Journal of Quantitative Economics, Vol. 10 No.2, July 2012

EQUILIBRIUM PLAY AND LEARNING IN A STOCHASTIC PUBLIC GOODS EXPERIMENT[1]

PRIYODORSHI BANERJEE[2]
SUJOY CHAKRAVARTY[3]
RUCHIKA MOHANTY[4]

Abstract

In the experimental game we study, the benefit received from a public good varies non-linearly and stochastically with effort. We examine the effect of a quality bound under which group members receive very negligible benefits from the public good. Our theory predicts that the presence of such a "pass mark" on quality of the good ultimately provided may under certain conditions reduce the incidence of free-riding. We obtain very mixed support for this hypothesis using both mixed strategy as well as quantal response equilibrium benchmarks. In contrast, a reinforcement-learning model does better at predicting laboratory behavior.

Keywords: free rider problem, public goods, quantal response, reinforcement learning

JEL Classification: H41, C91, C72, D83

1. Introduction

Public goods may either be provided by the government using tax revenues or by the community itself through a voluntary contribution of material resources or effort. Economic situations with public goods lead to a problem that is unique to the nature of these goods, i.e. - the free-rider problem. Since the benefits from these goods are typically conferred to all but the

---

[1] This research was made possible by grants from the Planning and Policy Research Unit (PPRU) of the Indian Statistical Institute (ISI), Delhi and the excellent laboratory facilities at the Humanities and Social Science (HUSS) Department at the Indian Institute of Technology (IIT), Delhi and the Centre for Experiments in Social &Behavioral Sciences (CESBS),JadavpurUniversity. We are grateful to Avinash Bhardwaj for programming the computer interface, and Remi Ray and Arjun Sengupta who provided able research assistance. We thank Professor Gautam Gupta for hosting us at the CESBS and being a source of support during the experiments and Dr Syamala Kallury of the HUSS Department at IIT for being generous with the laboratory facilities. We are especially indebted to an anonymous referee for many valuable comments. The usual disclaimer applies.

[2] Economic Research Unit, Indian Statistical Institute (ISI), Kolkata, E-mail: banpriyo@gmail.com

[3] Centre for Economic Studies and Planning, School of Social Sciences, Jawaharlal Nehru University (JNU), New Delhi and IMT Ghaziabad, E-mail: sujoy@mail.jnu.ac.in

[4] E-mail: ruchika.mohanty@gmail.com

cost payments are non-enforceable (due to voluntary contribution or provision), there is over-utilization or under-contribution of these public goods.

Public goods problems are typically analysed using deterministic N-player Prisoner's Dilemma (NPD) games with one Pareto dominated Nash equilibrium, using a linear production function for the public good. Voluntary Contribution Mechanism and Common Property Resource games are amongst the games in this class which have received maximum attention (see Ostrom 1990 and Ostrom, Gardner and Walker 1994a, 1994b for detailed reviews).

In contrast, the public goods problem we examine (see Banerjee 2007and Section 2 below), has benefit varying *non-linearly* and *stochastically* with effort. It can be viewed as a stochastic generalization of the 'best-shot', 'dragon-slaying' and similar games (see Hirshleifer 1983, 1985 and Bliss and Nalebuff 1984 respectively; see also Lipnowski and Maital 1983, Cornes 1993 and Chowdhury, Lee and Sheremata 2011).

We specifically examine the effect of a quality bound (for the public good provided) below which community members receive insignificant benefits. The theory predicts that the presence of such a "pass mark" on the quality of the good provided may sometimes reduce the incidence of free-riding. This result has significant implications for the design of mechanisms and institutions that provide public goods and services.

Given the profusion of public goods experiments in the literature, there is to our knowledge very little work on public goods whose production is stochastic and/or non-linear. The few exceptions include Harrison and Hirshleifer (1989) and Berger and Hershey (1994). Harrison and Hirshleifer (1989) study a deterministic version of the game we employ, the best-shot game. The significant differences in the design used and questions asked render our results largely incomparable with those from this study. Berger and Hershey (1994) find less efficient behavior among players as compared to experiments reported in the linear public goods literature, in an insurance market setting with moral hazard.[5] The main difference between their game and the standard public goods game is the stochastic return on investment.

