) = (\mu_{-i}(m), \sigma^2_{-i}(m))$. Because r^i does not depend on m^i and $X(m) = m^i + (N-1)\mu_{-i}$, $i's$ utility can be written as

$$W^i(m^i|r^i(m),\gamma) = V^i(m^i + (N-1)r^i_1(m), \alpha^i - T^i(m^i, r^i(m),\gamma)). \quad (1)$$

Different values of γ imply different outcome functions and, therefore, different mechanisms. So by letting γ range over values from 0 to ∞, one creates an entire class of mechanisms.

2.2 Prior Results and New Questions

The theoretical equilibrium properties of Groves-Ledyard mechanisms are well understood. The tax and allocation rules are specifically designed so that, if the agents follow Nash equilibrium behavior, then the equilibrium outcome of a one-shot game will be a Pareto optimal allocation. Formally, if m^* is a vector of messages such that $m^{*i} \in$ argmax $W^i(m^i|r^i(m^*),\gamma)$ for all i, then $[X(m^*), y(m^*)] \in$ argmax $\sum V^i(X, y^i)$ subject to $cX + \sum y^i = 0$. In environments with quasi-linear preferences, the Pareto optimal level of public good is unique and the equilibrium outcome level of the public good is independent of γ. But if one is interested in actually using these mechanisms, it is necessary to understand their dymamics. For example, if one is interested in the ability of the mechanisms to attain optimal levels of utility in repeated situations, then one must ask whether and how fast individuals will converge to the Nash equilibrium, since faster convergence implies higher aggregate welfare.

Theory is mostly silent on the dynamics of Groves-Ledyard mechanisms. Two exceptions are the papers by Muench and Walker (1983) and by Chen and Tang (1998), both of which suggest that the parameter γ plays a major role in those dynamics for agents following adaptive strategies. Based on the work of Milgrom and Roberts (1990) on strategic complementarities, Chen and Tang (1998) derive a sufficient condition for quadratic preferences, $\gamma \geq 2NB^i$ for all i, for the convergence of the mechanism in a sequence of repeated one-shot games if agents use adaptive learning.[4] Another sufficient condition for global convergence to Nash equilibrium, if agents use best replies, can be derived from a theorem of Gabay

[4] Adaptive learning is defined in Milgrom and Roberts (1990) and includes Cournot best response, fictitious play, Bayesian Learning and others. The sufficient condition for convergence under adaptive learning is $\partial^2 V^i / \partial m^i \partial m^j \geq 0$

and Moulin (1980) using a dominant diagonal condition.[5] For quadratic preferences that condition holds if $\gamma \geq [N(N-2)/(N-1)]B^i$ for all i. However, neither the strategic complementarity nor the dominant diagonal condition provide any insight into how the speed of convergence might depend on γ. Such knowledge is particularly important for practical implementations.

Experiments with human subjects also do not shed very much light on the dynamic properties of the Groves-Ledyard mechanism. The key results are reported by Chen in Chen and Plott (1996)and Chen and Tang (1998). In their experiments, messages did not converge in 100 experimental periods to Nash equilibrium for a low value of $\gamma = 1$, but did converge for a high value of $\gamma = 100$ (Chen and Tang, 1998). It is, however, not possible to determine from these experimental results whether the cut points, $\gamma \geq \max\{2NB^i\}$ provided by strategic complementarity or $\gamma \geq \max\{[N(N-2)/(N-1)]B^i\}$ provided by dominant diagonals, were a determining factor in the dynamics.

These interesting, but limited, prior results do not appear to answer the more important questions about the impact of different values of γ on the dynamics of the Groves-Ledyard mechanism when implemented with humans. Does the mechanism with $\gamma = 1$ ever converge to Nash equilibrium? Is there a value of γ such that for smaller values, the mechanism does not converge to the Nash equilibrium and for larger values it does? Is convergence related to either the dominant diagonal condition or the strategic complements condition? Does the time that it takes to converge to equilibrium vary systematically with an increase in γ? Is there a value of γ that results in the fastest convergence?

We will address these questions using a computer testbed based on a new flexible behavioral model.

3 The IEL Computer Testbed

Our approach to modeling behavior is based, to some extent, on evolutionary algorithms[6] such as genetic algorithms, classifier systems, genetic programming, evolutionary programming, etc. A large number of applications using these algorithms have focused on models of social learning where a population of agents (each agent is represented by a single strategy) evolves over a long period of time such that the entire population of agents jointly implements a behavioral algorithm. However, in some applications (see Arifovic, 2000 for a survey), 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 this approach and build on the framework of Arifovic (1994).

We describe a very simple, synchronous computer testbed that is sufficient to provide answers to the mechanism design questions we are interested in. One begins with an environ-

[5] We thank Paul Healy for the Gabay-Moulin reference. See Healy (2005) for a use of the theorem in the context of public good mechanism design.

[6] 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.

ment (Z, V, X) and a mechanism (M, g, r). The action then occurs in a sequence of rounds numbered $t = 1, ..., t'$. In every round, each agent[7] selects a message m_t^i randomly by using a mixed strategy, $\pi_t^i(m^i)$, a probability distribution on M^i. The mechanism then determines an outcome, using $g(m_t^1, ..., m_t^N)$, and informs the agents. Each agent learns $X(m_t), T^i(m_t)$, and the signal $r_t^i = r^i(m_t)$. Each agent then computes a new mixed strategy, $\pi_{t+1}^i(m^i)$ and a new round begins.

Over time a sequence is generated: $(\pi_1^i \to m_1) \to (z_1, u_1) \to (\pi_2^i \to m_2) \to ... \to (\pi_{t'}^i \to m_{t'}) \to (z_{t'}, u_{t'})$. The mixed strategy of each agent co-evolves, through the mechanism, with the mixed strategies of the other agents. It only remains to describe the agents' behavior - how they choose their mixed strategies π_t^i on M^i. We turn to that now.

3.1 Individual Evolutionary Learning

The primary variables of our behavioral model are a finite set of actions (messages) for each agent i at each round t, A_t^i, and a probability measure π_t^i on A_t^i. A_t^i consists of J alternatives[8] where $a_{j,t}^i \in M$, for $j \in \{1, \ldots, J\}$. In each t, each agent selects an alternative randomly from A_t^i using the probability density π_t^i on A_t^i and sends that message m_t^i to the mechanism. One can think of (A_t^i, π_t^i) as a mixed strategy.

In each round, after receiving m_t^i from each i, the outcome $g(m_t)$ is determined by the mechanism.[9] Agents are then informed about $X_t(m_t), y_t^i(m_t)$ and $r^i(m_t)$. This ends round t. At the beginning the next round $t+1$, each agent computes a new A_{t+1}^i and π_{t+1}^i. This computation is at the heart of our behavioral model and consists of two pieces: experimentation, and replication.

Experimentation comes first. 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. For each $j = 1, ..., J$, with probability[10] ρ, a new message is selected at random from M and replaces $a_{j,t}^i$. We use a normal density for this experimentation. 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.

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 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.

[7] This is a computer agent and not a human agent.

[8] 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.

[9] We describe how the first period message choice are made by the agents until the end of this section.

[10] ρ is a free parameter of the behavioral model that can be varied in the simulations.

It is worthwhile to point out that this is different from the experimentation or mutation traditionally discussed in the literature on learning or evolutionary game theory.

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.

Directed experimentation is not as random as it may look. While it is true that an alternative is selected at random from M, the alternative selected must have a reasonably high 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.[11]

Replication comes next. Replication reinforces strategies that would have been good choices in previous rounds. It allows potentially better paying alternatives to replace those that might pay less. Here we use $W^i(a|r^i_t)$, defined in (1), as the measure of "potentially better paying."[12] For $j = 1, \ldots, J$, $a^i_{j,t+1}$ is 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

$$a^i_{j,t+1} = \left\{ \begin{array}{c} a^i_{k,t} \\ a^i_{l,t} \end{array} \right\} if \left\{ \begin{array}{l} W^i(a^i_{k,t}|r^i_t) \geq W^i(a^i_{l,t}|r^i_t) \\ W^i(a^i_{k,t}|r^i_t) < W^i(a^i_{l,t}|r^i_t) \end{array} \right\}.$$

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.

[11] 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 about 1% of the messages of M 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.

[12] 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 games. Such strategies are unnecessary for the analysis in this paper. We intend to address these issues in our future research.

Selection is last.

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

for all $i \in \{1, \ldots, N\}$ and $k \in \{1, ..., J\}$ and where[13]

$$\varepsilon^i_{t+1} = min_{a \in A^i_{t+1}}\{0, W^i(a|r^i_t)\}.$$

Our model is now almost complete. We have shown how to move from A^i_t and π^i_t to A^i_{t+1} and π^i_{t+1}.

But we still have to explain how the process begins. One way to start would be with random selection. But subjects begin the experiment with information about their payoff functions and the game. So they have time to think before the first round about strategies that might work. So we use a more sophisticated initialization process.

Initialization starts everything off. First, for each i we generate a set, A^i_{-1}, of J messages using a uniform distribution over the feasible range. Second, for each message j in the initial set, we draw, again randomly 100 pairs of values of μ_{jk} (within the range [-4,6]) and σ_{jk} (within the range [0,5]). The payoff of message j is then calculated as $v_j = (1/100)\sum_k V^i(m_j + (N-1)\mu_{jk}, \alpha^i - T^i(m_j, \mu_{jk}, \sigma_{jk}, \gamma))$. Third, Replication takes place based on these average payoffs. This gives us A^i_0. Finally, we select the actual message using the π generated as in Selection.

