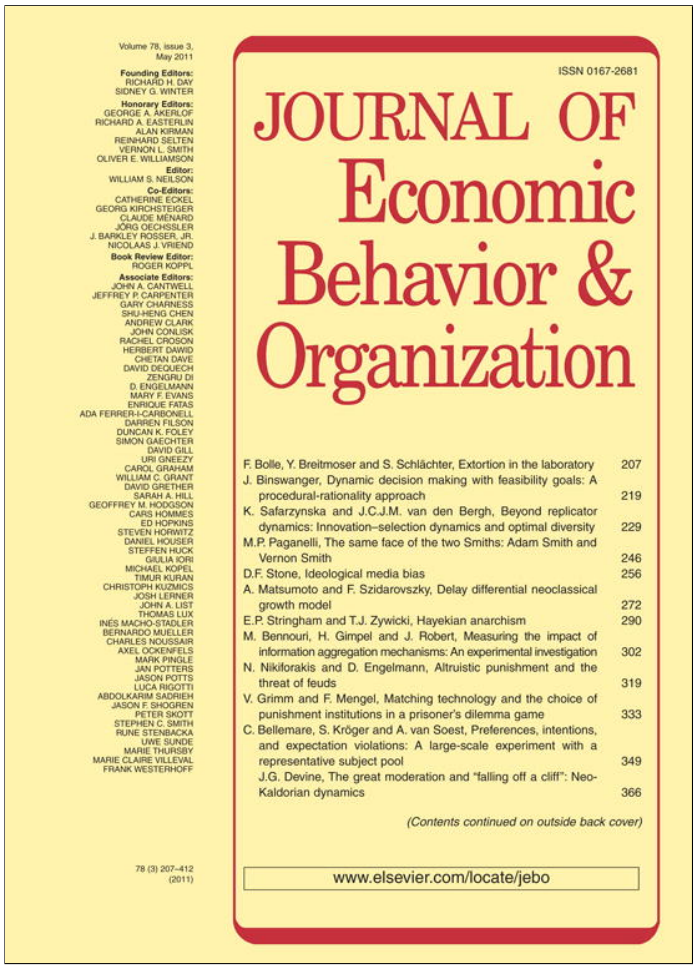

Volume 78, Issue 3,
May 2011

78 (3) 207–412
(2011)

ISSN 0167-2681

JOURNAL OF Economic Behavior & Organization

F. Bolle, Y. Breitmoser and S. Schlächter, Extortion in the laboratory 207
J. Binswanger, Dynamic decision making with feasibility goals: A procedural-rationality approach 219
K. Safarzynska and J.C.J.M. van den Bergh, Beyond replicator dynamics: Innovation–selection dynamics and optimal diversity 229
M.P. Paganelli, The same face of the two Smiths: Adam Smith and Vernon Smith 246
D.F. Stone, Ideological media bias 256
A. Matsumoto and F. Szidarovszky, Delay differential neoclassical growth model 272
E.P. Stringham and T.J. Zywicki, Hayekian anarchism 290
M. Bennouri, H. Gimpel and J. Robert, Measuring the impact of information aggregation mechanisms: An experimental investigation 302
N. Nikiforakis and D. Engelmann, Altruistic punishment and the threat of feuds 319
V. Grimm and F. Mengel, Matching technology and the choice of punishment institutions in a prisoner's dilemma game 333
C. Bellemare, S. Kröger and A. van Soest, Preferences, intentions, and expectation violations: A large-scale experiment with a representative subject pool 349
J.G. Devine, The great moderation and "falling off a cliff": Neo-Kaldorian dynamics 366

(Contents continued on outside back cover)

www.elsevier.com/locate/jebo

Journal of Economic Behavior & Organization 78 (2011) 374–395

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

A behavioral model for mechanism design: Individual evolutionary learning

Jasmina Arifovic[a], John Ledyard[b,*]

[a] Simon Fraser University, Vancouver, British Columbia, Canada
[b] California Institute of Technology, Pasadena, CA, USA

ARTICLE INFO

Article history:
Received 5 April 2010
Received in revised form 16 January 2011
Accepted 25 January 2011
Available online 9 February 2011

JEL classification:
D83
C63
C92
H41

Keywords:
Learning
Experiment
Public goods
Mechanism design

ABSTRACT

We are interested in how Groves–Ledyard mechanisms perform when used repeatedly in a sequence of one-shot games where agents know only their own preferences. In particular, how fast do the mechanisms converge to the stage game Nash equilibrium and how does that speed of convergence depend on the mechanism parameter γ. Prior theoretical and experimental work provide little guidance. Neither do existing behavioral models designed for small games with a small finite number of strategies. For example, even though experience weighted attraction learning is very successful in modeling behavior in one-shot games with very small, finite strategy spaces, it is not successful in modeling behavior in repeated games with a continuum strategy space unless one wants to be involved in fine tuning. We provide a behavioral model, individual evolutionary learning. The time to first convergence is predicted to be smooth and U-shaped in γ. These predictions are robust to a wide range of parameter values. To test the IEL predictions, we ran our own experiments at the California Institute of Technology. Qualitatively, the data from those experiments are consistent with the IEL predictions about convergence and the U-shaped curve. Quantitatively, the human subjects are a little faster, a little less stable, and slightly less efficient than IEL. But for $\gamma = 50$ and 100, the differences between humans and IEL are very small.

1. Introduction

Mechanism design has become very sophisticated since its introduction by Hurwicz (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.

An important question is whether individuals in a group are better off under one mechanism than another. To answer this we must be able to evaluate the performance of alternative mechanisms. To do that we need a model of behavior—how agents choose their actions given the mechanism and the environment. But, there is no single, generally agreed upon standard 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 learning. There are many others.

* Corresponding author. Tel.: +1 626 395 8482.
E-mail addresses: arifovic@sfu.ca (J. Arifovic), jledyard@caltech.edu (J. Ledyard).

0167-2681/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2011.01.021

Absent a unique, compelling model of behavior some economists have turned to 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.

An example of this approach can be found in an interesting paper by Chen and Plott (1996) where they raise a very important mechanism design issue: dynamics in repeated play. Chen and Plott (1996) and Chen and Tang (1998) implement the Groves–Ledyard mechanisms (1977), in a laboratory setting with human subjects, as a repeated stage game over 100 stages. Because the Nash equilibria are Pareto-optimal, the speed of convergence is important for mechanism performance. Faster convergence implies higher welfare. 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. But they left open the question as to which value of the parameter γ (a real number greater than zero) is the best? For which value of γ is convergence the fastest? Of course, running experiments for all values of γ is not an option: it is too expensive. We need a way to determine which limited set of γ we should examine in the lab to determine the best design.

One might look to theory for help, but standard theory is mostly silent on the dynamics of such mechanisms 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 the free parameter. 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. But these theories provide sufficient conditions for asymptotic convergence. Little is known about rates of convergence.

Behavioral theory is also of little help in identifying a best value for γ. Both reinforcement learning, RL, see Roth and Erev (1995), and experience-weighted attraction, EWA, see Camerer and Ho (1999), 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.[1] The strategy space of the Groves–Ledyard mechanism is the real line.

In this paper we present a model we believe has the potential to guide those interested in practical mechanism design to those mechanisms worth testing further in the lab. We propose a model of learning, called individual evolutionary learning (IEL) that deals with large strategy spaces. IEL is based, to some extent, on evolutionary algorithms such as genetic algorithms, classifier systems, and evolutionary programming, but it is different from them. We call it "individual" because each agent evolves their own set of successful strategies instead of there being a social learning process in which successful individuals are evolved.

When we apply IEL to Groves–Ledyard mechanisms, it predicts, contrary to standard approaches, convergence for all values of the parameter. Further it predicts that the time to converge is smooth and U-shaped in the parameter γ. Thus, there is a value of the parameter that minimizes the time to converge and, also, maximizes the efficiency of the chosen allocations. The U-shaped pattern of the convergence times is robust to structural variations and changes in the algorithm's parameter values.

Because this is significantly different from what we expected based on the theoretical analyses, we decided to run several experiments with human subjects for different values of the mechanism parameter. Remarkably, those data confirm the IEL analysis qualitatively. The experimental data replicate the U-shape and the parameter that minimizes the convergence time. But we also found, quantitatively, 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.

IEL was able to identify the values of γ that should be explored in the lab. Further it was exactly right in predicting which value of γ produced the fastest average time to converge and, therefore, the largest average efficiency.

2. A mechanism design problem

The Groves–Ledyard mechanisms were designed to achieve an optimal allocation and financing of public goods in one-shot situations. 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 the one-shot game will be a Pareto optimal allocation. In this section, we provide a quick reminder about public goods environments with quadratic, quasi-linear preferences and about Groves–Ledyard mechanisms. We then turn to the mechanism design problem of interest.

2.1. Environments and mechanisms

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).

[1] See Arifovic and Ledyard (2004). Also see Appendix in this paper.

2.1.1. 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.[2] There are N agents, $i \in \{1, \ldots, N\}$. Let $z = (X, y^1, \ldots, y^N)$ 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 $t = 1, \ldots, t'$, amounts of the public good and net consumptions $(X_t, y_t) \in Z$ are chosen. An agent's payoff from the sequence $(X_1, y_1, \ldots, X_{t'}, y_{t'})$ 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.

2.1.2. 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, $M \subset (-\infty, +\infty)$. $m^i \in M$ can be thought of as i's incremental demand for the public good.

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

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

The tax to be paid by i is:

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

where γ is an arbitrary free parameter greater than 0, $\mu_{-i} = (\sum_{h \neq i} m^h)/(N-1)$ is the mean value of the messages of the other agents, and $\sigma_{-i}^2 = (\sum_{h \neq i} (m^h - \mu_{-i})^2)/(N-2)$ is the squared deviation from this mean.

The outcome rule of the mechanism is:

$$g(m, \gamma) = (X(m), y^1(m, \gamma), \ldots y^N(m, \gamma)) = (X(m), -T^1(m, \gamma) \ldots, -T^N(m, \gamma))$$

Let $r^i = r^i(m) = (\mu_{-i}(m), \sigma_{-i}^2(m))$. 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

$$W^i(m^i | r^i(m), \gamma) = V^i(m^i + (N-1)r_1^i(m), \alpha^i - T^i(m^i, r^i(m), \gamma)) \quad (1)$$

Different values of γ imply different outcome functions and, therefore, different mechanisms. Letting γ range over values from 0 to ∞, one creates a class of mechanisms.