We contribute to this small literature by studying a hitherto untested public goods game whose production is stochastic and non-linear. Our objectives are to test how well Nash predictions describe outcomes, particularly in relation to alternative approaches such as Quantal Response and Learning models.

Our public goods game deals with an environment where output is stochastically dependent on the maximum of individual efforts. Consider as an example the family through the lens of evolutionary-genetic theory, which postulates that animal behavior is ultimately determined by the maximization of reproductive survival.[6] Thinking of the survival of the offspring as a public good for both parents, suppose a situation has arisen where a decision needs to be made regarding future course of family action, with any choice stochastically affecting survival

---

5 When an insured individual provides costly effort to prevent (or reduce) loss, reductions in insurance premium can benefit all policyholders. In this way, investing in loss prevention may be considered to be a public good. Like in the standard public goods games, though it is not individually rational to invest in loss prevention once insured (this is the moral hazard problem in the Principal-Agent literature), agents may still invest in order to realize the public good benefits conferred on the group and thus mitigate the moral hazard problem.

6 The example is based on one in Hirshleifer (1983).

probability. Our construction fits this scenario in which the parents propose options independently with quality dependent on effort, where the better of the two is selected. In another example, suppose a government wishes to promulgate policy in some area. Better implemented policy will benefit all, though stochastically. Stakeholders can each propose a policy, whose quality depends on effort. Our model fits this setting if the best of such proposals is implemented.

On average we obtain much lower levels of cooperation than those predicted by the Nash equilibrium. We also obtain very weak support for the theoretical prediction that under certain conditions imposing the quality pass mark increases the incidence of cooperation. In fact the subjects overall show very little evidence of attempting to play Nash strategies. We find Quantal Response also has poor predictive power, and subject behavior is described better by reinforcement learning models of Roth and Erev (1995).

2. The Theoretical Model and Discussion

We describe the theory underlying our experiment in a simple intuitive way. The full extensive form of the game, together with a discussion of the assumptions can be found in Banerjee (2007). The reader is referred to that paper for detailed development, analysis, statements of results, and proofs.

In any period, two players are randomly matched to form a pair. Each player has two possible actions: to *invest* or to *not invest*. Players can also play a mixed strategy, choosing investment with some probability. These decisions are not publicly observable and made simultaneously.

A player's investment contributes towards building her *investment quality (*or simply*, quality)*. If player i does not invest, she bears no private cost, and her quality is 0. If she invests, her private cost is c> 0, and her ability is the realization of a random variable distributed over [0, 1] according to the continuously differentiable, strictly increasing distribution function F, whose density is denoted as f.

Given the profile of realized qualities, α_1 and α_2 for the two players, and a pre-determined *cut-off level* $\mu \in [0, 1)$, the gross payoff for any player i is $\max(\alpha_1, \alpha_2)$ if $\max(\alpha_1, \alpha_2) \geq \mu$, and 0 otherwise. This payoff function is easy to understand if we think of the two players as members of a committee producing a public good. Each player gets the same gross payoff. The payoff is a function of the highest ability amongst them ($\max(\alpha_1, \alpha_2)$), provided this ability exceeds a minimum cut-off μ.

We now discuss some results from the analysis. First of all, it is easy to show that a unique symmetric Nash equilibrium exists for all c and μ. If private cost is sufficiently low, both players choose investment with probability 1, while if private cost is very high, neither invests. For intermediate cost, players choose investment with probability less than 1, with the equilibrium probability of choosing high effort decreasing in the cost.