If there are dominant strategies or a set of strategies that might provide a good payoff in expected value against a randomizing opponent, then our initialization process should populate A^i_0 with those messages. In some respects this is similar to using a fairly naive cognitive hierarchy model.

3.2 Some Remarks

Free Parameters Our model is not entirely determined a priori. The IEL model requires the choice of one set, one function, four processes, and two parameters.

The set to be chosen is A, which we have taken to be a subset of the message space M. In future research we intend to change A into a more general set of strategies. The function to be chosen is $W^i(a|r^i_t)$, which we have taken to be the payoff to i when they use a and the others use m^{-i} where $r^i = r^i(m^{-i})$. It is worth noting that the entire interplay between environment, mechanism, and behavior rests on the foregone utility computation of $W^i(a|r^i_t)$. The four processes are initialization, experimentation, replication, and selection.

The two parameters are J, the size of the set A, and μ, the rate of experimentation. In a very real sense, there are only two free parameters of this model.

[13] This implies that if there are negative foregone utilities 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.

Later in Section 6 we will vary all of the processes and parameters in order to examine the robustness of our results to changes in the model's specification. We defer further discussion until then but note now that the performance measures we are interested in are robust to a wide range of changes in the model's parameter values.

Calibration It is standard, in using learning models to explain experimental data, to use some of the data set to calibrate the parameters to that situation and then to test that calibration with the rest of the data. If IEL, or any other behavioral model, is to have any utility as the basis for a computer test bed for mechanism design, it seems to us that it must require, ideally, no calibration. That is, one should be able to plug in IEL, an environment, and a mechanism and produce results similar to those one would get with the same environment and mechanism using human subjects. As will be seen later in the paper, we believe that IEL with $\rho = 0.033$ and $J = 500$ provides such a model.

Other Learning Models and Large Strategy Sets Our model shares some common features with other learning models in the literature. For example, as in the Experience Weighted Attraction Model (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. One important one is IEL's ability to handle large strategy spaces, such as the continuum of possible messages for each agent in Groves-Ledyard mechanisms. 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.[14] Our model handles that problem well. It does start out with randomly chosen sets of alternatives for each agent, but 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.

Directed Experimentation and Changing Environments One interesting feature of directed experimentation is that it gives our model the ability to quickly adjust to changes in the environment. 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.

[14] In Appendix I we detail how this affects EWA. We explore more fully the implications of scaling up the strategy sets on a number of models in Arifovic and Ledyard (2004).

4 The IEL Testbed Results

In this section, we describe how we used the IEL testbed to discover answers to the mechanism design questions we are interested in - properties of the dynamics of Groves-Ledyard mechanisms. We explain how we setup and ran the testbed, the measurements we took, and the results.

4.1 Setup

We simulated the repeated one-shot game for each mechanism, γ. We call a particular simulation, a *run*. Each run is based on an environment, a behavioral model, and a mechanism.

The environment To be able to compare our results to those generated with humans, we used the utility functions and cost of production in Chen and Tang (1998). The cost, c, of producing a unit of the public good, which determines Z, is set to 100 and the utility parameters are given in table 1.

We used this environment for all the simulations and experiments reported in this paper.

The behavioral model We used the IEL model described in the previous section. For the runs reported in this section, we set the memory capacity $J = 500$. We set the rate of experimentation $\rho = 0.033$. There is nothing particularly special about this choice of parameters. In Section 6, we will examine what happens if we change these specifications.

The mechanisms Each value of γ determines a different mechanism. For these simulations we used $\gamma \in \{1, 10, 30, 50, 100, 260\}$. These choices were based on a large number of earlier simulations and the design of comparative human experiments. We will expand on these issues below.

The simulations For each mechanism, γ, we implemented 10,000 runs. Each run was terminated 100 periods after a *convergence criterion* was fulfilled.[15] We use a convergence criteria that is defined in terms of how close all agents' messages are to the equilibrium messages. This 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 a number 0.2; i.e., when $|m_t^i - m^{ie}| \leq 0.2$ for all i.

4.2 Performance measures

There are many possible measures of performance but to begin with we focus on:
(1) the time of first passage through equilibrium, (2) an index of equilibrium stability, and (3) efficiency.

[15] The maximum number of periods for each run was set at $t_{max} = 10,000$. If the convergence criterion is not fulfilled by that time, a run is terminated.

Time of convergence Rather than simply asking whether convergence occurs, we want to know how fast it occurs. The period when the convergence criterion is first fulfilled is called *the time of the first passage through equilibrium*, $T^{\gamma,r}$ for run r and given γ. The average time of the first passage through equilibrium for R runs, $\bar{T}^{\gamma}$, is given by:

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

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

Stability of convergence What happens to the agents' choices once the first passage through equilibrium has been recorded? The time of first passage would not be very interesting if the agents just rushed on by and the messages 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 in actions* S^a. It measures the frequencies with which messages close to equilibrium values are chosen by the agents during the 100 periods after the first passage through equilibrium. It is given by

$$S^{a\gamma} = \sum_{t=T^{\gamma,r}+1}^{T^{\gamma}+100} S_t^a$$

where S_t^a is an index variable that equals 1 if $|m_t^i - m^{ie}| \leq d$ for all i and otherwise equals 0. That is, $S^a = 0$ whenever at least one agent chooses a message that is more than d away from the equilibrium. We denote the average of this index across the R runs by $\bar{S}^{\gamma}$. We denote the standard deviation from this value, across the R runs, by $\sigma_{S^{\gamma}}$.

Another stability measure that we use is based on the individual's action sets - the set of strategies from which the agents choose their messages. We call this the *index of equilibrium stability in strategies* S^s. It measures the percentage of all possible choices that are close to that agent's equilibrium message.

$$S^s = \frac{\sum_{t=T^{\gamma,r}+1}^{T^{\gamma,r}+100} \sum_{i=1}^{N} \sum_{j=1}^{J} S_{j,t}^i}{NJ}$$

where $S_{j,t}^i$ is an index that equals 1 if $|a_{j,t}^i - m^{ie}| \leq d$ and otherwise equals 0.

Efficiency One measure commonly used to compare mechanism performance is the average efficiency of the outcomes across a number of iterations. Efficiency is the sum of the payoffs received by the agents divided by the maximum possible payoff, which for Groves-Ledyard mechanisms is attained at the equilibrium messages, m^e. Our measure of efficiency for a single period t is the standard

$$E_t^{\gamma} = \Big\{ \sum_{i=1}^{N} W^i(m_t^i | r^i(m_t), \gamma) \Big\} / \Big\{ \sum_{i=1}^{N} W^i(m^{ei} | r^i(m^e), \gamma) \Big\}$$

The measure of average efficiency over T periods is then just

$$E^{T\gamma} = \{\sum_{t=1}^{T} E_t^{\gamma}\}/T$$

4.3 Results

There are four main findings from our simulations with the IEL testbed. (1) There is convergence to Nash equilibrium messages of the stage game for all of the values of γ that we simulated, including $\gamma = 1$. Convergence is fast for a much larger set of the values of γ than that predicted by either the strategic complementarity condition or the dominant diagonal condition. (2) The time to first convergence, is smooth and U-shaped in γ with the minimum average convergence time occurring at $\gamma = 50$. (3) Convergence is stable in the sense that once the model first nears the equilibrium, it remains in its neighborhood. (4) Efficiencies are high and, as one might expect, inversely related to T^{γ}. That is, efficiencies have an inverted U-shape attaining a maximum at $\gamma = 50$.

Convergence to Nash Equilibrium Table 2 contains detailed data on the average time of first passage through equilibrium for $10,000$ runs for each for $\gamma \in \{1, 10, 30, 50, 100, 260\}$. The convergence criterion used here is $|m_t^i - m^{ie}| \leq 0.2$ for all i. The first column gives the value of γ, the second column presents the average values of times of first passage through equilibrium (averaged over $10,000$ runs), $\bar{T}^{\gamma}$, and the values of standard deviations, $\sigma_{T^{\gamma}}$, in the parenthesis.

The first thing to note is that our simulations converge on average for all the values of γ. Convergence is particularly fast for $\gamma = 10, 30, 50,$ and 100. Neither $\gamma = 80$, the minimum value of γ for which super-modularity holds, nor $\gamma = 30$, the minimum value of γ such that the dominant diagonal condition holds, appear to play any role in the rate of convergence for the IEL model. Convergence is in fact quicker for $\gamma = 10$ than for $\gamma = 260$.

Non-monotonicity of convergence time The really interesting fact is that T^{γ} is not monotonic with increases in γ. Instead, T^{γ} is U-shaped in γ. The fastest rate of convergence occurs at $\gamma = 50$. Further the standard deviations of the times to convergence behave in the same way.

Stability of equilibrium The third column in Table 2 gives the values of the action-based measure, S^a, of stability of equilibrium after the first passage. This is of interest particularly because of the random, but directed, experimentation of our testbed agents and the use of mixed strategies by those agents. It is possible that either random choice or experimentation could create instability and drive things very far from equilibrium for very long times. But that does not happen. IEL is remarkably stable. The values of this measure are above 90% for all of the values of γ. It is above 90% for $\gamma \geq 3$ and above 99% for $\gamma \geq 30$. Once IEL attains equilibrium, it stays there. The standard deviations from S^{γ} are reported in the

parentheses in the third column of Table 4. It can be seen that these are remarkably low for $\gamma \geq 10$. The standard deviations are also U-shaped in γ.