We know from the theorem of Groves and Ledyard (1977) that the Nash equilibrium allocations of the mechanisms are Pareto-optimal, no matter what the value of γ is. In environments with quasi-linear preferences, the Pareto-optimal level of the public good is unique. Therefore, the Nash equilibrium outcome level of the public good is independent of γ.

Theorem 1. *For public goods environments with quasi-linear preferences, if m^* is a vector of messages such that $m^{*i} \in$ argmax $W^i(m^i \mid r^i(m^*), \gamma)$ for all i, then $[X(m^*), y(m^*)] \in$ argmax $\sum V^i(X, y^i)$ subject to $cX + \sum y^i = 0$.*

2.2. Mechanism dynamics

We are interested in how Groves–Ledyard mechanisms perform when used repeatedly in a sequence of one-shot games where agents know only their own preferences. Under these conditions, human subjects do not instantaneously play Nash equilibrium strategies. Instead messages converge to equilibrium over time and get close in a finite amount of time. As Healy (2006, p. 114) points out, "In a repeated-interaction public goods economy, incomplete information and dynamic behavior may affect the realized outcomes of mechanisms known to be efficient in a complete information one-shot game." If one is interested in actually using these mechanisms repeatedly, it is necessary to understand their dynamics to understand their potential performance. Faster convergence to the Nash equilibrium implies higher average welfare over time. We need to know not only whether strategies converge as $t \to \infty$ but how fast they converge. From a mechanism design point of view, we want to determine for which γ convergence is fastest and efficiency is highest.

[2] 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.

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

Let $m(t)$, be the sequence of strategies generated by repeated play of some mechanism under some behavior. How soon do the equilibrium strategies get close to the equilibrium strategies? Our concept of closeness is $|m^i(t) - m^{ie}| \leq d$ for all i, where m^{ie} is the Nash equilibrium message of i. We call this the *convergence criterion*. The period when the convergence criterion is first fulfilled is called *the time of the first passage through equilibrium*. For a particular mechanism γ, the time of first passage through equilibrium is

$$T(\gamma) = \min t \text{ subject to } |m^i(t, \gamma) - m^{ie}| \leq d$$

We want to understand how $T(\gamma)$ varies with γ. Is $T(\gamma)$ continuous in γ? If so is there a specific pattern? Or, is there a γ' such that $T(\gamma) = \infty$ if $\gamma < \gamma'$ and $T(\gamma) \leq T^* < \infty$ if $\gamma \geq \gamma'$? Is there a γ^* that yields the best performance? Is there a γ^* such that $T(\gamma^*) = \min_\gamma T(\gamma)$?

In this paper, we provide some answers to these questions.

3. Traditional approaches

In this section, we explore what we might learn about mechanism dynamics from previous research. We look at standard theory, experiments, and a leading learning model.

3.1. Standard theory

Standard theory is mostly silent on the dynamics of Groves–Ledyard mechanisms.[3] Two exceptions are the papers by Muench and Walker (1979, 1983) and by Chen and Tang (1998), both of which suggest that the parameter γ plays a major role in the dynamics in quadratic environments. Muench and Walker (1983) base their analysis on Cournot–Nash best response dynamics and focus on large economies. They show two types of limiting results: (1) for a fixed γ, as $N \to \infty$, the best reply dynamics become unstable,[4] and (2) if one reacts to this by setting $\gamma \geq N$, then the dynamics are always stable but as $N \to \infty$, the parameter $\gamma \to \infty$ and so there becomes little incentive to tell the truth in large populations.[5]

Chen and Tang (1998) base their analysis on more general dynamics from adaptive learning[6] found in the work of Milgrom and Roberts (1990) on strategic complementarities. They show that the supermodularity condition, $\partial^2 W^i / \partial m^i \partial m^j \geq 0$, is a sufficient condition for the convergence of the mechanism under adaptive learning in a sequence of repeated one-shot games.[7] For quadratic preferences and Groves–Ledyard mechanisms, the condition is $\gamma \geq 2NB^i$ for all i.

Another sufficient condition for global convergence to Nash equilibrium, if agents use best replies, can be derived from a theorem of Gabay and Moulin (1980) using a dominant diagonal condition, $|\partial^2 W^i / \partial m^i m^i| > \sum_{i \neq j} |\partial^2 W^i / \partial m^i m^j|$. For quadratic preferences and Groves–Ledyard mechanisms the condition is $\gamma \geq [N(N-2)/(N-1)]B^i$ 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.

3.2. Experiments

Prior experiments with human subjects have not shed very much light on the dynamic properties of the Groves–Ledyard mechanism.[8] The key results are reported in Chen and Tang (1998), which expands on Chen and Plott (1996). In each of the Chen and Tang experiment sessions, 5 subjects participated in a Groves–Ledyard mechanism for 100 rounds. Subjects were given quasi-linear, quadratic preferences. We list the parameters for these utility functions in Table 1. The message space was $[-4, 6]$. A total of 14 experimental sessions were run: 7 each with $\gamma = 1$ and $\gamma = 100$.

[3] See Healy (2006) for a good summary of the theory covering the dynamics of mechanisms for public goods economies.
[4] In fact what is true is that as $\gamma/N \to 0$, the dynamics become unstable.
[5] We recommend the interested reader see Muench and Walker (1979). There they carry out a deep and revealing analysis of the trade-off between existence and optimality, stability, strength of preferences, and manipulability.
[6] Adaptive learning is defined in Milgrom and Roberts (1990) and includes Cournot best response, fictitious play, Bayesian learning and others.
[7] This is not as nice a result as one might imagine. The Milgrom-Roberts result is that supermodularity guarantees convergence of adaptive behavior to a vector of strategies that lies between the largest and smallest Nash Equilibrium. As Healy and Mathevet (2009) point out, if the strategy spaces, M^i, are compact (as is usually the case) then there can be boundary equilibria. So, in many cases convergence is guaranteed only to somewhere in M, not much of a restriction.
[8] See Chen (2003) for a good summary of experiment covering mechanisms for public goods economies.

Table 2
Chen and Tang experimental results.

γ	$T(\sigma_T)$	$S^a(\sigma_S)$	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

For $d = 0.2$. NA(7/7) – 7 of the 7 sessions did not converge in 100 rounds.

Using their data,[9] we have computed time to converge, $T(\gamma)$, using a convergence criterion of $d = 0.2$.[10] A summary of these computations is presented in Table 2. For $\gamma = 1$, the choices of the agents did not converge to the equilibrium within 100 experimental periods for any of the 7 sessions that were conducted. For $\gamma = 100$, convergence did occur. The average time to converge was 9.0 with a standard deviation of 5.7. It is, however, not possible to determine from these experimental results whether the cut points, $\gamma \geq 80$ provided by strategic complementarity or $\gamma \geq 30$ provided by the dominant diagonal condition, were a determining factor in the dynamics.

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* the system is after the first passage. If messages get close to equilibrium, do they stay close to equilibrium? To track this, 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. For a given mechanism γ, let $S^a(t, \gamma)$ be an index variable that equals 1 if $|m^i(t, \gamma) - m^{ie}| \leq d$ for all i and otherwise equals 0. For a given mechanism γ, the index of equilibrium stability in actions is given by

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

The computations of our index of stability,[11] $S^a(\gamma)$, for the Chen–Tang data can be found in Table 2. For $\gamma = 100$ the seven Chen and Tang experiments yield an average stability of 72. That is, slightly more than 25 percent of messages after convergence are not within the convergence criterion. Maybe things have not "converged".

In the end, the main reason we are interested in time to converge and stability is that faster convergence and higher stability imply higher average welfare over time. We use an *index of efficiency*, E, to measure welfare. It 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 Nash equilibrium messages, m^e. Our measure of efficiency over T periods is:

$$E^T(\gamma) = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N} W^i(m^i(t, \gamma)|r^i(m(t)), \gamma)}{T\sum_{i=1}^{N} W^i(m^{ei}|r^i(m^e), \gamma)}$$

We computed average efficiency for the first 100 periods, E^{100}, and for the first 10 periods, E^{10}, in the Chen and Tang experiments. For the mechanism with $\gamma = 100$, human subjects achieved 98.1 percent on average for the first 100 rounds. For just the first 10 periods, human agents achieved 93 percent. These are very high numbers considering that convergence to equilibrium did not occur on average for 9 periods and stability was at 74. It is also interesting to note that average efficiencies over 100 rounds are also high for the mechanism with $\gamma = 1$. Even though convergence has not occurred by 100 rounds, humans achieve 93.4 percent. Of course, the efficiencies are much lower for the first 10 rounds at 80 percent, as one might expect with slow convergence.

3.3. A classic learning model—EWA

It is clear, from both the Chen and Tang (1998); Healy (2006) data that human subjects are learning as they play. They do not instantaneously play the one-shot Nash Equilibrium strategies but they do seem to try to learn them. Sometimes they

[9] Their data can be found at http://www.si.umich.edu/ yanchen/data/jpe98data/jpe98data.htm. Similar results are reported by Healy (2006) for $\gamma = 100$.

[10] In their experiments, Chen and Tang allowed only integer messages from their subjects. To implement this without having 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 only messages in multiples of 0.2 in our simulations. Using $d = 0.2$ requires messages in Chen and Tang to converge within 1 of the equilibrium. We have used other values of d and have found little difference in the qualitative properties of the results.

[11] Because these experiments only last 100 rounds, we take the remaining number of periods of a particular session once the first passage through equilibrium is achieved and compute the percentage of those messages that are near equilibrium messages.

Table 3
Simulations experience weighted attraction learning $\lambda = 0.35$, $N_0 = 10$, $\delta = 0.96$, $\rho = 0.95$, $\phi = 0.991$, $d = 0.2$.