Specifically, let X_k denote the highest order statistic of a sample of size k drawn from [0, 1] according to distribution function F, k = 1, 2. Let F_k be the distribution function of X_k. Define

$$b_1(\mu) = \int_\mu^1 x dF_1(x), b_2(\mu) = \int_\mu^1 x dF_2(x) - \int_\mu^1 x dF_1(x) \quad \text{... (1)}$$

$b_1(\mu)$ is the expected ability of a player who has invested, conditional on quality being at least μ, while $b_2(\mu)$ is the increment in the expected value of the highest order statistic, conditional on the random variable taking a value no less than μ, when the sample size is increased from 1 to 2. It can be proved that that a mixed strategy equilibrium exists if and only if $c \in [b_2(\mu), b_1(\mu)]$. In this equilibrium, σ- the probability a player invests, is the solution to the equation:

$$(1-\sigma)b_1(\mu) + \sigma b_2(\mu) = c \quad \text{... (2)}$$

Secondly, as is standard in many public good contribution games, we find that players have a tendency to free-ride on other players' investments, and so all players could be better off by committing to some investment level ex ante.[7] In particular, free-riding always exists if the equilibrium is in mixed strategies. To understand this result, observe that gross payoff is a function of the maximum of the two qualities because each player's quality is a public good. Thus, if a player believes that the other is very likely to invest, and therefore generate high quality with a large probability, she may have a tendency to invest with a lower probability and thereby save on private investment cost.

Thirdly, the degree of free-riding is dependent on the pre-committed cut-off level μ, and it may be optimal from the perspective of ex ante net expected payoff to commit to a positive μ. Notice that any cut-off $\mu > 0$ is inefficient ex post. This is because given quality levels, if $0 < \max(\alpha_1, \alpha_2) < \mu$, each player gets 0 gross payoff, while if the cut-off is not applied, players get positive gross payoff. In spite of this, the players may be better off committing to a positive cut-off ex ante. This is because setting such a cut-off can improve ex ante investment incentives for any player by jointly penalizing both players when realized quality levels are very low. Hence, a positive cut-off can increase the expected maximum level of quality. In particular, this improvement in ex ante net expected payoff is always possible if *c* is greater than, yet close to $b_2(0)$. In such situations, the payoff loss from the cut-off, stemming from the ex post inefficiency, is outweighed by the payoff gain from reduction in free-riding ex ante.

The game described above is a variant on the standard public good contribution problem which has been extensively studied theoretically and experimentally. The key features distinguishing the current version from the standard form lie in stochastic production and selective aggregation. Normally, in public good games, the amount of the public good produced is some function of a linear combination of all contributions taken together. In the context of informational public goods, such a model is not always appropriate. Here, contributions are probabilistically linked to ability and only the highest ability plays a role in the public good production process. In an environment where the fruit of one's labor is uncertain, in a community with a well-functioning selection procedure, typically it is the highest achievers who assume leadership roles. It is their information and decisions resulting thereof which guide social outcomes, affecting everyone else's payoffs. The free-riding problem thus manifests itself in this setting in novel ways.

The results of the analysis above are taken to the laboratory environment for experimental verification. We compare results from a baseline version without any cut-off ($\mu = 0$) to a treatment with a cut-off ($\mu > 0$). Following the discussion above, our focus will be on

7 The terms *ex ante* and *ex post* are used with respect to the timing of effort choice.

parameter values which generate equilibrium in mixed strategy, requiring some care in the experimental design. The laboratory version of the above game is described below in the first part of the section on experimental design.

3. Experimental Design

3.1 The public goods game used in the laboratory

We use a version of a game presented in the last section, in the laboratory. Two players (who constitute a group) have to simultaneously specify probabilistically whether to invest or not invest in the production of a good. They can of course make their choice non-stochastic by specifying either 0 (no investment) or 100 (invest with certainty). However they also have the option of specifying a likelihood, i.e. - a number in the set {1, 2, .., 99}, and depending on the number chosen, the computer would implement investment with that probability. Thus our design allows a player to directly select a non-degenerate mixed strategy in addition to pure strategies. Once the two likelihoods are obtained, the computer determines whether an investment outcome is obtained for both players. If an investment outcome is obtained, the player pays the investment cost (see table 2 for the different investment costs for different treatments). Finally the computer makes a draw from [50, 250] using the uniform distribution to select an investment quality for each player. If any player's outcome was not to invest (either due to deterministic choice or due to the luck of the draw) his investment quality is at the lowest possible level of 50. The maximum of the two qualities drawn is used to calculate the payoff of both players. Table 1 illustrates the four possible payoffs of the game. Here, A_1 is the quality realized for player 1, A_2 is the quality realized for player 2 and C is investment cost.