The 4th column provides a partial explanation for this high degree of stability. Here are listed the values of the strategy-based measure, S^s, of stability. This shows that over 98% of all actions in the sets A_t^i for $t = T^a + 1, ..., T^a + 100$, are within 0.2 of their equilibrium values. This reflects a high degree of homogeneity in the sets of alternatives A_t^i after convergence, despite the fact that experimentation is present.

Efficiency We provide two measures of the efficiency of the mechanisms. E^{100} measures the average efficiency attained by the mechanism over the first 100 rounds. E^{10} measures the average efficiency attained by the mechanism over the first 10 rounds. The average efficiencies over 100 periods are very high, even for $\gamma = 1$ which converges slowly. When we only consider only the first 10 periods (when the mechanism may not have had time yet to converge), the efficiencies are not as high. But for $\gamma = 30, 50, 100$ the efficiencies are over 90%, peaking at 94% for $\gamma = 50$.

Summary *For $\gamma = 30, 50,$ and 100, IEL is fast, stable and efficient.*

5 Comparison of IEL to experimental data

The IEL 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). These data provide strong, but limited, support for the IEL model. A second set was generated by us after seeing how the results of the testbed compared to their data.

5.1 Chen and Tang's data

Chen and Tang (1998) conducted 7 experiment sessions each with $\gamma = 1$ and $\gamma = 100$. A summary of the results from these experiments is presented in Table 3 along with the relevant data from the IEL simulations.The good news is that the experimental convergence times are remarkably consistent with the predictions of IEL. The bad news is that the human subjects are less stable than IEL.

Convergence to Nash Equilibrium A look at the average time to converge, T^γ, reveals a remarkable correspondence between their data and our IEL testbed results. 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 IEL testbed produced a really high average convergence time close to 340 periods. For $\gamma = 100$, convergence did occur, according to the 0.2 criterion in all of the sessions reported in Chen

and Tang. The average time to converge was 9.0 with a standard deviation of 5.7. The IEL simulations converged on average in 8.0 rounds with a standard deviation of 1.8.

Stability The second statistic of interest is the message based measure[16] of stability, $S^{a\gamma}$. Here the correspondence between the experimental data and IEL falls apart. For $\gamma = 100$, the seven Chen and Tang experiments yield an average stability of 43 while the IEL testbed produced an average of 95. One reason for this discrepancy may be that Chen and Tang required messages in multiples of 0.2 while we allowed all messages in the IEL simulations.[17]

Efficiency It is interesting to note that in both the human experiments and in the IEL simulations, average efficiencies over 100 rounds are high for $\gamma = 1$, even though convergence has not occurred by 100 rounds. IEL achieves 97% efficiency and humans achieve 93.4%. Of course, the efficiencies are much lower for the first 10 rounds, as one might expect with slow convergence.

At $\gamma = 100$, the experimental and simulation data are very similar. Human subjects achieved 98.1% on average for the first 100 rounds while IEL achieved 99.3% for the same time period. For just the first 10 periods, human and computer agents achieved 93%.

Next The close correspondence between IEL and experimental subjects with respect to the time of convergence suggests further questions. For example, the Chen and Tang data suggest a significant difference in the performance of the Groves-Ledyard mechanisms when $\gamma = 1$ and when $\gamma = 100$. One hypothesis put forward by both theorists and experimentalists for this difference is based on stragtegic complements. That is, they predict we should expect (reasonably rapid) convergence for $\gamma \geq 80$ and, perhaps, a lack of convergence for $\gamma < 80$. On the other hand our IEL testbed results suggest no such dramatic difference. Rather, relative to $\gamma = 100$, convergence should be faster for, say, $\gamma = 50$ and slower for $\gamma \geq 100$. To see if we could settle this issue, we generated a new data set at the California Institute of Technology.

5.2 Our data

Our experiments were conducted between May and September 2007. We used the experimental software Z-tree.[18] Our experimental design is very similar to Chen's and Tang's with two

[16] Because we did not always have 100 periods in an experimental session after the first pass by equilibrium, we need to use a slightly different measure than 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 and compute the the percentage of actual messages that are near equilibrium messages.

[17] Actually, Chen and Tang allowed only integer messages from their subjects. To implement this without a very coarse set of actions, they took the messages chosen by the subjects and divided them by 5 before applying the Groves-Ledyard rules. That is, their subjects chose integers s_t^i which Chen and Tang converted to $m_t^i = s_t^i/5$ and then applied the Groves-Ledyard rules summarized in (1). This is equivalent to allowing messages in only multiples of 0.2 in our simulations.

[18] The experiment instruction is provided in the Appendix II.

modifications. In Chen and Tang, subjects could make only integer number choices, equivalent to allowing only multiples of 0.2. We allowed subjects make real number choices with a two decimal points restriction, equivalent to allowing multiples of 0.01. Second, we added a 'what-if' calculator[19] that gave the subjects an opportunity to examine what payoffs they would obtain for different choices of m^i and pairs of μ_i's and σ_i's. This 'what-if-calculator' is similar to the evaluation of hypotheticals in our IEL algorithm.

We conducted 4 experiment sessions each with $\gamma = 1$, 10, 30, 50, 100, and 260. Each experimental session lasted for 100 periods. Subjects earned on average about \$38 which included a show-up fee of \$10. A summary of the results from these experiments is presented in Table 4 along with the relevant data from the IEL simulations. The good news is that the data are consistent with the qualitative IEL predictions about convergence and the U-shaped curve. The bad news is that the human subjects are both faster and less stable than IEL.

Convergence to Nash Equilibrium IEL predicts that convergence should occur in less than 100 periods when $\gamma \geq 5$ and should take more than 100 periods otherwise. Our data certainly support that. For $\gamma \geq 10$, our subjects always converged to the Nash Equilibrium and at a fairly rapid rate. For $\gamma = 1$ our subjects did not converge in 100 periods all 4 sessions.

IEL predicts a U-shaped curve for T^γ as a function of γ with the minimum time to converge occurring at 50. Our data certainly support that. In Figure 1 look at both the experiment data and the IEL predictions using two convergence criteria, $d = 0.1$ and $d = 0.2$. All are U-shaped with a minimum at 50.

The experimental data are below (i.e, faster) and flatter than the IEL predictions, but IEL does remarkably well.

Stability From Table 4 it can be seen that the human subjects are less stable than IEL when using the measure S^a. For IEL, with the exception of $\gamma = 1$, in the first 100 periods after convergence more than 98% of the IEL messages sent are within 0.1 of the equilibrium messages. The human subjects are close to this for $\gamma = 50$ and 100 but for the other values of γ humans are certainly more prone to send messages that are away from their equilibrium values. It is an open question whether this lack of stability is the result of random errors, more experimentation, or something else. Random errors could occur for example as in the Quantal Response model. However, to achieve this with IEL we would have to find a way to have more variety in the sets A_t^i. If we allowed experimentation in IEL after replication, remember replication eliminates many of the experiments, then we could have more variations. This is set aside for future research.

[19] Because a feature of Z-tree keeps a log of the use of this scenario calculator, we were able to record the number of times that the calculator was used prior to each actual choice. The usage of the calculator was relatively high in the initial periods, and then decreased over time, as subjects converged towards a particular action.

Efficiency Remarkably, efficiencies are fairly high for the human experiments. This can be easily seen in Figure 2. This is particularly true for $\gamma = 50$ and 100 where the efficiencies from the human experiments are just slightly under the efficiencies for the IEL simulations. (This is also true for $\gamma = 100$ in the Chen-Tang experiments.) So even though humans are slower and less stable than IEL, the utility levels attained are very similar in the range of $\gamma \in [50, 100]$. Outside that range, humans seem to be slower, less stable, and less efficient than IEL.

Summary The qualitative predictions of the IEL model are replicated in the experimental data. Convergence occurs for a wide range of γ, T^γ is U-shaped in γ, and efficiencies are high. But, the quantitative predictions do not hold up. Humans seem to be faster and less stable than IEL for $\gamma > 10$ and efficiencies are much lower outside the range of $\gamma \in [30, 100]$.

Next It is worth considering whether variations in IEL might produce simulations that better conform to the experimental results. We turn to that now.

6 Sensitivity

We would ultimately like to be able to use the IEL 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.

Remember there are three main groups of findings in earlier sections. (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. (2) The time to first convergence, is U-shaped in γ. (3) 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.

We will see that the above features characterize the dynamics of all of the different versions of the model that we examine.

6.1 The baseline simulations

We begin with a baseline set of simulations using the following parameters: the size of A_t^i is $J = 100$, normal experimentation with $\rho = 0.033$, tournament replication, proportional selection and random initialization. Only the first and last are different from the model that is described in Section 3 and used for simulations in Section 4. Random initialization generates a set of J values using a uniform distribution over the feasible range. This is simpler (and quicker) than the process used earlier.

Table 5 contains detailed data for the average time of first passage through equilibrium, for 10,000 runs, using the convergence criterion that $|m_t^{ai} - m_t^{ei}| \leq 0.1$ for all i. 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), T^γ, and the values of standard deviations, σ_{T^γ}, in the parenthesis.

In Figure 3, we present, graphically, the results for $\gamma \in \{1, \ldots 100\}$. In Figure 4 are the data for $\gamma \in \{10, 100\}$. In Figure 5 are the data for γ between 120 and 1000 in the increments of 20. To provide a feel for the distribution of times to first passage, we also provide Figure 6. It contains histograms of the frequencies of convergence times for $\gamma = 1$, 50, 80 and 100.