γ	R	$T(\sigma_T)$	S^a (σ_S^a)
1	0	dnc	dnc
10	2431	56.35 (278.84)	15.08 (25.59)
20	5527	45.34 (195.54)	23.17 (32.02)
30	9014	26.56 (135.78)	46.74 (36.91)
40	9007	28.78 (135.98)	49.94 (36.62)
50	9013	28.57 (166.74)	53.22 (37.08)
60	8927	27.88 (204.59)	56.54 (38.74)
70	8102	36.63 (309.00)	48.21 (40.04)
80	6755	42.02 (374.63)	46.21 (40.62)
90	5066	57.92 (463.45)	45.99 (40.52)
100	4216	70.61 (566.82)	46.81 (40.64)
110	3975	41.34 (378.15)	52.45 (42.16)
120	3038	50.33 (441.12)	48.73 (41.11)
130	2968	43.62 (427.49)	48.25 (42.36)
140	2464	63.81 (511.78)	48.89 (41.96)
150	2240	67.79 (526.67)	50.02 (42.02)

R– number of simulation runs, out of total of 10,000 runs, that converged within 10,000 periods.

do not succeed, when $\gamma = 1$. But when they do, for $\gamma = 100$, they converge on the Nash equilibrium in an average of 9 periods. Can standard learning models explain these dynamics?

To answer that question we turned to the leading example of a successful behavioral model of learning, the Experience Weighted Attraction model of Camerer and Ho (1999).[12] EWA has been shown to be an accurate predictor of behavior in one-shot games with a small finite number of strategies. We wondered how it would do in repeated games with a continuum for a strategy space.

3.3.1. Setup

We simulated what EWA agents would do in repeated play of Groves–Ledyard mechanisms for a variety of values of γ. Each value of γ determines a different mechanism. We used $\gamma \in \{1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150\}$. We call a particular simulation, a *run*. Each run is based on an environment, a behavioral model, and a mechanism. For each value of γ, we conducted 10,000 runs, using different random number generator seed for each run. Each run lasted 10,000 periods (unless it was terminated earlier as the convergence occurred). Each run was terminated 100 periods after a *convergence criterion* was fulfilled. We used a convergence criterion of $d = 0.2$.

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. In order to implement EWA, as for any model designed to deal with finite strategy sets, we had to discretize the message space. We followed Chen and Tang (1998) and constructed a set of 51 messages,[13] equally spaced within the range of $[-4, 6]$.

We used the EWA behavioral model described in Appendix I. We conducted 8 simulations, each with a different set of parameters for EWA. For the runs reported in this section, we used the parameter set: $\lambda = 0.35$, $N_0 = 10$, $\delta = 0.96$, $\rho = 0.95$, and $\phi = 0.991$.[14] This is the set that gave the best results: faster convergence and higher stability.

3.3.2. Results

We report the results in Table 3. Column 1 gives the value of γ, column 2 reports on the number of simulations, R (out of 10,000) that converged within 10,000 periods, column 3 presents the average values of times of first passage through equilibrium (averaged over the R runs that converged), $T(\gamma)$, and the values of standard deviations, $\sigma_T(\gamma)$, in the parenthesis, and finally column 4 reports on our measure of equilibrium stability, $S^a(\gamma)$.

The key finding is that there is neither much convergence nor much stability. There is no real pattern to $T(\gamma)$. For $\gamma = 1$, no convergence occurs over 10,000 periods in any of the 10,000 simulations. There is no value of γ that results in convergence of all of the 10,000 simulations. The percentage of simulations that converge varies from lows of 22–24 percent for $\gamma = 10$ and 150 to highs around 90 percent for $\gamma = 30$–60. Because the variances are high and convergence is rare, it is impossible to pin down any systematic relationship in $T(\gamma)$.

There is also very little stability. Even when the process manages to get near to a Nash-equilibrium it does not stay there. The measure, that we use to study the stability of equilibrium after the first passage, takes a low value of 15 percent for $\gamma = 1$.

[12] We have provided, in Appendix I, the formal structure of EWA for those who might not remember it.
[13] See footnote 11.
[14] 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 parameter sets for EWA in Arifovic and Ledyard (2004).

For $\gamma > 10$, the measures stays around 50 percent. In the best of circumstances, the process remains in the neighborhood of the equilibrium less than 56 percent of the time. It never really settles down.

There is also not much correlation between the EWA simulations for $\gamma = 100$ and the Chen and Tang experiment results. Chen and Tang's subjects converged every time with an average time to converge of 9 periods with a 5.7 standard deviation. EWA's subjects converged only 70 percent of the time. When they did converge it took on average 70 periods with a huge standard deviation of 566. Clearly EWA is not the right model for human subject behavior in the repeated Grove-Ledyard mechanisms.

It is worth pointing out that the problem of convergence of EWA (as well as reinforcement learning) could be due to how the parameter λ is set.[15] If one allows for a time varying value of λ such that λ becomes larger with time, these models will converge. The difficulty however is choosing the function $\lambda(t)$ which depend on how the strategy space is discretized and on the particular payoff functions for the experiment. This requires a level of tuning that IEL does not.

3.4. Summary to here

We are interested in how Groves–Ledyard mechanisms perform when used repeatedly in a sequence of one-shot games where agents know only their own preferences.

(1) There are theoretical conditions that are sufficient for convergence, dominant diagonal and strategic complementarity, but these conditions provide no information about rates of convergence.
(2) Prior experimental work merely separates $\gamma = 1$ from $\gamma = 100$: convergence does not happen in the former but does in the latter. Nothing is known about other values of γ.
(3) Even though EWA is very successful in modeling behavior in one-shot games with very small, finite strategy spaces, it is not successful in modeling behavior in repeated games with a continuum strategy space unless one wants to be involved in fine tuning.

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? 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?

To provide answers, we need a better model.

4. The IEL approach

An experiment can usually be modeled as a repeated game (G, R). The repeated game has a stage game G and a number of rounds, T. The idea is that G will be played for T rounds. In $G = \{N, X, u, r\}$, N is the number of subjects, X^i is the action space of i, $u^i(x^1, \ldots, x^N)$ is i's payoff if the joint strategy choice is x, and $r^i(x_t)$ describes the information reported to subject i at the end of a round. These are all controlled by the experimenter. In round t, each subject chooses $x_t^i \in X^i$. At the end of round t, subject i will be told the information $r^i(x_t)$ about what happened. Then the next round will be played. A behavioral model must explain how the sequence of choices for i, $(x_1^i, x_2^i, \ldots, x_R^i)$ is made, given what i knows at each round t.

Our approach to modeling behavior is based, to some extent, on evolutionary algorithms.[16] 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, 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 individual approach and build on the framework of Arifovic (1994).

In our model, which we call individual evolutionary learning (IEL), each agent retains their own finite set of remembered strategies. After each iteration, they update this set through experimentation and replication. Experimentation involves replacing, with low probability, some of the elements of the set with a strategy chosen at random from the entire strategy space. Experimentation introduces strategies that might otherwise never have a chance to be tried. Replication goes through the set of remembered strategies and, through a series of random paired comparisons within the set, replaces the strategies which would have provided a low payoff in the previous period with copies of those that would have yielded a higher payoff. Over time the remembered set becomes homogeneous with copies of the "best reply" strategy. To generate their strategic choices, individuals choose strategies from the set of remembered strategies, at random, proportionately to the payoff they would have received had they been played in the last round. The end result of IEL is a Markov process in mixed strategies

[15] We thank an anonymous referee for pointing this out to us.
[16] For surveys, see Arifovic (2000) for applications to macroeconomic models, see LeBaron (2000) for applications in finance, and see Dawid (1999) for general overview of applications in economics and game theory.

where the remembered set of each agent, the state space of the process, co-evolves with the sets of all other agents. IEL is particularly well-suited to repeated games with large strategy spaces such as subsets of the real line.

4.1. Individual evolutionary learning

The primary variables of our behavioral model are a finite set of remembered strategies for each agent i at each round t, $A_t^i \subset X^i$ and a probability measure, π_t^i on A_t^i. A_t^i consists of J alternatives. J is a free parameter of IEL that can be loosely thought of as a measure of the processing and/or memory capacity of the agent. In round t, each agent selects an alternative randomly from A_t^i using the probability density π_t^i on A_t^i and then chooses the action $x_t^i = a_t^i$. One can think of (A_t^i, π_t^i) as inducing a mixed strategy on X^i at t. At the end of each round t, agents are told $r(x_t)$. At the beginning of 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 three pieces: experimentation, replication, and selection.

We begin at the end of round, t, knowing A_t^i, π_t^i, and $r^i(x_t)$.

Experimentation comes first. Experimentation introduces new alternatives that otherwise might never have a chance to be tried. This insures that a certain amount of diversity is maintained. For each $j = 1, \ldots, J$, with probability ρ, a new contribution is selected at random from X^i and replaces $a_{j,t}^i$. We use a normal density, conditional on X^i, for this experimentation. For each j, the mean value of the normal distribution is set equal to the value of the alternative, $a_{j,t}^i$ that is to be replaced by a 'new' idea. That is the new alternative $a \sim N(a_{j,t}^i, \sigma)$. ρ and σ are free parameters of the behavioral model that can be varied in the simulations.

Replication comes next. Replication reinforces strategies that would have been good choices in previous rounds. It allows potentially better paying strategies to replace those that might pay less. The crucial assumption here is the measure of "potentially better paying strategies". We let $W^i(a_{jt}^i | r^i(x_t))$ be the forgone utility of alternative j at t given the information $r^i(x_t)$, where W^i comes from Eq. (1). This measures the utility i thinks she would have gotten had she played a_j in round t. $W^i(a_j \,|\, r^i(x_t))$ is a counter-factual valuation function and must be specified for each application.

Given a forgone utility function, W^i, we can describe how replication takes place. For $j = 1, \ldots, J$, $a_{j,t+1}^i$ is chosen as follows. Pick two members of A_t^i randomly (with uniform probability) with replacement. Let these be $a_{k,t}^i$ and $a_{l,t}^i$. Then

$$a_{j,t+1}^i = \left\{ \begin{matrix} a_{k,t}^i \\ a_{l,t}^i \end{matrix} \right\} \text{ if } \left\{ \begin{matrix} W^i(a_{kt}^i | r^i(x_t)) \geq W^i(a_{lt}^i | r^i(x_t)) \\ W^i(a_{kt}^i | r^i(x_t)) < W^i(a_{lt}^i | r^i(x_t)) \end{matrix} \right\}$$

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

Selection is last. Each contribution $a_{k,t+1}^i$ is selected with the following probability[17]:

$$\pi_{k,t+1}^i = \frac{W^i(a_{k,t+1}^i | r^i(x_t)) - \varepsilon_{t+1}^i}{\sum_{j=1}^{J} (W^i(a_{j,t+1}^i | r^i(x_t)) - \varepsilon_{t+1}^i)}$$

for all $i \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, J\}$ and where, to avoid negative probabilities,[18]

$$\varepsilon_{t+1}^i = \min_{a \in A_{t+1}^i} \{0, W^i(a | r^i(x_t))\}$$

We now have a complete description of the way that an IEL agent moves from A_t^i and π_t^i to A_{t+1}^i and π_{t+1}^i. The only remaining things to pin down are the initial values, A_1^i and π_1^i.