Table 1. The Laboratory Game

Player 2 *Player 1*	*INVESTMENT*	*NO INVESTMENT*
INVESTMENT	$Max(A_1, A_2) - C$ $Max(A_1, A_2) - C$	$A_1 - C$ A_1
NO INVESTMENT	A_2 $A_2 - C$	50 50

The payoff matrix is almost identical for the treatments with a quality-cut off, with one crucial difference: if the computer chose invest for at least one player, and the maximum quality drawn is less than the cut-off, the return is 50 for either player. Of course, in that case, any player for whom the computer chose invest obtains a net payoff of 50 – C.

Our design requires direct elicitation of mixed strategies. The first attempts at studying mixed strategies in the laboratory are from Ochs (1994) and Bloomfield (1994).We follow the procedure proposed in Shachat (2002). Subjects are asked to give a probability distribution over a set of actions instead of picking an action. A single random draw is then made from this distribution, and the realization becomes the subject's action (in our case, investment).[8]

3.2 Laboratory Protocol

The experiments use a paired strangers matching protocol on *z-tree* (Fischbacher, 2007) where a static game form is repeated over several periods (15-20 depending on the number of subjects and time constraints). In each period each subject is matched with another (an even number of subjects were chosen in any experimental session). There is a very small probability that a person may be matched with the same person more than once. All subjects have to make only one decision per period, i.e. – specify the probability of investment. Depending on that probability the program draws a decision to invest or not in the public good. The treatments are organized in within-subject B*X* and *X*B formats (a certain number of repetitions of B followed by a certain number of repetitions of *X* or vice-versa for the same cohort) where B is the baseline treatment with no quality cutoff. *X* represents a certain treatment in the form of a quality cutoff that either induces higher contribution (Treatment Increase = *I*) or lower contribution (Treatment Decrease = *D*). The next section details these treatments.

The participants are undergraduate students in Engineering and Sciences from the Indian Institute of Technology, Delhi and Undergraduate and Master's students in Liberal Arts, Sciences and Engineering from Jadavpur University, Kolkata. Sessions were run in both Delhi and Kolkata from April to June 2008. The recruitment of subjects in Kolkata was done primarily over e-mail, while in Delhi, the experimenters went to undergraduate classes to sign up students. The subjects were paid a show-up fee of Rs 50 and averagely earned between Rs. 200 and Rs. 500 from a session, which took between 90 and 120 minutes.[9,10]

The instructions were read aloud to the subjects and special care was taken in making sure that they understood the mixed strategy elicitation process.[11] After the instructions were read and explained, the subjects took a quiz that tested their understanding of the decision and payoff

8 This procedure is more faithful to the way the implementation of mixed strategy is traditionally viewed in game theory. Here players choose a probability distribution over the set of actions, and then a realization is drawn according to the specified likelihood. On the other hand, Ochs (1994) and Bloomfield (1994) have subjects making several choices every period and construct the mixed strategies from the relative proportions of these choices. However the successive "rounds" of choices within a period are not independent making the construction and interpretation of these proportions as mixed strategies debatable. See Wooders and Shachat (2001) and Shachat (2002) for a detailed discussion.

9 At purchasing power parity (PPP) exchange rates, the range of earnings was approximately between US$ 13 (Rs. 200) and US$ 33 (Rs. 500). The PPP data was for 2009 and can be found in the Penn World Tables (Heston et al., 2011)

10 The reason for using more than one location was to avoid contamination of the subject pool. The first sessions were run in IIT, which is a residential institute with low social distance among students. Running more than six sessions was not felt to be optimal as subjects typically discuss their strategies with prospective participants. Since our design required 12 sessions, it was felt that a location change (given that the students' demographic profiles in Delhi and Kolkata were not that different) would minimize bias in the results.

11 The instructions not reproduced here but a longer working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2047675 has instructions for the Baseline and the Increase treatments.

calculation processes. Only after the experimenters were convinced that the subjects had understood all aspects of the experimental protocol were the sessions allowed to begin.[12]

4. Treatments and Hypotheses

The main three treatment groups are *Baseline* (*B*), *Increase* (*I*) and *Decrease* (*D*) where the equilibrium mixing probabilities for contribution are higher in *I* and lower in *D* (as compared to *B*). These parameters are displayed in the table below. For the analysis later, the treatments in each group have been pooled. Henceforth *B* refers to all the observations on the baseline treatment and *I* refers to all observations on the Increase treatment.