Convergence to Nash The simulations converge on average to Nash equilibrium for all the values of γ. The average convergence time, T^γ is smooth and U-shaped in γ. This can be most easily seen in the figures. There are no discrete jumps; i.e., no values of γ at which the system switches between stability and instability. This smoothness is consistent with the findings of Chen and Gazalle (2004). But, in our case, it is not because we are close to the strategic complements condition of $\gamma = 80$. Indeed, IEL shows relatively rapid convergence for values of γ as low as 10, lower even than the dominant diagonal condition of $\gamma = 30$. Neither the dominant diagonal condition for stability nor the strategic complementarity condition for stability appear to play any role in the simulations.

There are a couple of other facts to note. First, the average times to convergence are not much different for the values of γ between 10 and 100. For example, T^{20} is within a minus one standard deviation of T^{50} and also, T^{50} is within a plus one of its own standard deviation of T^{20}. Similar relationships can be observed for all the values of γ up to 100. Second, the standard deviation of T^γ decreases with the increases in γ, drops to the lowest values for $\gamma \in \{40, 60\}$, but then slightly increases for the values of $\gamma > 60$. Thus, the value of γ that results in the fastest passage through equilibrium also results in the least amount of variation.[20]

Stability of equilibrium The fourth column in table 2 gives the values of the strategy measure that we use to study the stability of equilibrium after the first passage. The values of this measure are above 85% for all of the values of $\gamma \in \{1, \ldots, 100\}$. It is above 90% for $\gamma \geq 3$ and above 95% for $\gamma \geq 22$. This reflects a high degree of homogeneity in the sets of alternatives A_t^i after convergence, despite the fact that experimentation is present. Standard deviations from $\bar{E}_s^\gamma$ that are reported in the parentheses in the fourth column of Table 2 are generally low, but are also decreasing as γ increases.

6.2 The variations

Next we proceed to consider various parameters and processes we could have used for IEL in order to see how robust our results are. Unless otherwise specified below, when we change

[20] It is evident from Figures 4 & 5, the distributions of convergence times is not symmetric. They have a large righthand tail. This is not surprising since all observations are bounded from below because T $\geq$ 1. Because of this it is not surprising that as the mean moves closer to 0, the variance also shrinks. Of course the variance could increase with more observations in the right-hand tail but that doesn't happen.

one parameter, we keep the other baseline parameters fixed.

Experimentation Our baseline simulations used an experimentation process based on the normal distribution whose mean is history determined as the value of the action in A_t^i being experimented with. We compare that to a history independent process: we use the uniform density over all possible strategies. We also consider two very different rates of experimentation: $\rho = 0.033$, our baseline rate, and $\rho = 0.25$. We compare these two approaches in Table 6.

Experimentation using the uniform distribution results in significantly higher values of T^γ than for the normal distribution. The time to convergence under the history independent experimentation with the uniform distribution 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. (As long as $\gamma > 1$) the effects are on the order of only 10-30% for an 750% increase in ρ - and decreasing in γ. These effects seem to be considerably magnified at $\gamma = 1$.

Comparing the results in Table 5 with the experimental data in Tables 2 and 3, it is obvious that experimentation from the normal distribution with $\rho = 0.033$ is favored. And that the choice of the value of ρ is probably not that crucial.[21]

Replication For the **replication process**, in addition to the tournament replication described in Section 3, we tried using proportional replication. In proportional replication, an alternative $a_{j,t}^i$ occupies an interval of a measure equal to its probability, $\pi_{j,t}^i$. Then, for each member j, a random number between 0 and 1, r_j, is drawn from a uniform distribution. An alternative that occupies the range of values where r_j belongs is determined. Then a copy of that alternative is assigned to member j.

We provide in Table 7 a comparison of the two procedures with uniform experimentation at a rate $\rho = 0.033$. As one can see, proportional replication yields really high times to first passage through equilibrium. They are 45-60 times longer than under tournament replication.

Since the times of convergence when we use proportional replication are totally inconsistent with the experimental evidence, we will not consider this variation further.

Selection We 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 earlier. For exponential selection, one computes selection probabilities, using exponentialized payoffs,

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

[21] One thing we did not test was the effect of changing the standard deviation of that normal to something other than 1. A larger variance might not change the times to convergence very much but might add in that level of instability that the experimental data suggest humans have over IEL.

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 utilities 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 8, the convergence times for 4 different values of λ, for $\gamma = 1, 50$ and 100 using uniform experimentation with the rate $\rho_u = 0.033$.

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 T^γ. 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. 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.[22] The reason the selection process is pretty much irrelevant in our testbed is that it really only plays a role when the set of provisional strategies is heterogeneous. But that set becomes homogeneous fairly quickly through the replication process. Consequently which action is selected from A_t^i doesn't matter very much.

Since the data are pretty much the same whether we use exponential selection or proportional selection, we have elected to stay with proportional. That way we have one less free parameter to worry about.

Strategy Space Size One of the important free parameters that does matter is J, the size of A_t^i.

Increases in J speed up the rate of convergence at a decreasing rate, at least for the range we have considered. The data are in Table 9. Increasing J from 50 to 100 reduces the time to convergence on average by 32%. An increase from 100 to 200, reduces average convergence times by 23%. From 200 to 500, we get a 16% drop in times. The reduction is only 6% for an increase in J from 500 to 1000 with virtually nothing happening for $\gamma = 50$ and $\gamma = 100$, the values of γ for which we see the fastest convergence with IEL.

Because IEL is slower than human subjects we have focused on those $J \geq 500$. And because the time of computation increases significantly for $J = 1000$, we have chosen $J = 500$ for the analysis in the text.

Initialization We have used two initialization procedures: a purely random initialization and the one described in Section 3, a modified procedure in which a hypothetical strategies are evaluated using hypothetical plays of the others before choosing the first message.

In Table 10 we provide the data from these two approaches for two values of J and three values of γ. As one can easily see, the effect of the modified procedure is to flatten out the time to convergence as a function of γ. There is virtually no effect at $\gamma = 50$ for either value

[22] 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).

of J, but there is an effect as we move away in either direction. If the time to converge is low (remember it is minimized at $\gamma = 50$), then preliminary screening of potential actions does not seem to be very important. But if the time to converge is larger, then some initial screening can increase the rate of convergence.

Since the times to converge under the modified initialization process are closer to the experimental data, we chose to use that in Section 3.

Summary of Sensitivity Analysis of IEL Our baseline model of IEL consisted of $J = 100$, random initialization, experimentation using the normal distribution with $\rho = 0.033$,, tournament replication, and proportional selection. We considered several alternatives.

Requiring consistency with the experimental data leads us to reject random initialization in favor of a modified approach, uniform experimentation, proportional replication, and exponential selection. The sensitivity analysis suggests that the choice of ρ is not too important although the lower value of $\rho = 0.003$ is more consistent with the data. Only the choice of $J = 100$ is potentially an issue. We chose $J = 500$ for consistency with the data and ease of computation.

7 Final Thoughts

Traditional behavioral models such as EWA do not behave very well when there is a large strategy space. We have proposed a new behavioral model for repeated play. It has two characteristics that are not common to other learning models. First, it can handle large, even infinite, strategy sets without additional instruments. Second, it uses directed, not random, experimentation which minimizes time wasted on trying obviously unproductive strategies.

To demonstrate the potential usefulness of IEL, we used it to create a testbed for studying mechanisms. We applied that testbed to the study of the dynamics of Groves-Ledyard mechanisms. The key findings were: (1) the average time to convergence is U-shaped and smooth in γ, the free parameter of this class of mechanisms, and (2) neither of the cut-points of the sufficient conditions for stability based on strategic complementarities and dominant diagonal analysis seem to play much of a role in whether a mechanism converges or at what rate it converges. The latter is a bit surprising since our behavioral model has a significant element of best reply in its formulation.

To test the accuracy of the testbed, we compared the data generated by the testbed with data generated in economic experiments. The data certainly support the existence of the U-shaped curve for the average time to converge as a function of γ. Further, it does appear that the testbed model with modified initialization, tournament replication, historically dependent normal experimentation and proportional selection is the right one. However, it is also clear that the testbed is a bit slower than experimental subjects.

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.[23] Measurements from a good 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 of the time and monetary resources required for their implementation.

[23] We have had some success in taking the testbed as it is and using it to analyze the performance of two call market designs for a private good environment. See Arifovic and Ledyard (2003).

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 (2003), Information and Dynamics: Sequences of Call Markets, *Information Systems Frontiers*, 5(1): 39-45, 2003.

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.

Gabay, Moulin (1980)

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.

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

Milgrom P., and Roberts (1990) Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities, Econometrica, 58, 1255-1277.

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.

Appendix I:

Analysis of Experience-weighted Attraction Learning

Experience Weighted Attraction (EWA) is probably the leading example of a successful behavioral model of learning. One might reasonably ask: why not replace IEL in our testbed with EWA? Even though it has exhibited superior performance in a number of environments (mainly with small strategy spaces), EWA has difficulties with application to the GL environment common to other commonly studied models of individual behavior: it has to be tuned, i.e., parameters must be set differently for different games or market conditions in a way that is not necessarily known a priori, and it does not scale; that is, as the number of strategies increases the models behave less and less like real subjects do. We illustrate this issue by reporting on the results of a set of EWA simulations in our GL environment.

In order to implement EWA, we first discretize the message space and construct a set of 51 messages, equally space within our range of $[-4, 6]$. Once determined, this set does not change over time. What changes are the probabilities that each one of the messages in the set is selected to be played.

In EWA there are two main variables that are updated after each round of experience: N_t, the number of "observation-equivalents" of past experience (called the *experience weight*); and $\zeta^i_{j,t}$, the attraction of message $m^j_{i,t}$ at the end of period t. Their initial values N^i_0 and $\zeta^i_{j,0}$ can be interpreted as prior game experience and/or principal's predictions.