One way to start would be with random selection. However, subjects begin the experiment with information about their payoff functions and the game. They also have time to think before the first round about strategies that might work. So we use a more sophisticated initialization process.

---

[17] An alternative selection model would change the probabilistic choice function to $\pi(a^k) = (e^{\lambda u^i(a^k)})/(\sum_j e^{\lambda u^i(a^k)})$. We have found (see Arifovic and Ledyard, 2004) that the behavior predicted for any λ differs very little from that generated by our proportional selection rule. This is because the set A tends to become homogeneous fairly fast, at which point the selection rule is irrelevant. We therefore use the proportional rule since it eliminates another parameter.

[18] 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.

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} and σ_{jk} (within the feasible range). The payoff of message j is then calculated as $v_j = (1/100)\sum_{k=1}^{100}[W^i(m^i_j|r^i(\mu_{jk}, \sigma_{jk}))]$. 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 (level 1) cognitive hierarchy model (Camerer et al., 2004).

4.2. Some remarks

4.2.1. 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, 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 models of learning, such as EWA, the continuum must be discretized. However, discretization causes problems when there are very fine differences in equilibrium values between different mechanisms. (Note that this remains a problem even if we model λ as a function of time.) 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.

4.2.2. Free parameters

Our model is not entirely determined a priori. The IEL model requires the choice of one set, one function, 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 $r^i = r^i(m^{-i})$. The entire interplay between environment, mechanism, and behavior rests on this computation of hypotheticals.[19]

The two parameters are J, the size of the set A, and ρ, the rate of experimentation. They are the only free parameters of this model. In Section 5.4, we will examine what happens to the performance of the mechanisms as we vary the parameters. 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.

4.2.3. Experimentation

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 the appropriate one 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 due to replication only those newly generated alternatives that appear promising are actually tried out. This is different from the experimentation or mutation traditionally discussed in the literature on learning or evolutionary game theory.

As pointed out earlier, 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 hypothetical payoff 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.[20]

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.

---

[19] This is a 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.

[20] This actually can happen fairly fast. A rough, approximate calculation using expected values, based on the assumptions that $\rho = 0.03$ (and is uniform) and that initially about 1 percent 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.

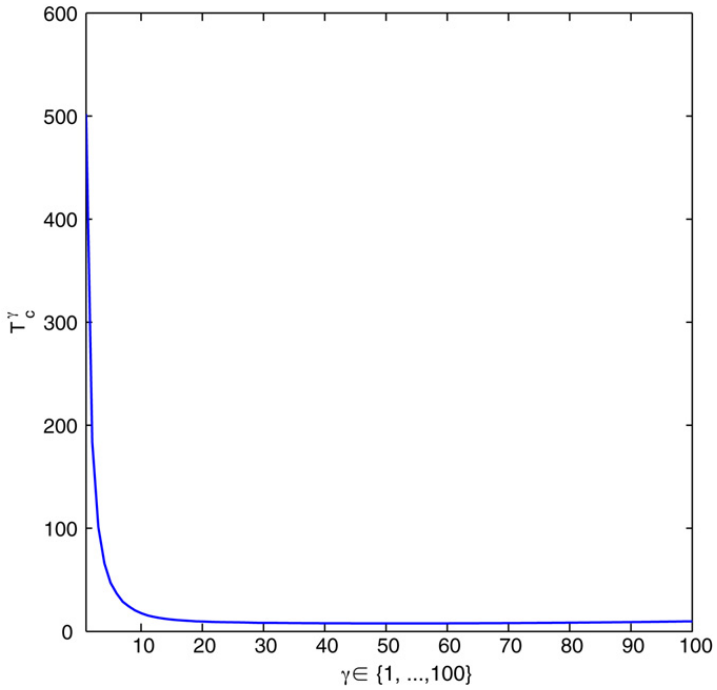

Fig. 1. Average time of first passage through equilibrium, $\gamma \in \{10, \ldots, 100\}$.

4.2.4. Selection

We have chosen to use a particularly simple selection process so as to avoid adding more parameters to the model. As we indicated in the previous paragraph, the set of remembered alternatives, A_t^i in IEL becomes homogeneous very quickly. Once any level of homogeneity is attained, selection becomes irrelevant: if all alternatives in A are the same, it does not matter at all which is chosen. Thus, especially after a few rounds, the particular method of selection is not important.

It is possible that a tighter selection process, such as selecting the alternative with the highest hypothetical utility (the best reply) or probabilistically selecting in proportion to the hypothetical utility (Quantal response), might change how IEL behaves in early rounds of a repeated game. But, as we will see below, the proportional selection process we have chosen fits the data and does not require additional parameters.

5. Simulations with IEL

What does IEL predict about the dynamics of behavior in repeated games with the Groves–Ledyard mechanisms. To answer that we run a number of simulations: laboratory experiments using IEL instead of human agents. We uncover a systematic relationship in $T(\gamma)$.

5.1. Setup

We simulated what IEL agents would do in repeated play of Groves–Ledyard mechanisms for a variety of values for γ. Each value of γ determines a different mechanism. We used $\gamma \in \{1, \ldots, 100\}$ and $\gamma \in [120, 1000]$ in increments of 20. We call a particular simulation, a *run*. Each run is based on an environment, a behavioral model, and a mechanism. For each mechanism, γ, we implemented 1000 runs. Each run was terminated 100 periods after a *convergence criterion* was fulfilled. For these simulations we let $d = 0.1$. The maximum number of periods for each run was set at $t_{\max} = 10,000$. If the convergence criterion was not fulfilled by that time, a run was terminated.

To be able to compare our results to those generated with humans, we used the utility functions, cost of production, and strategy spaces in Chen and Tang (1998). The cost, c, of producing a unit of the public good, which determines Z, is set to 100. The utility parameters are given in Table 1. The strategy space is $M = [-4, 6]$.

We used the IEL behavioral model described in the previous section. The foregone utility function is W from equation (1). For the runs reported in this section, we set the memory capacity $J = 500$. We set the rate of experimentation $\rho = 0.033$. In Section 5.4, we will indicate why we have chosen these numbers and examine what happens if we change them.

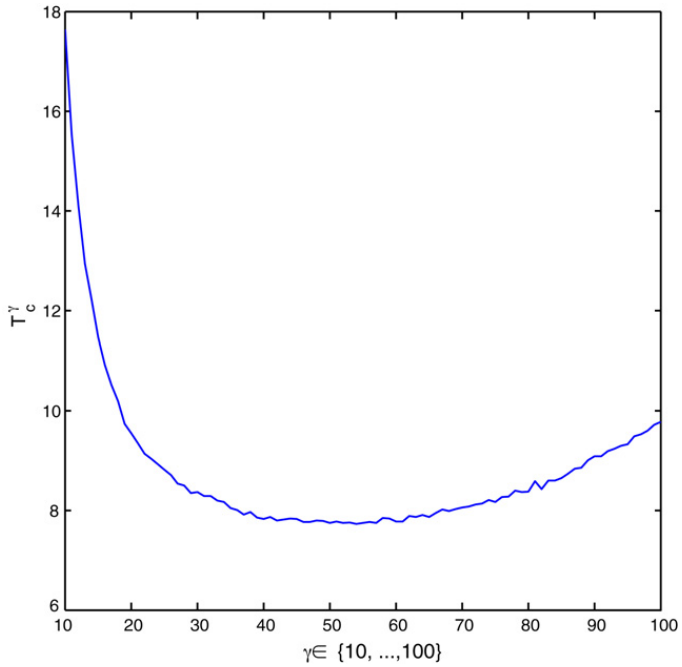

Fig. 2. Average time of first passage through equilibrium, $\gamma \in \{10, \ldots, 100\}$.

5.2. Results

In Fig. 1, we present, graphically, the average time to first passage through equilibrium for $\gamma \in \{1, \ldots 100\}$. In Fig. 2 are the data for $\gamma \in \{10, 100\}$ and in Fig. 3 are the data for γ between 120 and 1000 in the increments of 20. All runs converge in finite time for all γ.

Table 4 contains detailed data for the average time of first passage through equilibrium for $\gamma \in \{1, \ldots 100\}$. The first column gives the value of γ, the second column presents the average values of times of first passage through equilibrium (averaged over 1000 runs), $T(\gamma)$, and the values of standard deviations, $\sigma_T(\gamma)$, in the parenthesis. In the third column are the values for a modified stability measure based on strategies. In Figs. 4–9, we exhibit the densities of the time to converge for $\gamma \in \{1, 10, 30, 50, 100\}$. With the exception of $\gamma = 1$ and $\gamma = 260$, these are remarkably concentrated.

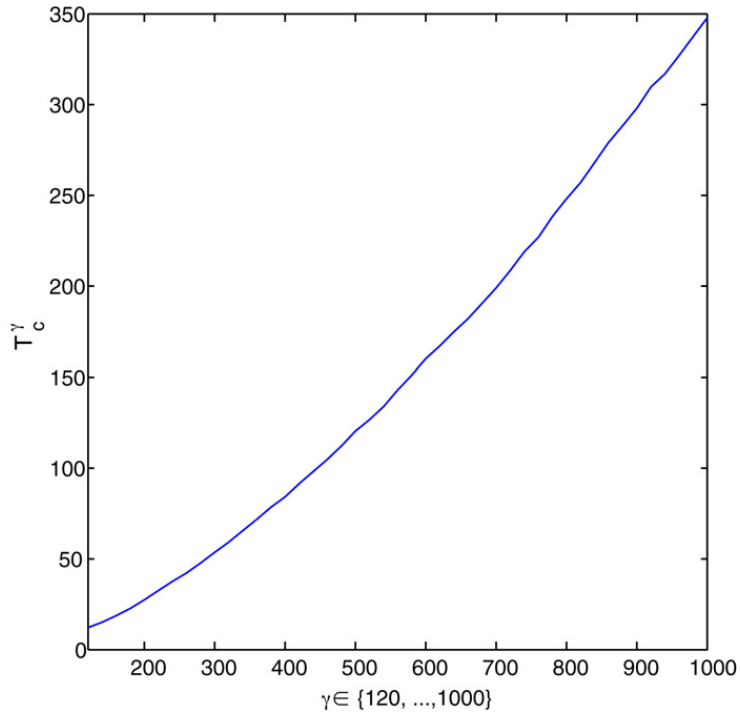

Fig. 3. Average time of first passage through equilibrium, $\gamma \in \{100, \ldots 1000\}$, increments of 20, $\rho_n = 0.033$.