Table 2. Treatments

Treatment Name	*Investment Cost*	*Quality Cutoff*	*Nash equilibrium probability*
Baseline (*B*0)	40	None	0.90
Baseline (*B*1)	42	None	0.87
Baseline (*B*2)	45	None	0.825
Increase (*I*1)	42	150	0.99
Increase (*I*0)	40	100	0.955
Decrease (*D*)	45	185	0.57

Table 2 gives information about the different treatments that are used in this study and their associated equilibrium point predictions. Notice that we present Nash equilibrium benchmarks for *risk neutral* players. For a moderate to high level of risk aversion, the probability of investment falls to zero for all treatments with a quality cut-off regardless of whether they increase (*I*) or decrease (*D*) the probability of investment in the risk neutral model.[13] Accordingly, for a moderate amount of risk aversion, the investment probability for the*I*treatment may sometimes be lower than that of the *B* treatment. In fact, to obtain point predictions in equilibrium, we need to calibrate the coefficient of risk aversion for the specific population playing our public goods game. This exercise was not performed and thus we do not attempt to report the large set of probability estimates from the numerous potential values that the coefficient of risk aversion could take.

We have experimental cohorts playing the treatments in different orders. This is to safeguard against a potential confound that can arise if playing *D* or *I* before and after the no-cutoff baseline leads to a difference in observed investment behavior.

---

[12] Demographic data was collected for the subjects in the Kolkata sessions. These were not used in the analysis presented in this section, as most of the covariates such as income; age and education of parents were roughly similar among participants. About 45 percent of the sample was female but no gender effect was observed with respect to investment behavior. Though demographic data was not collected in Delhi, the demographics of the subject pool did not appear to be significantly different from that in Kolkata.

[13] with a constant relative risk aversion (CRRA) utility function $u(x) = x^{1-r}/(1-r)$,we classify preferences as risk averse, risk neutral and risk loving as $r > 0$, $r = 0$ and $r < 0$ respectively. Here r is the coefficient of relative risk aversion and $r \neq 1$. The ex-post payoffs are calculated in a quasi-linear manner, i.e., $\pi(c, x) = x^{1-r}/(1-r) - c$, if the investment outcome occurs and $\pi(0, x) = x^{1-r}/(1-r)$ if the investment outcome does not occur.

Based on the above treatments our main hypotheses are:

H0: Investment probabilities are not significantly different from Nash equilibrium predictions.

H1: Investment probabilities in Treatment *D* are lower than those in treatments *B*0, *B*1 and *B*2

H2: Investment probabilities in Treatment *I*0/*I*1 are higher than those in treatments *B*0, *B*1 and *B*2

H3: Investment probabilities in Treatment *I*0/*I*1 are higher than those in treatment *D*.

Table 3. Session Specific Details

Session name	*Number of subjects*	*Location*	*Treatment*	*Observed*	*Nash Prediction (risk neutral)*	*Quantal Response Equilibrium (QRE)*
4151	16	Delhi	Periods 1 – 20 (B2)	0.56	0.83	.533
			Periods 21 – 40 (D)	0.36	0.57	.536
4152	16	Delhi	Periods 1 – 20 (B2)	0.72	0.83	NC
			Periods 21 – 40 (D)	0.43	0.57	.526
4221	18	Delhi	Periods 1 – 20 (D)	0.53	0.57	.556
			Periods 21 – 39 (B2)	0.74	0.83	.556
4222	18	Delhi	Periods 1 – 20 (D)	0.64	0.57	.533
			Periods 21 – 40 (B2)	0.58	0.825	.56
4231	14	Delhi	Periods 1 – 18 (B1)	0.68	0.87	.534
			Periods 19 – 36 (I)	0.51	0.99	.529
4232	18	Delhi	Periods 1 – 18 (I)	0.61	0.99	.53
			Periods 19 – 36 (B1)	0.59	0.87	.531
6121	18	Kolkata	Periods 1 – 20 (B1)	0.69	0.87	.525
			Periods 21 – 40 (I)	0.60	0.99	.527
6131	18	Kolkata	Periods 1 – 20 (I)	0.59	0.99	.529
			Periods 21 – 40 (B1)	0.62	0.87	.53
6191	14	Kolkata	Periods 1 – 18 (B0)	0.48	0.90	.529
			Periods 19 – 36 (I1)	0.55	0.955	.535
6201	18	Kolkata	Periods 1 – 24 (I1)	0.68	0.955	.533
			Periods 25 – 44 (B0)	0.66	0.90	.525
6261	14	Kolkata	Periods 1 – 18 (B0)	0.71	0.90	.534
			Periods 19 – 36 (I1)	0.72	0.955	.534
6262	14	Kolkata	Periods 1 – 18 (I1)	0.61	0.955	.533
			Periods 19 – 36 (B0)	0.60	0.90	.532