The experience weight, N_t, is updated according to

$$N_t = \rho N_{t-1} + 1 \tag{2}$$

for any $t \geq 1$, where ρ is a depreciation rate or retrospective discount factor.

The key part of evaluation of the attractions is the payoff that the message, m^i_j received when it was used or would have received if it had been used, taking the behavior of other agents as given. $W^i(m^i_j|r^i_t)$ is the value of this payoff which is i's utility for m^i_j given the information r^i_t.

The evaluation by agent i of an attraction of message m^i_j at time t, $\zeta^i(m^i_j)$, is computed as follows:

$$\zeta^i(m^i_j) = \frac{\left[\phi \cdot \zeta^i(m^i_j)\right] + \left[\delta + (1-\delta) I^i_j\right] W^i(m^i_j|r^i_t)\right]}{N_t} \tag{3}$$

$I^i_{t-1}(m^i_j, r^i_{t-1})$ is equal to 1 if the message, m^i_j, was chosen by i at time t and is equal to 0 otherwise. The parameter δ determines the extent to which hypothetical evaluations will be used in computing attractions. The factor ϕ is a discount factor or decay rate, which depreciates the previous attraction.

Finally, the selection probabilities for the following period, $t+1$ are computed using exponentialized payoffs,

$$\pi^i_{k,t+1} = \frac{e^{\lambda W^i(m^i_j|r^i_t)}}{\sum_{j=1}^{J} e^{\lambda W^i(m^i_j|r^i_t)}}.$$

The free parameters of this model are: $|S|, N(0), \zeta^i(m^i, 0), \rho, \phi, \delta$, and λ. If $\delta = 0$ and $\rho = 0$ then this is just the Reinforcement Learning (RL) model. If $\lambda \to \infty$, and $\delta = 1$ and $\phi = \rho = 0$, then this is the best-reply model and all the probability is put on the strategy that maximizes utility in response to s_t^i.

Testbed results with EWA We conducted a total of 8 sets of simulations. The one that we report in this paper is the one that gave the best results in terms of the convergence and stability measures. This set consists of the following EWA parameters: $\lambda = 0.35$, $N_0 = 10$, $\delta = 0.96$, $\rho = 0.95$, and $\phi = 0.991$.[24]

We simulate the EWA model for the values of γ equal to {1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150}. For each value of γ, we conduct 10,000 simulations, using different random number generator seed for each simulation. Each simulation was run for 10,000 periods (unless it was terminated earlier as the convergence occurred). We report below on the results obtained using our testbed measures. Column 1 gives the value of γ, column 2 gives the total number of simulations conducted for each γ, column 3 reports on the number of simulations, C (out of 10,000) that converged within 10,000 periods, column 4 presents the average values of times of first passage through equilibrium (averaged over $10,000$ runs), T^γ, and the values of standard deviations, σ_{T^γ}, in the parenthesis, and finally column 5 reports on our measure of equilibrium stability.

γ	R	R/C	T^γ (σ_{T^γ})	S^γ (σ_{S^γ})
1	10000	0	NaN (NaN)	NaN
10	10000	506	120.72 (570.53)	11.18
20	10000	1641	56.35 (200.82)	13.70
30	10000	2587	52.42 (262.52)	17.51
40	10000	0	NaN (NaN)	NaN
50	10000	3458	66.22 (477.32)	23.81
60	10000	3600	73.63 (536.81)	26.26
70	10000	3153	92.42 (618.03)	26.ld
80	10000	2627	98.33 (650.17)	29.06
90	10000	2249	94.75 (627.97)	31.10
100	10000	1950	113.49 (758.51)	32.13
110	10000	1719	102.25 (700.15)	35.47
120	10000	1533	102.50 (710.14)	37.85
130	10000	1397	92.62 (623.11)	38.22
140	10000	1311	105.51 (653.65)	40.48
150	10000	1243	136.24 (788.01)	42.01

[24]The other sets that we simulated are: $\phi = 0.7$ $\delta = 0.7$, and $\lambda = 0.35$; $\phi = 0.7$ $\delta = 0.7$, and $\lambda = 0.035$; $\phi = 0.7$ $\delta = 0.8$, and $\lambda = 0.35$; $\phi = 0.99$ $\delta = 0.7$, and $\lambda = 0.35$; $\phi = 0.99$ $\delta = 0.7$, and $\lambda = 0.035$; $\phi = 0.99$ $\delta = 0.8$, and $\lambda = 0.35$; and, $\phi = 0.99$ $\delta = 0.96$, and $\lambda = 0.35$.

For the values of γ equal to 50, 100, and 150, we report on more EWA robustness checks in Arifovic and Ledyard (2004).

Convergence to Nash Equilibrium First thing to note is that for $\gamma = 1$, and $\gamma = 40$ no convergence occurs in any of the 10,000 simulations. Second, there is no value of γ that results in convergence of all of the 10,000 simulations. The number of simulations that converge C varies from the lowest of 506 for $\gamma = 10$ to the highest for $\gamma = 60$ of 3,600. At first, the average time to convergence drops from 120.72 for $\gamma = 10$ to 56.35 for $\gamma = 20$, and 52.42 for $\gamma = 30$. Then, there is no convergence for $\gamma = 40$ [25] After that, starting with $\gamma = 50$, there is a monotonic increase in times to convergence until we reach $\gamma = 150$ with T^{γ} equal to 136.24. This pattern is followed by the similar pattern for the values of $\sigma_{T^{\gamma}}$, the standard deviations from the average time to convergence. They have fairly large values which follow the pattern of average times to convergence, i.e. decrease for $\gamma = 20$, and 30, and then start increasing with $\gamma = 50$.

It is interesting that, with the exception of $\gamma = 40$, this is exactly the same pattern implied by IEL.

Stability of equilibrium The last column in Table 5 gives the values of the measure that we use to study the stability of equilibrium after the first passage. The measure takes a low value of 11.18 for $\gamma = 1$, and then increases slightly for $\gamma = 10$, and 20. After a breakdown for $\gamma = 40$, the measure continues increasing steadily until it reaches the value of 42.01 % for the highest value of γ.

These low values of this measure indicate that the algorithm does not exhibit much stability once the first time passage through equilibrium has been recorded. The value equal to 42% means after that first passage, for the next 100 periods, agents choose strategies that are in the neighborhood of the equilibrium only 43% of the time. This is in stark contrast to very high values of this measure for both the experimental and the IEL data.

[25] Note that non-convergence for this value of γ is characteristic of the other set of simulations for $\phi = 0.7$ as well. We also tried values of $\gamma = 39$, and $\gamma = 41$, and we obtained results in terms of average convergence times and stability of equilibrium that fall in between these values for $\gamma = 30$, and $\gamma = 50$.

Appendix II: Experiment Instructions

Experiment Instructions *ID = PLAYER 1*

Introduction

You are about to participate in a decision process in which one of numerous competing alternatives will be chosen. This is part of a study intended to provide insight into certain features of decision processes. If you follow the instructions carefully and make good decisions you may earn a considerable amount of money. You will be paid in cash at the end of the experiment.

Your final payoff will be determined by a project level which will be chosen by the group, and by your individual expenditure on the project. The decision process will proceed as a series of rounds during each of which a project level will be determined and financed. The "level" can be negative, zero or positive "units", the exact level of which must be determined.

In your folder, you will find a chart which describes the payoffs to you of various decisions, called the Payoff Chart. *You are not to reveal this information to anyone.* It is your own private information.

The Situation

The payoff each period, which is yours to keep, is the difference between the value to you of the project level which is chosen, and your individual expenditure on the project. All values are stated in francs and can be converted into cash at a rate of $\underline{\$2}$ dollars per $\underline{100}$ francs at the end of the experiment. Note that in some cases your values can be negative. It is also possible that your expenditures can be negative (that is, rather than paying for the project you are paid.). These will be explained in turn.

Project level determination (Y) Each round each individual will choose a proposed addition *(x)* to the status quo of zero project level. This proposed addition can be any amount ranging from $\underline{\text{-4}}$ to $\underline{6}$. For these experiments you will be restricted to values of one-hundreths, $x.xx$. These amounts will be added together to get the total of proposed additions *(Y)*. This total is the project level that will be chosen. For example, if $x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 4,\ x_5 = 5$, then the project level is $Y = x_1 + x_2 + x_3 + x_4 + x_5 = 15$.

Payoffs Your payoff will depend on your individual value for the project less your share of the cost minus an additional expenditure. That is,

Your Payoff = Value - Cost - Expenditure.

The various pieces of this are described below. For your convenience we provide you with a payoff chart.

Individual Value Each individual has a different set of values for each different project level. The formula for your value $= 26Y - Y^2 + 11$. This value can be positive or negative. *The formula for your value may be different from others'.*

Project Costs Each unit of the project costs 100 francs. Your share of this cost is $100/5 =$ 20. Hence, total cost for a project is 100 times the project size. In our example, it would be 1,500 francs. And your share of this cost would be $(1500/5) =$ 300 francs.