Table 4
$J = 500$, modified initialization, 1000 runs.

γ	$T(\gamma)$	σ_T	S^s	σ_S^s
1	501.93	96.28	98.35	3.63
2	183.87	52.53	98.97	2.08
3	101.24	33.48	99.22	1.53
4	65.92	23.40	99.42	0.92
5	47.14	18.10	99.53	0.71
6	36.77	14.58	99.56	0.65
7	28.77	11.86	99.59	0.55
8	24.22	9.73	99.62	0.52
9	20.33	7.93	99.66	0.37
10	17.64	6.56	99.68	0.32
11	15.53	5.39	99.70	0.31
12	14.11	4.44	99.69	0.33
13	12.94	3.88	99.73	0.22
14	12.22	3.38	99.73	0.21
15	11.47	2.99	99.73	0.22
16	10.92	2.55	99.76	0.19
17	10.52	2.32	99.77	0.17
18	10.19	2.12	99.78	0.17
19	9.74	2.10	99.77	0.17
20	9.55	1.81	99.79	0.15
21	9.35	1.73	99.80	0.15
22	9.14	1.59	99.80	0.14
23	9.04	1.50	99.81	0.12
24	8.93	1.47	99.81	0.13
25	8.82	1.45	99.81	0.13
26	8.71	1.41	99.81	0.14
27	8.54	1.42	99.81	0.13
28	8.50	1.34	99.82	0.12
29	8.35	1.33	99.82	0.11
30	8.37	1.31	99.82	0.12
31	8.29	1.25	99.82	0.11
32	8.29	1.32	99.83	0.11
33	8.20	1.28	99.82	0.12
34	8.17	1.26	99.83	0.09
35	8.05	1.23	99.82	0.11
36	8.01	1.26	99.82	0.12
37	7.92	1.18	99.82	0.12
38	7.97	1.20	99.83	0.11
39	7.86	1.22	99.82	0.11
40	7.83	1.26	99.82	0.14
41	7.87	1.21	99.83	0.12
42	7.80	1.21	99.83	0.11
43	7.82	1.20	99.83	0.11
44	7.84	1.20	99.83	0.10
45	7.83	1.18	99.83	0.10
46	7.77	1.22	99.82	0.12
47	7.77	1.22	99.83	0.11
48	7.80	1.19	99.83	0.11
49	7.79	1.15	99.83	0.10
50	7.75	1.23	99.83	0.11
51	7.78	1.23	99.83	0.10
52	7.75	1.23	99.82	0.13
53	7.76	1.19	99.83	0.10
54	7.73	1.20	99.83	0.12
55	7.75	1.22	99.82	0.11
56	7.77	1.24	99.82	0.12
57	7.75	1.27	99.82	0.12
58	7.85	1.26	99.83	0.11
59	7.84	1.25	99.83	0.11
60	7.78	1.29	99.82	0.11
61	7.78	1.31	99.82	0.11
62	7.89	1.29	99.83	0.10
63	7.87	1.36	99.82	0.13
64	7.91	1.30	99.82	0.12
65	7.87	1.30	99.82	0.11
66	7.95	1.33	99.82	0.11
67	8.02	1.42	99.82	0.12
68	7.99	1.42	99.82	0.13

Table 4 (*Continued*)

γ	$T(\gamma)$	σ_T	S^s	σ_S^s
69	8.03	1.37	99.82	0.12
70	8.06	1.40	99.82	0.11
71	8.08	1.50	99.82	0.12
72	8.12	1.50	99.82	0.13
73	8.14	1.50	99.82	0.11
74	8.21	1.62	99.82	0.11
75	8.17	1.52	99.81	0.13
76	8.27	1.56	99.82	0.12
77	8.28	1.60	99.81	0.12
78	8.40	1.63	99.82	0.11
79	8.37	1.71	99.82	0.12
80	8.38	1.72	99.81	0.12
81	8.59	1.88	99.81	0.13
82	8.43	1.75	99.82	0.12
83	8.60	1.93	99.81	0.13
84	8.60	1.87	99.81	0.14
85	8.65	1.93	99.81	0.14
86	8.74	2.05	99.81	0.14
87	8.84	2.13	99.81	0.12
88	8.86	2.22	99.80	0.13
89	9.01	2.25	99.81	0.13
90	9.09	2.31	99.81	0.12
91	9.09	2.41	99.81	0.14
92	9.19	2.31	99.81	0.13
93	9.24	2.49	99.81	0.13
94	9.30	2.59	99.80	0.18
95	9.33	2.61	99.81	0.14
96	9.49	2.83	99.81	0.13
97	9.53	2.87	99.80	0.15
98	9.60	2.81	99.80	0.14
99	9.72	3.03	99.79	0.15
100	9.78	3.10	99.80	0.14

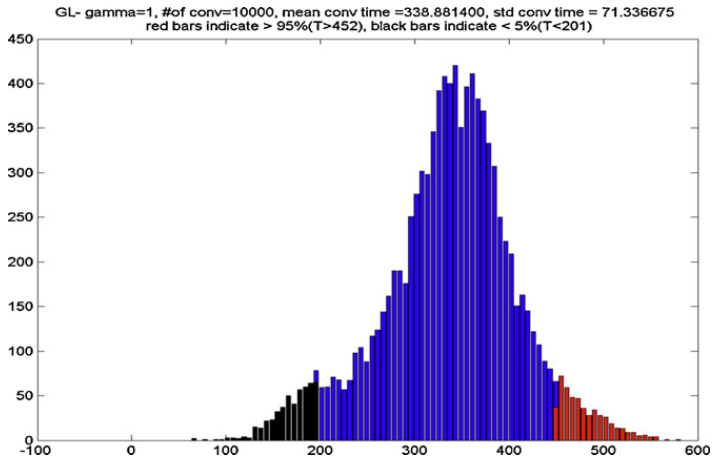

Fig. 4. Distribution of time of first passage through equilibrium, $\gamma = 1$.

The key observation about the dynamics of Groves–Ledyard mechanisms under IEL behavior is that $T(\gamma)$ is smooth and U-shaped in γ.

This can be most easily seen from the figures. There are no discrete jumps, i.e. no values of γ at which the system switches between stability and instability. The smoothness is consistent with the finding of Chen and Gazzale (2004). But, in our case, it is not because we are close to the strategic complementarity condition of $\gamma = 80$. Indeed, IEL shows relatively rapid convergence for values of γ as low as 20, 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.[21]

The standard deviations of $T(\gamma)$ are very small, less than 3, for all γ greater than 15. The variances are also U-shaped in γ, decreasing initially with increases in γ, dropping to the lowest values for $\gamma \in \{40, 60\}$, but then slightly increasing for the

[21] Of course, both the dominant diagonal and strategic complementarity conditions are sufficient but not necessary for convergence.

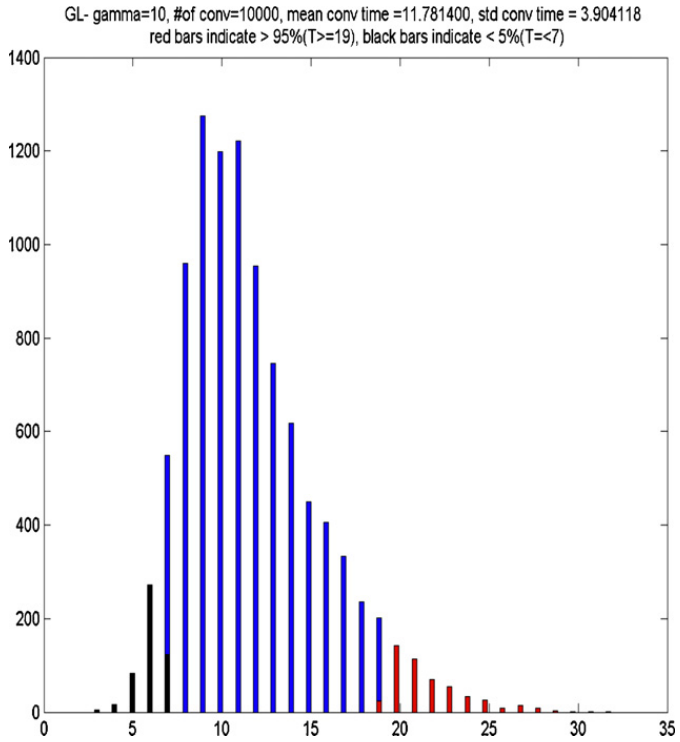

Fig. 5. Distribution of time of first passage through equilibrium, $\gamma = 10$.

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.

The average times to convergence are not much different for the values of γ between 20 and 100. For example, $T(25)$ 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 $T(25)$. Similar relationships can be observed for the values of γ between 23 and 90. As the values of γ increase above 100, the average times to convergence start increasing.

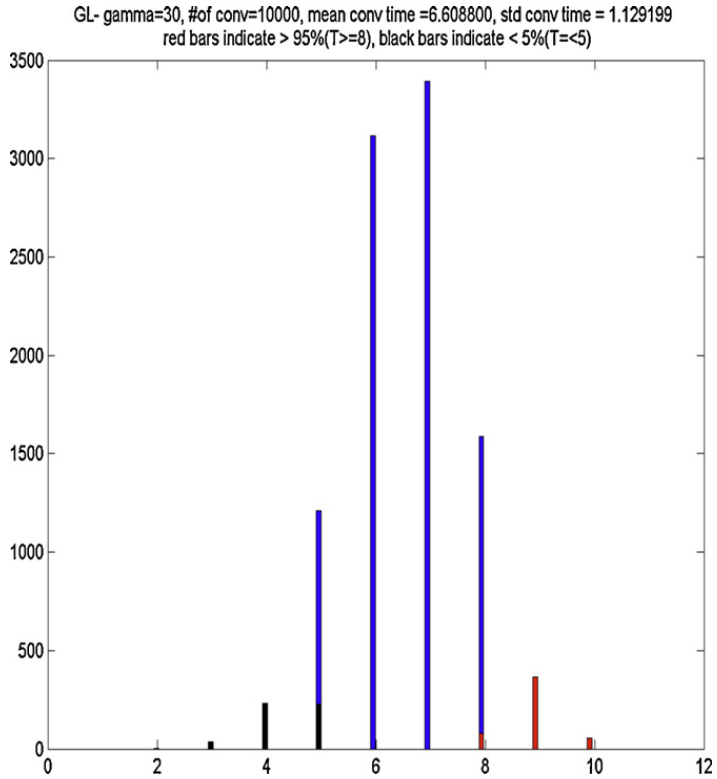

Fig. 6. Distribution of time of first passage through equilibrium, $\gamma = 30$.

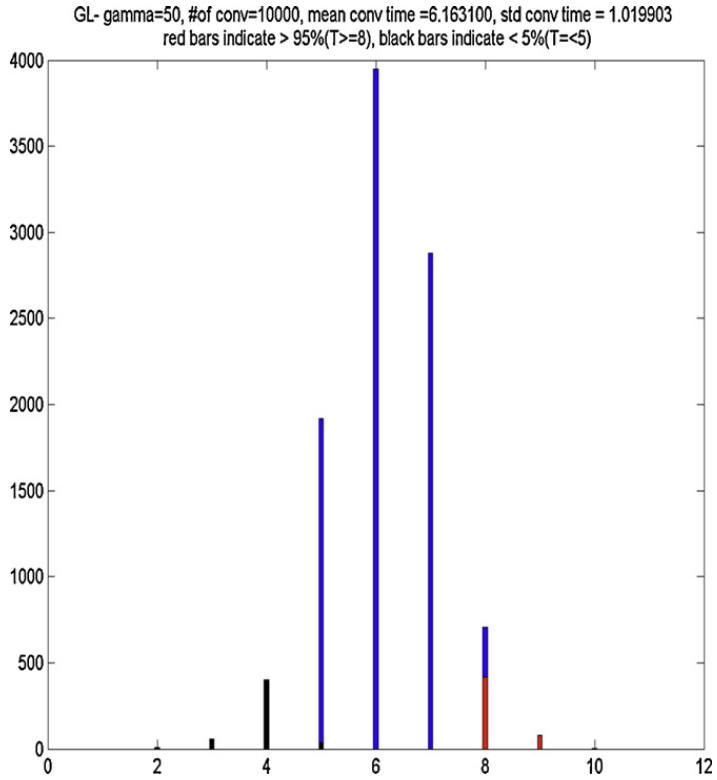

Fig. 7. Distribution of time of first passage through equilibrium, $\gamma = 50$.

It is perhaps understandable that for low values of γ, especially below 30, the time to converge will be higher than for values between 30 and 80 since best reply dynamics are unstable (due primarily to overreactions) and so a period of averaging is necessary to get convergence. But this does not explain the fact that $T(\gamma)$ increases for high values of γ. The explanation lies in the particular role γ plays in the Grove-Ledyard tax function. As γ grows, there is stronger pressure for all agents to

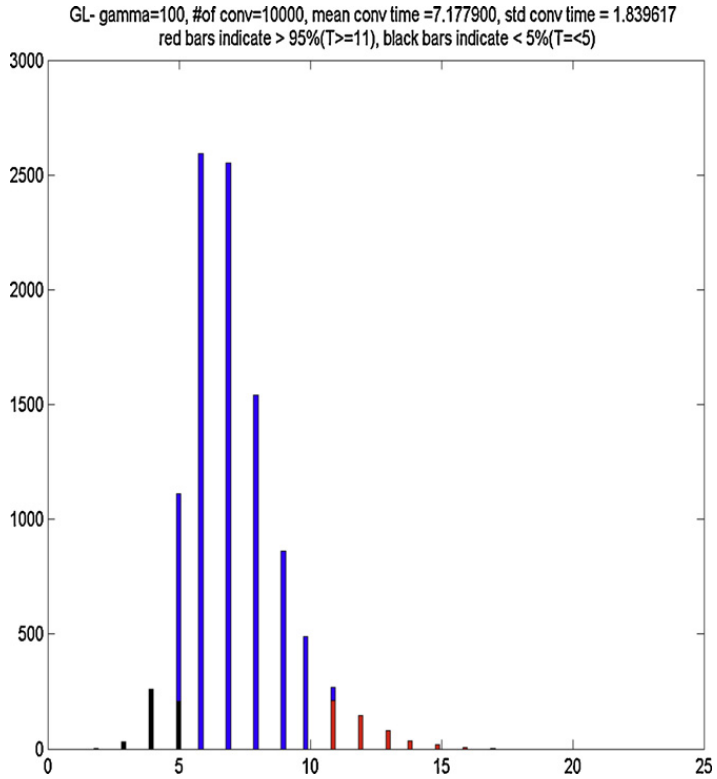

Fig. 8. Distribution of time of first passage through equilibrium, $\gamma = 100$.

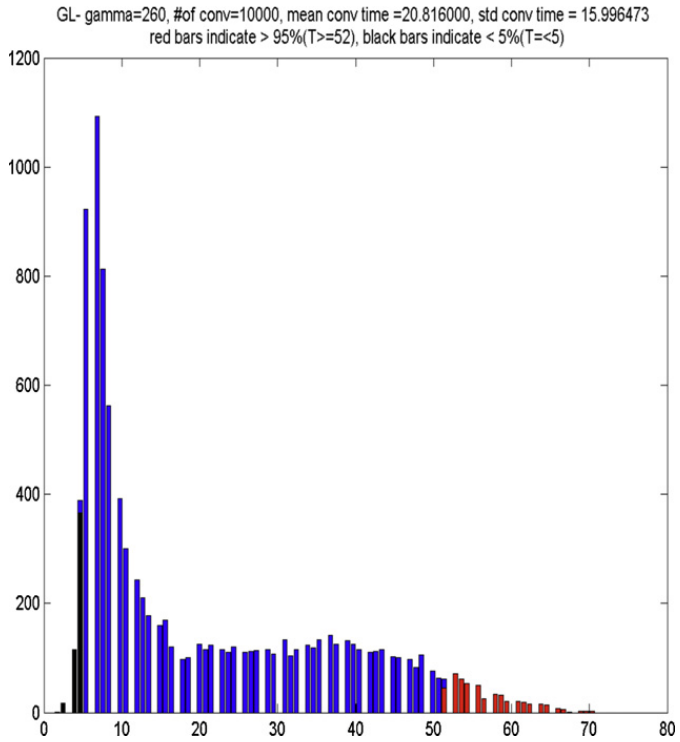

Fig. 9. Distribution of time of first passage through equilibrium, $\gamma = 260$.

coordinate on a common value, typically not the equilibrium one, in order to minimize the difference between their own and average contribution of others. Once the coordination takes place, strategies that participated in this receive relatively high payoffs and are then copied, increasing in frequency. So, collections of strategies become homogenized. At that point, experimentation is required in order to introduce strategies close to the equilibrium ones that will get the mechanism out of the non-equilibrium outcome. However, given the circumstances, it takes time for the required experimentation to succeed. Most new values that deviate from the one everyone has coordinated on will most likely receive lower foregone payoffs and thus, disappear from the collections. The larger γ is, the more important this effect becomes, increasing the time to converge to equilibrium.

Stability,[22] the percent of strategies in the remembered set that are close to equilibrium values during the 100 periods after the first passage through equilibrium, are very high for all values of γ with little variance across runs. For $\gamma \geq 3$, the stability measure is above 99 percent. Unlike EWA, once IEL reaches equilibrium, it pretty much stays there. So it is not luck that gets IEL to converge so quickly.

5.3. Consistency with data

How do the sessions with IEL agents compare to the sessions with human subjects run by Chen and Tang? The relevant comparison data from the IEL simulations are in Table 5.[23] The good news is that the experimental convergence times and efficiencies are remarkably consistent with the predictions of IEL. However, the human subjects are a bit less stable than IEL. The details follow.

[22] Because we have more information in the IEL simulations than we do for the experiments and because we are interested in the stability of the mixed strategies (A_t^i, π_t^i), we use a more rigorous stability index in our evaluations of IEL. It is an index based on the strategies in A_t^i and not on the actions taken, m_t^i. 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. For any one run, where T is the time of first passage to equilibrium, it is given by:

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

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

[23] To make results of simulations comparable to Chen and Tang, the convergence criterion used for IEL in Table 5 is $d = 0.2$.

Table 5
IEL simulations.

γ	$T(\gamma)(\sigma_T)$	$S^a(\sigma_{S^a})$	$S^s(\sigma_S^s)$	E^{100}	E^{10}
1	337.22 (71.47)	93.71 (13.24)	98.73 (2.66)	97.60	76.99
10	11.69 (3.84)	98.80 (1.33)	99.73 (0.24)	99.15	91.59
30	6.59 (1.13)	99.33 (0.85)	99.85 (0.11)	99.69	96.92
50	6.15 (1.02)	99.36 (0.81)	99.86 (0.10)	99.73	97.34
100	7.22 (1.88)	99.32 (0.86)	99.84 (0.12)	99.55	95.61
260	20.64 (15.79)	99.13 (1.09)	99.79 (0.19)	98.21	88.89

$d = 0.2$, $J = 500$ and $\rho_n = 0.033$.

5.3.1. Convergence to Nash equilibrium

A look at the average time to converge, $T(\gamma)$, 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 $d = 0.2$ 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.

5.3.2. Stability

The second statistic of interest is the action based measure of stability. 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 72 while the IEL testbed produced an average of 95. The standard deviation for the Chen and Tang data are also much larger than for the IEL simulations. One reason for this discrepancy may be that Chen and Tang required messages to be in multiples of 0.2, making it very hard to stay exactly within 0.2 of the equilibrium, while we allowed all messages in the IEL simulations.

5.3.3. Efficiency

We also computed average efficiency for the first 100 periods and for the first 10 periods in both the Chen and Tang experiments and the IEL runs. 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 percent efficiency and humans achieve 93.4 percent. 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 percent on average for the first 100 rounds while IEL achieved 99.3 percent for the same time period. For just the first 10 periods, human and computer agents achieved 93 percent.