Individual expenditures The level of your individual expenditures depends upon your individual proposed addition **(x)**, the average of proposed additions of other participants **(A)** and the variability among the proposed additions of the other participants **(V)**. For example, if $x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 4,\ x_5 = 5$, then for subject No. 1, the sum of others' proposed addition is $S_1 = x_2 + x_3 + x_4 + x_5 = 14$ and the average of the others' proposed additions is $(14/4) = 3.5$. (Notice that S_1 is also $Y - x_1$.) The variability measure reflects how scattered the additions of others are. For example if all of the other participants give the exact same number then there is no scatter at all and the variability is zero. Suppose that all of the other participants give a different number but all numbers differ very little, then the scatter is low as is the measure of variability. In the above example, the variability for subject No. 1 is $V_1 = [(x_2 - S_1/4)^2 + (x_3 - S_1/4)^2 + (x_4 - S_1/4)^2 + (x_5 - S_1/4)^2]/3 = 1.67$. Notice your proposed addition, x, does not affect the calculation of your own A and V.

The formula for your expenditure is $20(x - A)^2 - 25V$.

Payoff Chart The payoff chart summarizes your payoff as a function of value and cost to you of the level of the project chosen, and the level of individual expenditures that you will incur depending upon the choices of additions that you and other participants make.
The formula for your payoff is $26Y - Y^2 + 11 - 20Y - 20(x - A)^2 + 25V$ which is equivalent to

Your Payoff $= 6(x + 4A) - (x + A)^2 - 20(x - A)^2 + 25V + 11$.

On the payoff chart the horizontal axis is the average of others' proposed additions, A. The vertical axis is your payoff when the variability of others is $V = 0$. Each curve represents your payoff from a particular choice of proposed addition, x. The small box on the right hand side of the chart gives the color of the curve for the eleven different proposed additions charted. Since the chart would be difficult to read if all possible proposed additions were plotted, only eleven different ones equal distant from each other are given. The curves for proposed additions which are not given, lie between the given curves.

Notice the values on the vertical axis of your payoff chart are your payoffs when the variability of others is $V = 0$. Your actual payoff is the value on the vertical axis plus $50V$.

Different participants might have different Payoff Chart.

"What-if" Scenario Analyzer At the beginning of each round, before you submit your proposal, you can use the "What-if" Scenario Analyzer to help you compute your payoff for different combinations of your proposal, the sum and variability of others' proposal. The Analyzer allows you to accurately compute your payoffs when the variability of others' proposals is not zero, while the Payoff Chart gives you an overview of your payoff information for different combinations of your proposal and the sum of others' proposals when the variability of others' proposals is zero.

Procedure for Each Round At the beginning of each round, you will enter a proposal on the terminal. The central computer will then calculate the sum of others'

proposals, the variability of others' proposals and your net payoff, and send this information back to you.

It is crucial that you check your Payoff Chart before and after each decision. From the Chart you can see your choice determines which curve you use, and others' choices determine the level of A and the amount of shift (due to V) in your payoffs.

There will be 100 rounds using this mechanism. There will be no practice rounds. From the first round, you will be paid for each decision you make.

Feel free to earn as much cash as you can. Are there any questions?

Review Questions

1. If each of you propose the following units (the subscripts correspond to your real ID numbers): $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$, (therefore, $V_1 = 1.67$, $V_2 = 2.92$, $V_3 = 3.33$, $V_4 = 2.92$, $V_5 = 1.67$.)
   (1) The total level of the project, $Y =$
   (2) The sum of others' proposal, $S =$

2. Suppose all others have the same proposed addition as in Question 1, you alone raise your addition by 1 unit, then
   (1) The total level of the project, $Y =$
   (2) The sum of others' proposal, $S =$
   (3) The variability of others' proposed additions, $V =$

3. True or False: My payoff is determined by my own proposed addition only.

4. If your proposed addition is $x = 0.5$, the average of others' proposed addition is $S = 0.6$,
   (1) Find your payoff from your Payoff Chart when $V = 0$: $P =$
   (2) If $V = 10$, your payoff is _____.

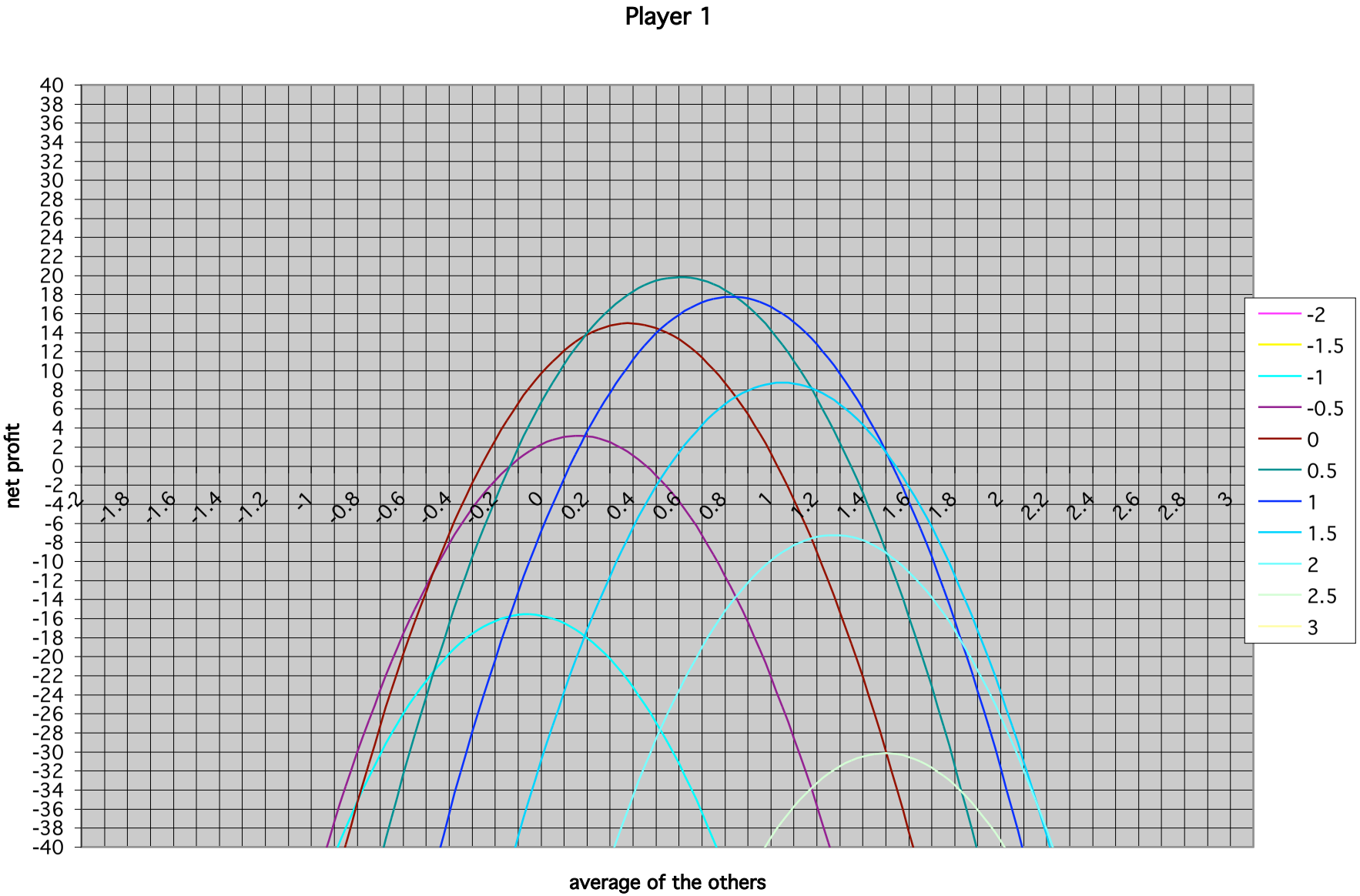
Player 1
net profit
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
-24
-26
-28
-30
-32
-34
-36
-38
-40
-2
-1.8
-1.6
-1.4
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
average of the others
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3

Table 1 - Consumer Utility Parameters

agent	1	2	3	4	5
A_i	26	104	38	82	60
B_i	1	8	2	6	4
α_i	200	10	160	40	100

Table 2 - IEL simulations

γ	$T^\gamma(\sigma_T)$	$S^a(\sigma_{S^a})$	$S^s(\sigma_S^s)$	E^{100}	E^{10}
1	339.88 (71.26)	93.84 (12.75)	98.62 (2.58)	97.01	71.18
10	12.59 (3.87)	98.81 (1.35)	99.58 (0.35)	98.78	87.87
30	7.52 (1.12)	99.33 (0.85)	99.72 (0.22)	99.36	93.61
50	7.12 (1.03)	99.40 (0.79)	99.73 (0.21)	99.42	94.27
100	8.00 (1.77)	99.34 (0.85)	99.71 (0.23)	99.29	93.02
260	19.89 (14.80)	99.21 (1.04)	99.66 (0.28)	98.18	87.86

The convergence criterion used is distance in messages with $d = 0.2$.

Table 3 - Chen and Tang Experimental Results

γ	T (σ_T)	S (σ_S)	E^{100}	E^{10}
1	NA(7/7)	NA(7/7)	93.4	80.5
100	9.0 (5.7)	72 (19.3)	98.1	93.1

The convergence criterion used is distance in messages with $d = 0.2$.
NA(7/7) = 7 of the 7 sessions did not converge in 100 rounds.

Table 4 - Arifovic and Ledyard Experimental Results

γ	T (σ_T)	S (σ_S)	E^{100}	E^{10}
1	NA(4/4)	NA(4/4)	81.1	-21.7
10	18.0 (21.6)	68.0 (39.7)	95.6	61.0
30	3.25 (2.22)	92.2 (8.4)	97.1	74.8
50	2.5 (0.6)	94.0 (4.5)	99.2	97.0
100	2.75 (2.22)	94.2 (8.9)	98.7	91.7
260	3.25 (2.63)	89.5 (11.4)	90.2	68.8

The convergence criterion used is distance in messages with $d = 0.2$.
NA(4/4) = 4 of the 4 sessions did not converge in 100 rounds.