5.4. Robustness

If we are going to be able to use IEL as a preliminary filter to decide which mechanisms to examine more closely in the lab, we need to be comfortable that the precise values of the parameters are not too important. In this section, we examine how the performance of the mechanisms change as we change various parts of the IEL model. We consider the effect of changing the initialization process and the type of experimentation. We also consider variations in the parameters J, the number of alternatives retained for consideration, and ρ, the rate of experimentation. We find that the qualitative results survive all variations. No matter how we change the parameters of IEL, $T(\gamma)$ remains smooth and U-shaped. The quantitative results do show a limited variation. For example, increasing J, the size of the remembered strategy set, does speed up the average rates of convergence. But even that effect disappears for larger values.

5.4.1. Initialization

We have tried two initialization procedures: a purely random initialization and the one described in Section 4.1, a modified procedure in which a hypothetical strategies are evaluated using hypothetical plays of the others before choosing the first message.

In Table 7 we provide the data from these two approaches for two values of J and three values of γ. As one can easily see, the qualitative property of the U-shape of $T(\gamma)$ is retained for all these values and for both initialization procedures. Modified initialization does decrease the average time to converge. Interestingly, it does so by more than 1 round; that is, more is being computed than simply a one round reaction. Modified initialization also reduces the rate of convergence in a way that is increasing in γ. We believe this is for the same reason as discussed in Section 5.2: the particular role γ plays in the Grove-Ledyard tax function. Since convergence is slower for higher γ, getting the initial choice correct is more important for higher γ.

Table 6
Arifovic and Ledyard experimental results.

γ	$T(\sigma_T)$	$S(\sigma_S)$	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

Note: $d = 0.2$. NA(4/4) – 4 of the 4 sessions did not converge in 100 rounds.

Table 7
Modified vs. random initialization. Average time to convergence.

Initialization	$J = 200$		$J = 500$	
	Random	Modified	Random	Modified
$\gamma = 1$	650.90	643.41	510.20	504.58
$\gamma = 50$	9.76	9.00	8.64	7.78
$\gamma = 100$	14.09	12.10	11.70	9.88

$J = 500$, $\rho = 0.033$, 10,000 runs and $d = 0.1$.

Table 8
Variations in experimentation process and rates.

Experimentation	γ			
	1 $T(\gamma)(\sigma_T)$	50 $T(\gamma)(\sigma_T)$	100 $T(\gamma)(\sigma_T)$	140 $T(\gamma)(\sigma_T)$
$\rho_n = 0.033$	504.58 (95.89)	7.78 (1.21)	9.88 (3.10)	15.22 (8.24)
$\rho_u = 0.033$	879.09 (215.15)	7.95 (1.33)	10.59 (4.56)	19.87 (15.25)
$\rho_n = 0.25$	1416.02 (1421.62)	10.40 (2.17)	12.50 (3.42)	16.83 (6.80)
$\rho_u = 0.25$	dnc	13.22 (2.92)	17.27 (5.96)	25.50 (13.12)

$J = 500$, $\rho = 0.033$, $d = 0.1$, modified initialization, 10,000 runs. dnc - did not converge.

5.4.2. Experimentation

In our simulations, we used an experimentation process based on the normal distribution whose mean is history dependent as the value of the action in A_t^i being experimented with. Here, we compare that to a history independent process: 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 the results of these variations[24] in Table 8. The qualitative result, that $T(\gamma)$ is smooth and U-shaped, survives all variations.

Experimentation using the uniform distribution does result in higher values of T^γ than for the normal distribution with the lowest effect occurring at $\gamma = 50$. That is, the uniform distribution increases the concavity of $T(\gamma)$. For a given distribution, increases in the rate of experimentation increase the time of convergence. As long as $\gamma > 1$ the effects are on the order of only 10–30 percent for a 700 percent increase in ρ and are decreasing in γ. These effects seem to be considerably magnified at $\gamma = 1$.

Comparing the results in Table 8 with the experimental data in Tables 2 and 6, it is obvious that experimentation from the normal distribution is favored. and that the choice of the value of ρ is probably not that crucial. So we have used the normal distribution with $\rho = 0.033$ in all the simulations in Tables 4 and 5.

5.4.3. Remembered set size

We tried a number of different values for J, the size of A_t^i. As before, the qualitative result, that $T(\gamma)$ is smooth and U-shaped, survive these variations.

Increases in J speed up the rate of convergence at a decreasing rate, at least for the range we have considered. The data are presented in Table 9. Increasing J from 50 to 100 reduces the convergence time by 40 percent on average across γ. An increase from 100 to 200, reduces average convergence times by 30 percent. From 200 to 500, we get a 20 percent drop in times. The reduction flattens out and is only 3 percent 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, see the data in Tables 2 and 6, 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.

[24] 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.

Table 9
Variations in the size of the set of alternatives, J.

γ	$J=50$ $T(\gamma)(\sigma_T)$	$J=100$ $T(\gamma)(\sigma_T)$	$J=200$ $T(\gamma)(\sigma_T)$
1	1296.79 (458.36)	863.35 (253.33)	646.37 (152.16)
50	23.82 (12.66)	12.95 (5.78)	8.98 (2.41)
100	30.90 (17.48)	17.72 (9.67)	12.09 (5.30)
150	50.55 (31.13)	31.48 (19.65)	22.32 (13.96)

γ	$J=500$ $T(\gamma)(\sigma_T)$	$J=1000$ $T(\gamma)(\sigma_T)$
1	504.58 (95.89)	449.50 (79.40)
50	7.78 (1.21)	7.55 (1.08)
100	9.88 (3.10)	9.40 (2.61)
150	16.92 (9.65)	15.13 (8.10)

$\rho=0.033$, $d=0.1$, modified initialization, 10,000 runs.

5.4.4. Other environments

Another way to check robustness is to apply this methodology to a significantly different environment. We have used IEL, with J = 200, $\rho=0.033$, and random initialization, for experiments with the voluntary contributions mechanism. See Arifovic and Ledyard (2009). Unlike the GL mechanism, the Nash equilibria of the VCM are dominant strategy equilibria and are directly opposed to the efficient allocation. Using a utility function with heterogeneous other-regarding preferences we generate simulations that provide average rates of contribution in

each iteration that compare favorably with those from experimental data.

We have also successfully applied our behavioral model in the context of the call markets where the IEL model generates the same types of price volatility and efficiencies as those generated in our experiments with human subjects (Arifovic and Ledyard, 2007). We used IEL with $J=100$, $\rho=0.033$ and random initialization. We varied the amount of information available to human subjects and IEL agents. In the *c* losed book design, agents could observe only the previous period's price level while in the *open* book design they could, in addition, observe all the previous period's asks and bids. Both experiments with human subjects and IEL simulations exhibited higher price volatility and lower efficiency in the open book call market.

5.5. Summary to here

We are interested in how Groves–Ledyard mechanisms perform when used repeatedly in a sequence of one-shot games where agents know only their own preferences.

(1) There are theoretical conditions that are sufficient for convergence, dominant diagonal and strategic complementarity, but these conditions provide no information about rates of convergence.
(2) The experimental work merely separates $\gamma=1$ from $\gamma=100$: convergence does not happen in the former but does in the latter. Nothing is known about other values of γ.
(3) Even though EWA is very successful in modeling behavior in one-shot games with very small, finite strategy spaces, it is not successful in modeling behavior in repeated games with a continuum strategy space unless one wants to be involved in fine tuning.
(4) We propose IEL as a behavioral model for repeated play in games with large strategy spaces. When applied to the Groves–Ledyard mechanisms, it predicts that $T(\gamma)$ is smooth and U-shaped in γ with the minimum occurring around $\gamma=50$. That is also the value of γ which maximizes the average efficiency both over the first 10 rounds and the first 100 rounds. IEL also predicts that convergence is stable in the sense that once the model first nears the equilibrium, it remains in its neighborhood.
(5) The predictions of IEL are consistent with the data from the experiments of Chen and Tang (1998).
(6) The predictions of IEL are robust to variations in the parameters of IEL.

IEL is consistent with existing experimental evidence. But the experiments are only for $\gamma=1$ and 100. What about other values of γ? For example, will the U-shape occur in the data? Will $\gamma=50$ be the value that minimizes average time to converge? Will human subjects make choices that are as stable and efficient as IEL at other values of γ? To get these answers, we conducted a sequence of experiments at Caltech.

6. Experimental testing

Our experiments were conducted between May and September 2007. We chose to focus on $\gamma=10$, 30, 50, 100, 260. We chose 50 as the value of γ for which convergence was relatively fast. We chose 30 as the lower bound of the two theoretical

sufficient conditions for stability. We chose 100 as the right hand side of the low convergence times. The values of 10 and 260 were chosen as tests of the U-shaped nature of average convergence times.

6.1. Experimental design

We conducted 4 experiment sessions for each value of γ. Each experimental session lasted for 100 periods. Subjects earned on average about $38 which included a show-up fee of $10. We used the experimental software Ztree (Fischbacher, 2007).[25] Our experimental design is very similar to Chen's and Tang's with two modifications. In Chen and Tang, subjects could make only integer number choices, equivalent to allowing only multiples of 0.2. We allowed subjects to make real number choices with a two decimal points restriction, equivalent to allowing multiples of 0.01. Second, we added a 'what-if' calculator[26] 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.

6.2. Results

A summary of the results from our Caltech experiments is presented in Table 6. The comparable results for IEL are found in Table 5.

6.2.1. Convergence to Nash equilibrium

Qualitatively the data from our experiments is similar to the predictions of IEL. In the previous section, we found that IEL predicts convergence to Nash equilibrium messages of the stage game for all of the values of γ that we simulated, including $\gamma = 1$. Further, 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 are certainly consistent with these predictions. 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 in all 4 sessions. Another prediction from IEL is that the time to first convergence is smooth and U-shaped in γ with the minimum average convergence time occurring at around $\gamma = 50$. Further, average convergence times are predicted to be relatively constant between $\gamma = 10$ and $\gamma = 100$. The data in Table 6 have exactly this pattern.

Quantitatively, the human subjects are a bit different from the IEL agents. Human subjects were faster.

6.2.2. Comparison to Chen–Tang

It is useful to compare our results to Chen–Tang for the 2 values of γ which are common. One can see that for $\gamma = 1$ neither sets of experiments see convergence in 100 periods. But the efficiency levels are higher in Chen–Tang than ours. The key difference occurs in the first 10 periods. In our experiments, the subjects got things fairly wrong early on, leading to negative payoffs for some of the subjects. The subjects in Chen–Tang avoided this. Efficiencies in rounds after 10 were similar in both experiments.

For $\gamma = 100$, the average time to converge is 9 in the Chen–Tang experiments and 2.75 in the Arifovic–Ledyard experiments. There is at least one possible reasons for the difference. Our subjects had a "what-if" calculator they could use to compute hypotheticals which the Chen–Tang subjects did not. It is true that subjects in both experiments had a graph showing payoffs which could be used to estimate hypotheticals, but that seems to be hard to use. We believe that the "what-if" did speed things up for our subjects (and also created the instability in the first rounds of the $\gamma = 1$ experiments. One might also conjecture the possibility of subject pool differences. But the Chen and Tang experiments also used Caltech subjects for 4 of the 7 sessions.[27]

6.2.3. Stability

From Tables 5 and 6, it can be seen that the human subjects are less stable than IEL. For IEL, with the exception of $\gamma = 1$, in the first 100 periods after convergence more than 98 percent of the IEL messages 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 occasionally 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, for example, be introduced as in the Quantal response model (McKelvey and Palfrey, 1995). However, to achieve the variation seen in the data, we would have to find a way to have more variety in the sets A_t^i. Remember replication eliminates many of the experiments.

[25] The experiment instruction is provided in Appendix II.

[26] 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.

[27] The others were run in Amsterdam.

6.2.4. Efficiency

Efficiencies are a slightly different measure of performance than time of first passage or stability. We have computed two measures of efficiency: one, E^{10}, is based on the first 10 periods of the mechanism and the other, E^{100}, is based on the first 100. What is remarkable is that efficiencies are fairly high for the human experiments for all values of $\gamma \neq 1$. For $\gamma = 50$ and 100, the efficiencies from the human experiments are essentially the same as the efficiencies for the IEL simulations. (This is also true for $\gamma = 100$ in the Chen–Tang experiments.) So even though humans are faster 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.

What is perhaps curious is that although the time of first passage for the humans is faster than that for IEL, the efficiencies in the first 10 periods are higher for IEL than for humans. The answer lies in the stability measure. Although IEL is a bit slower, once it gets there it stays. The humans are faster but they wander away more often. The net effect is that slower but more stable is better for efficiency.

7. Summary and final thoughts

7.1. Summary

We are interested in how Groves–Ledyard mechanisms perform when used repeatedly in a sequence of one-shot games where agents know only their own preferences.

(1) There are theoretical conditions that are sufficient for convergence, dominant diagonal and strategic complementarity, but these conditions provide no information about rates of convergence.
(2) The experimental work merely separates $\gamma = 1$ from $\gamma = 100$: convergence does not happen in the former but does in the latter. Nothing is known about other values of γ.
(3) Even though EWA is very successful in modeling behavior in one-shot games with very small, finite strategy spaces, it is not successful in modeling behavior in repeated games with a continuum strategy space unless one wants to be involved in fine tuning.
(4) We provide a behavioral model, IEL. The time to first convergence is predicted to be smooth and U-shaped in γ. Also, convergence is predicted to be stable in the sense that once the messages first near the equilibrium, they remain in its neighborhood.
(5) The predictions of IEL are consistent with the data from the experiments of Chen and Tang (1998) and robust to variations in the parameters.
(6) To test IEL, we ran our own experiments at the California Institute of Technology. Qualitatively, the data from those experiments are consistent with the IEL predictions about convergence and the U-shaped curve. Quantitatively, the human subjects are a little faster, a little less stable, and slightly less efficient than IEL. But for $\gamma = 50$ and 100, the differences between humans and IEL are very small.

In this paper, we have shown the capability of IEL to act as a filter in identifying those mechanisms to take to the lab. Referring to Section 3.4, we can see that IEL predicted answers the questions of interest which were later verified in the lab. (a) There is not a value of γ such that for smaller values, the mechanism does not converge to the Nash equilibrium and for larger values it does. The time it takes to converge to equilibrium does vary systematically with an increase in γ: it is U-shaped. The value of $\gamma = 50$ results in the fastest convergence and highest efficiency.

7.2. Final thought

We have shown that IEL is reasonably good at matching the behavior of experimental subjects in repeated games when that behavior converges to a Nash equilibrium of the stage game.[28] These games do not require very sophisticated strategies. However, there is a class of games that does—coordination games. In these games, the highest joint payoffs come if subjects can either coordinate on a highly Pareto ranked stage game equilibrium from a number of multiple equilibria (as in minimum effort games) or alternate over time among, say, two possible Nash equilibria of the stage game (as in the battle of the sexes). IEL is not well-suited to coordination games.

The main reason IEL does not do well in these games is that the set of remembered strategies is only populated with one-period actions. Coordination seems to require contingent strategies involving moves over a number of stages; strategies that extend beyond one period and that are intended to influence one's opponents. Classic examples of these are tit-for-tat and grim trigger. A simple and straight forward way to rectify this is to let the strategy space be a set of automata, an efficient codification of a contingent multi-period strategy, so that the items in the set of remembered strategies are automata. Then one could apply the experimentation, replication and selection procedures of IEL to this set.

[28] The Grove-Ledyard environments in this paper, the call market environments in Arifovic and Ledyard (2007), and the VCM environments in Arifovic and Ledyard (2009) are examples of such games.

But there remains a problem - how to evaluate the performance of a given strategy; that is, what is the appropriate foregone utility function. It is certainly no longer just the payoff in the one-stage game. If the IEL agent were to know the strategies of the others, then that agent could play each of her strategies against those over some number of periods and use the payoffs from these (hypothetical) tournaments as the foregone utility measure. This is effectively the approach taken by Hanaki et al. (2005) in a strategy learning model with which the authors succeed in replicating the behavior of subjects. But, in experiments, subjects do not know or see the strategies of others. They only see, at most, the actions taken in each round. This would not be a problem if all subjects had to pick a strategy and then wait in a coordinated manner for some number of rounds before switching strategies. But that also is not a feature of most experiments. Subjects are free to change strategies whenever they wish.[29]

Our approach has been to constrain our IEL agents to exactly the same information as is available to subjects in an experiment. Doing so rules out the use of tournaments to evaluate repeated game strategies. We have not yet found a satisfactory solution to this evaluation problem. That remains for future research.

Acknowledgements

We thank Olena Kostyshyna and Stephen Tse for their very able research assistance. We thank Peter Bossaerts, Colin Camerer, Yan Chen, Paul healy, Tom Palfrey, Charlie Plott, and several anonymous referees for helpful discussions. 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 NSF.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2011.01.021.

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., Ledyard, J., 2004. Scaling up learning models in public good games. Journal of Public Economic Theory 6, 205–238.
Arifovic, J., Ledyard, J., 2007. Call market book information and efficiency. Journal of Economic Dynamics and Control 31, 1971–2000.
Arifovic, J., Ledyard, J., 2009. Individual Evolutionary Learning, Other-regarding Preferences, and the Voluntary Contributions Mechanism. Working Paper, California Institute of Technology.
Camerer, C., Ho, T., 1999. Experience weighted attraction learning in normal form games. Econometrica 67, 827–873.
Camerer, C., Ho, T., Chong, J., 2004. A cognitive hierarchy model of games. The Quarterly Journal of Economics 119, 861–898.
Chen, Y., 2003. Incentive-compatible mechanisms for pure public goods: a survey of experimental research. In: Plott, C., Smith, V. (Eds.), The Handbook of Experimental Economics Results. North-Holland.
Chen, Y., Gazzale, R., 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.
Chen, Y., Plott, C., 1996. The Groves–Ledyard mechanism: an experimental study of institutional design. Journal of Public Economics 59, 335–364.
Chen, Y., Tang, F., 1998. Learning and incentive compatible mechanisms for public goods provision: an experimental study. Journal of Political Economy 106, 633–662.
Dal Bó, P., Frechette, G.R., 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. American Economic Review 101, 411–429.
Dawid, H., 1999. Adaptive Learning by Genetic Algorithms, Analytical Results and Applications to Economic Models, 2nd ed. Springer-Verlag, Berlin.
Fischbacher, U., 2007. z-Tree—Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.
Gabay, D., Moulin, H., 1980. On the uniqueness and stability of Nash-equilibria in noncooperative games. In: Kleindorfer, P.S., Tapiero, C.S. (Eds.), Applied Stochastic Control in Econometrics and Management Science. North Holland.
Groves, T., Ledyard, J., 1977. Optimal allocation of public goods: a solution to the 'free rider' problem. Econometrica 45, 783–809.
Hanaki, N., Sethi, R., Erev, I., Peterhansl, A., 2005. Learning strategies. Journal of Economic Behavior & Organization 56 (4), 523–542.
Healy, P., 2006. Learning dynamics for mechanism design: an experimental comparison of public goods mechanisms. Journal of Economic Theory 129, 114–149.
Healy P., Mathevet, L., 2009. Designing stable mechanisms for economic environments. Manuscript. Available at http://healy.econ.ohio-state/research.html.
Hurwicz, L., 1960. Optimality and efficiency in resource allocation processes. In: Arrow, K.J., Karlin, S., Suppes, P. (Eds.), Mathematical Methods in the Social Sciences. Stanford University Press, Stanford, CA.
LeBaron, B., 2000. Agent based computational finance: suggested readings and early research. Journal of Economic Dynamics and Control 24, 679–702.
McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games and Economic Behavior 10, 6–38.
Milgrom, P., Roberts, D.J., 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica 58, 1255–1277.
Muench, T., Walker, M., 1979. Identifying the free rider problem. In: Laffont, J.-J. (Ed.), Aggregation and Revelation of Preferences. North-Holland Publishing Co..
Muench, T., Walker, M., 1983. Are Groves–Ledyard equilibria attainable? The Review of Economic Studies 50, 393–396.
Roth, A., Erev, I., 1995. Learning in extensive games: experimental data and simple dynamic model in the intermediate term. Games and Economic Behavior 8, 164–212.

[29] This observation is not unique to us. See, for example, Dal Bó and Frechette (in press), where they have designed a very interesting experiment to circumvent these observational constraints.