Table 5 - A Set of Baseline Runs
The convergence criterion is $d = 0.1$.

γ	R	T_γ^c ($\sigma_{T_\gamma^c}$)	S_s^γ ($\sigma_{E_s^\gamma}$)
1	10000	903.38 (273.97)	85.13 (8.41)
2	10000	320.09 (111.78)	89.4 (7.06)
3	10000	178.95 (65.64)	91.17 (5.72)
4	10000	118.89 (45.4)	92.11 (4.78)
5	10000	87.36 (34.92)	92.73 (4.12)
6	10000	68.71 (28.39)	93.17 (3.56)
7	10000	55.84 (23.81)	93.48 (3.19)
8	10000	46.64 (20.58)	93.72 (2.93)
9	10000	39.37 (17.73)	93.967 (2.65)
10	10000	34.81 (15.98)	94.12 (2.51)
11	10000	30.51 (14.35)	94.23 (2.40)
12	10000	27.40 (13.06)	94.34 (2.28)
13	10000	24.91 (11.68)	94.43 (2.20)
14	10000	23.18 (11.12)	94.55 (2.02)
15	10000	21.57 (10.04)	94.69 (1.83)
16	10000	20.47 (9.47)	94.71 (1.88)
17	10000	19.33 (8.77)	94.80 (1.73)
18	10000	18.43 (8.39)	94.86 (1.67)
19	10000	17.87 (8.01)	94.93 (1.57)
20	10000	17.18 (7.61)	94.95 (1.59)
21	10000	16.78 (7.46)	94.99 (1.54)
22	10000	16.51 (7.39)	95.00 (1.54)
23	10000	16.10 (7.01)	95.05 (1.51)
24	10000	15.69 (6.82	95.08 (1.44)
25	10000	15.37 (6.59)	95.10 (1.44)
26	10000	15.18 (6.55	95.14 (1.39)
27	10000	15.01 (6.43)	95.15 (1.37)
28	10000	14.91 (6.28)	95.17 (1.32)
29	10000	14.72 (6.28)	95.18 (1.32)
30	10000	14.61 (6.17)	95.21 (1.35)
31	10000	14.42 (6.14)	95.18 (1.45)
32	10000	14.29 (6.06)	95.23 (1.29)
33	10000	14.22 (6.12)	95.26 (1.21)
34	10000	14.13 (6.05)	95.24 (1.33)
35	10000	13.94 (5.82)	95.27 (1.26)

Table 5, cont.d

γ	R	T^c_γ $(\sigma_{T^c_\gamma})$	S^{γ_s} $(\sigma_{E^\gamma_s})$
36	10000	13.92 (5.79)	95.25 (1.34)
37	10000	13.92 (5.81)	95.27 (1.27)
38	10000	13.83 ((5.74)	95.28 (1.21)
39	10000	13.79 (5.85)	95.29 (1.28)
40	10000	13.67 (5.65)	95.29 (1.22)
41	10000	13.68 (5.8)	95.31 (1.25)
42	10000	13.80 (5.93)	95.32 (1.22)
43	10000	13.55 (5.66)	95.3 (1.26)
44	10000	13.65 (5.85)	95.32 (1.18)
45	10000	13.52 (5.69)	95.31 (1.24)
46	10000	13.50 (5.7)	95.3 (1.25)
47	10000	13.52 (5.76)	95.31 (1.23)
48	10000	13.57 (5.79)	95.3 (1.27)
49	10000	13.60 (5.86)	95.29 (1.29)
50	10000	13.48 (5.76)	95.31 (1.21)
51	10000	13.54 (5.76)	95.33 (1.18)
52	10000	13.60 (5.79)	95.29 (1.21)
53	10000	13.64 (5.83)	95.31 (1.18)
54	10000	13.63 (5.85)	95.32 (1.20)
55	10000	13.61 (5.76)	95.29 (1.26)
56	10000	13.63 (5.78)	95.3 (1.28)
57	10000	13.68 (5.87)	95.32 (1.18)
58	10000	13.61 (5.86)	95.31 (1.21)
59	10000	13.75 (5.85)	95.3 (1.27)
60	10000	13.79 (5.98)	95.28 (1.29)
61	10000	13.81 (5.92)	95.26 (1.36)
62	10000	13.82 (6.03)	95.27 (1.35)
63	10000	13.85 (6.04)	95.29 (1.26)
64	10000	13.86 (6.00)	95.26 (1.31)
65	10000	13.92 (6.05)	95.28 (1.32)
66	10000	14.07 (6.24)	95.26 (1.32)
67	10000	14.10 (6.12)	95.27 (1.23)
68	10000	14.24 (6.19)	95.24 (1.39)
69	10000	14.37 (6.33)	95.24 (1.33)
70	10000	14.48 (6.4)	95.24 (1.33)

Table 5, cont.d

γ	R	T_γ^c ($\sigma_{T_\gamma^c}$)	E_s^γ ($\sigma_{S_s^\gamma}$)
71	10000	14.56 (6.48)	95.22 (1.45)
72	10000	14.65 (6.61)	95.23 (1.34)
73	10000	14.64 (6.49)	95.25 (1.35)
74	10000	14.81 (6.65)	95.21 (1.37)
75	10000	14.93 (6.79)	95.21 (1.41)
76	10000	15.02 (7.05)	95.18 (1.51)
77	10000	15.35 (7.11)	95.19 (1.42)
78	10000	15.32 (7.08)	95.18 (1.45)
79	10000	15.43 (7.06)	95.16 (1.54)
80	10000	15.50 (7.15)	95.18 (1.43)
81	10000	15.84 (7.49)	95.16 (1.47)
82	10000	15.89 (7.50)	95.15 (1.52)
83	10000	16.06 (7.55)	95.16 (1.43)
84	10000	16.14 (7.65)	95.14 (1.46)
85	10000	16.31 (7.82)	95.13 (1.51)
86	10000	16.52 (8.01)	95.14 (1.52)
87	10000	16.82 (8.05)	95.12 (1.54)
88	10000	16.86 (8.29)	95.12 (1.46)
89	10000	17.22 (8.39)	95.13 (1.50)
90	10000	17.42 (8.68)	95.12 (1.50)
91	10000	17.71 (8.85)	95.07 (1.64)
92	10000	17.76 (8.99)	95.07 (1.62)
93	10000	18.11 (9.33)	95.07 (1.59)
94	10000	18.28 (9.3)	95.06 (1.64)
95	10000	18.61 (9.58)	95.07 (1.55)
96	10000	18.57 (9.63)	95.03 (1.74)
97	10000	18.84 (9.75)	95.03 (1.63)
98	10000	19.21 (10.04)	95.03 (1.69)
99	10000	19.30 (10.05)	95.04 (1.60)
100	10000	19.65 (10.52)	95.00 (1.74)

Table 6

First passage through equilibrium varying experimentation process and rates

experimentation γ	1	50	100	140
$\rho_n = 0.033$	903.38 (273.97)	13.48 (5.76)	19.65 (10.52)	34.75 (20.46)
$\rho_u = 0.033$	2556.53 (999.18)	28.59 (23.56)	46.28 (35.39)	83.43 (59.89)
$\rho_n = 0.25$	8661.86 (5652.00)	18.39 (6.77)	25.02 (11.27)	36.35 (18.78)
$\rho_u = 0.25$	dnc*	33.62 (16.39)	54.74 (33.35)	89.49 (58.82)

dnc* - did not converge. Out of $10,000$ runs, only one run converged in less than 20,000 periods

Table 7

First passage varying replication process
with $\rho_u = .033$

replication γ	1	50	100
tournament	2556.53 (999.18)	28.59 (23.56)	46.28 (35.39)
proportional	dnc*	1804 (1851)	1973 (1921)

dnc* - did not converge. Out of $10,000$ runs, only one run converged in less than 20,000 periods; ρ_u = uniform, ρ_n = normal;

Table 8 - First passage through equilibrium for different values of λ for $\rho_u = .033$

λ	$\gamma = 1$	$\gamma = 50$	$\gamma = 100$
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 10

First passage through equilibrium for different values of the size of the set of alternatives, J

	$J = 50$	$J = 100$	$J = 200$
γ	$T^\gamma(\sigma_{T^\gamma})$	$T^\gamma(\sigma_{T^\gamma})$	$T^\gamma(\sigma_{T^\gamma})$
1	1268.70 (457.52)	903.38(273.97)	715.70 (182.35)
50	22.52 (11.69)	13.48(5.76)	9.98 (2.53)
100	29.99 (16.82)	19.65(10.52)	14.78 (6.83)
150	52.13 (31.22)	38.89 (22.81)	30.69 (17.95)

	$J = 500$	$J = 1000$
γ	$T^\gamma(\sigma_{T^\gamma})$	$T^\gamma(\sigma_{T^\gamma})$
1	594.18 (128.86)	543.37 (111.23)
50	8.85 (1.36)	8.68 (1.25)
100	12.41 (4.55)	11.77 (3.91)
150	24.67 (13.89)	22.20 (12.15)

Table 11

Modified vs. Random Initialization

	J = 200		J = 500	
Initialization:	Random	Modified	Random	Modified
$\gamma = 1$	715	645	594	NA
$\gamma = 50$	9.98	9.95	8.85	8.70
$\gamma = 100$	14.78	12.7	12.41	10.65

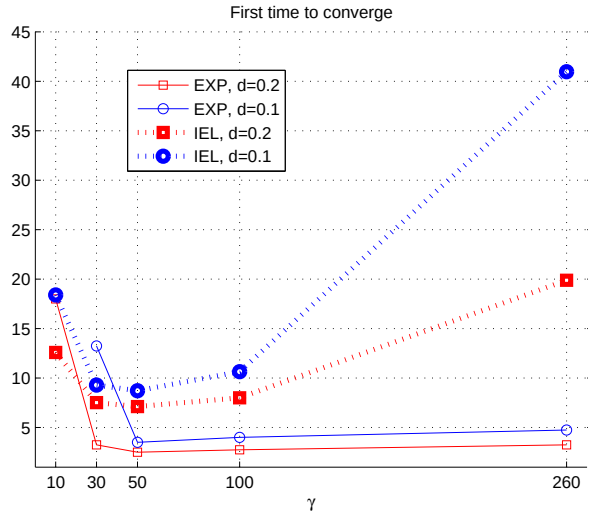

Figure 1: AL convergence data compared to IEL

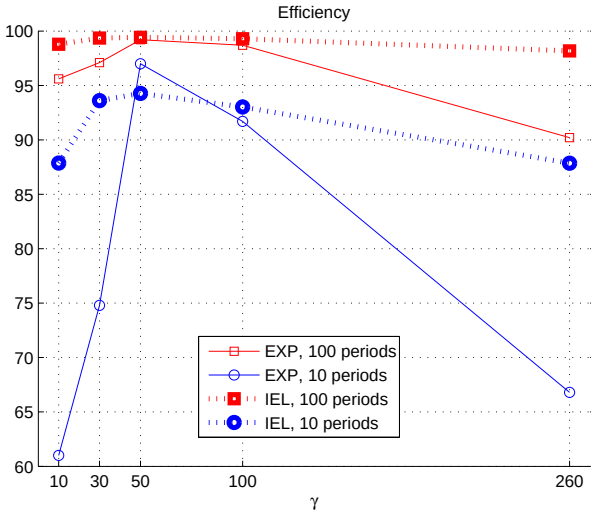

Figure 2: AL efficiency data compared to IEL data

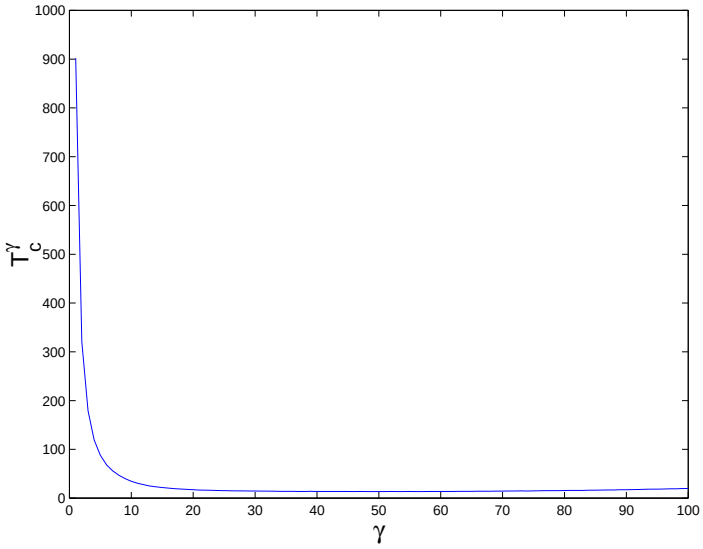

Figure 3: Average time of first passage through equilibrium, $\gamma \in \{1, \ldots 100\}$, $\rho_n = 0.033$

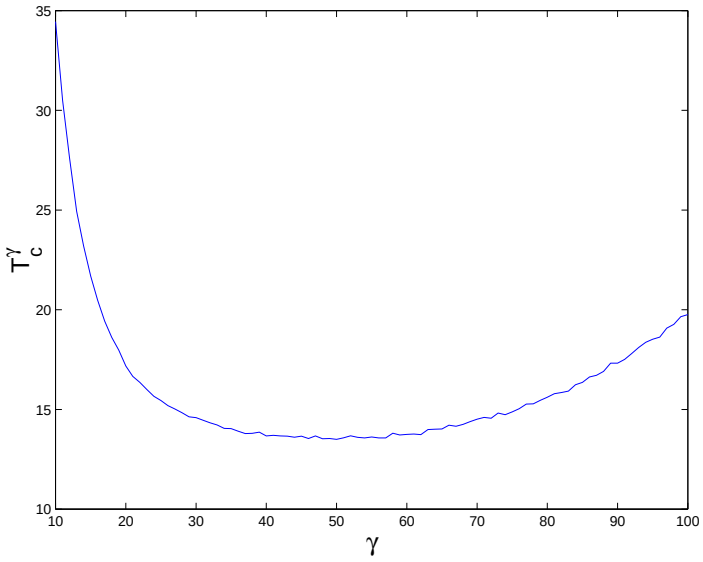

Figure 4: Average time of first passage through equilibrium, $\gamma \in \{10, \ldots, 100\}$, $\rho_n = 0.033$

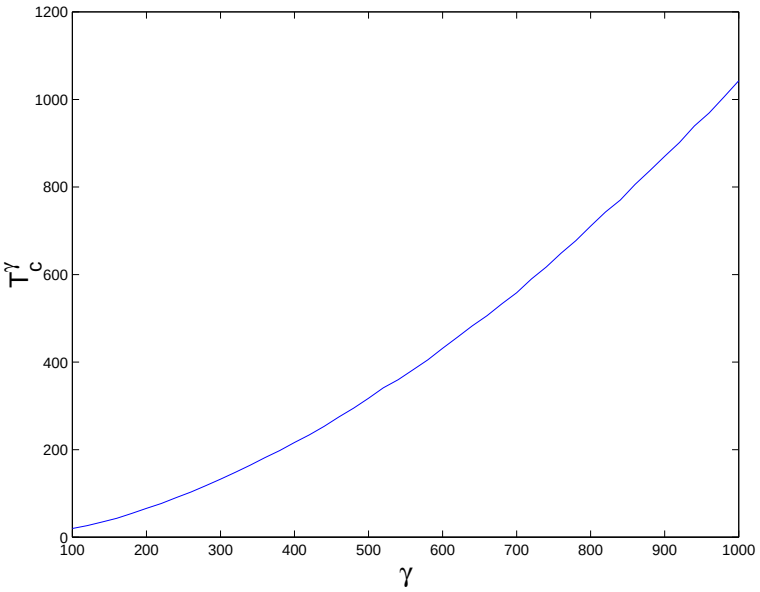

Figure 5: Average time of first passage through equilibrium, $\gamma \in \{100, \ldots 1000\}$, increments of 20, $\rho_n = 0.033$

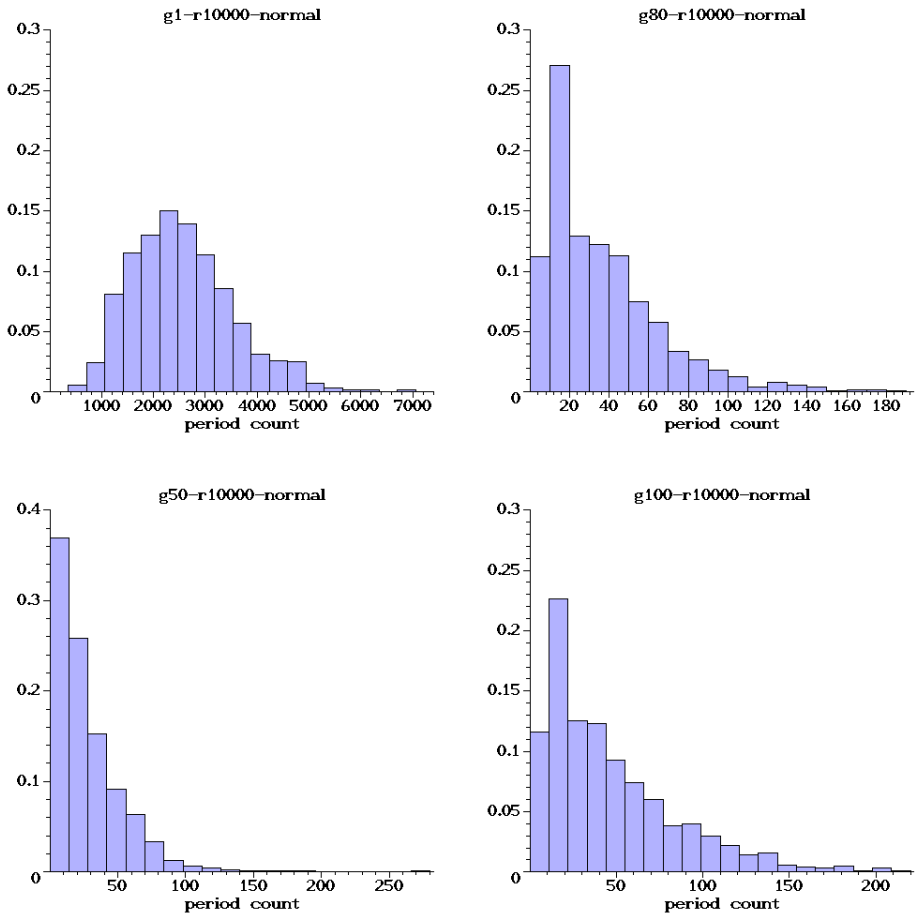

Figure 6: Histogram for runs with $\rho_n = 0.033$,$\gamma = 1$, 50, 80 and 100