5. Experimental Results

From table 3 above it is clear that point predictions from theory are not borne out by the experimental observations. Specifically, the observed likelihoods for contribution understate the theoretical predictions significantly. Thus we find no support for hypothesis H0. Furthermore, we find that the mixed strategy responses are significantly lower when the *D* treatment is played, validating hypothesis H1. This is robust across the 6 sessions that employ this treatment. However when the *I* (or *I*1) treatment is played as either the first or the second treatment within a cohort we find that the treatment effect is sometimes in a direction opposite to that predicted by theory. In 5 out of 8 sessions that employ the *I* treatment the observed probabilities are higher than in *B*, but these differences are largely statistically insignificant. In the other 3 sessions, the contribution probabilities are significantly lower than those in treatment *B*. Thus on average, we do not robustly find support for hypothesis 2. Figures 1-4 compare period-wise average probabilities for the Delhi and Kolkata sessions. Figure 1 represents the time-series of average probabilities for those observations where the three treatments were played first (sequence 1). In figure 2, we tabulate the average probabilities for those observations in which these treatments were played second (sequence 2). Figures 3 and 4 do likewise for the Kolkata data. This allows us to check for order effects in section 5.1

On average the *B* and *I* treatments from the Kolkata sessions are for the most part indistinguishable as seen in figures 3 and 4 below. The time series from Delhi with the *D* treatment however (especially for observations where the treatments are played as second in the sessions, i.e. - figure 2) shows the baseline clearly above the *D* observations. Pooling the data over both locations by treatment, we find support for hypothesis H3, i.e. - the mixed strategy responses from treatment *D* are significantly lower than those in treatment *I.* See table 5 for the probability values from the t-tests that check for treatment effects.

Though they are not directly comparable it may be instructive to compare the results of our experiment with those of the standard linear public goods games in the literature such as Isaac and Walker (1988) and Kim and Walker (1984) and Ostrom et al. (1992). One stark observation that emerges from our data is that *the per-period contribution probability as a percentage of the optimal contribution probability does not show a decreasing trend over the experimental periods. Likewise the number of strong free riders is not seen to increase as the endgame approaches.* These observations hold true regardless of location, group size or the number of periods in the experiment. A possible explanation could be the small group size (two). It has been observed that success in sustaining cooperation in the Prisoner's Dilemma diminishes as group size increases. Dawes (1980) notes this for the NPD game (also see Kollock (1998) for an overview of dilemma games). Dufwenberg and Gneezy (2000) show that with two players, the static Bertrand prediction is not observed in the laboratory and prices stay well above the theoretical equilibrium even in the last period. However with three or four competitors the price tends towards the Nash prediction especially towards the end of the multi-period interaction.

5.1 Order, Location and Treatment Effects

We first check whether playing a treatment as the first one in a session yields different contribution probabilities than playing it second. Table 4 checks for these order effects.The two sided p-values from the parametric t-tests and the non-parametric Wilcoxon-Mann-Whitney tests are reported in table 4. We observe that for each treatment, the later periods (sequence 2) display lower probabilities of cooperation, though the difference is significant at the 5 percent level only for treatment *D*. However since both values are lower than the baseline average probability, we pool the observations.

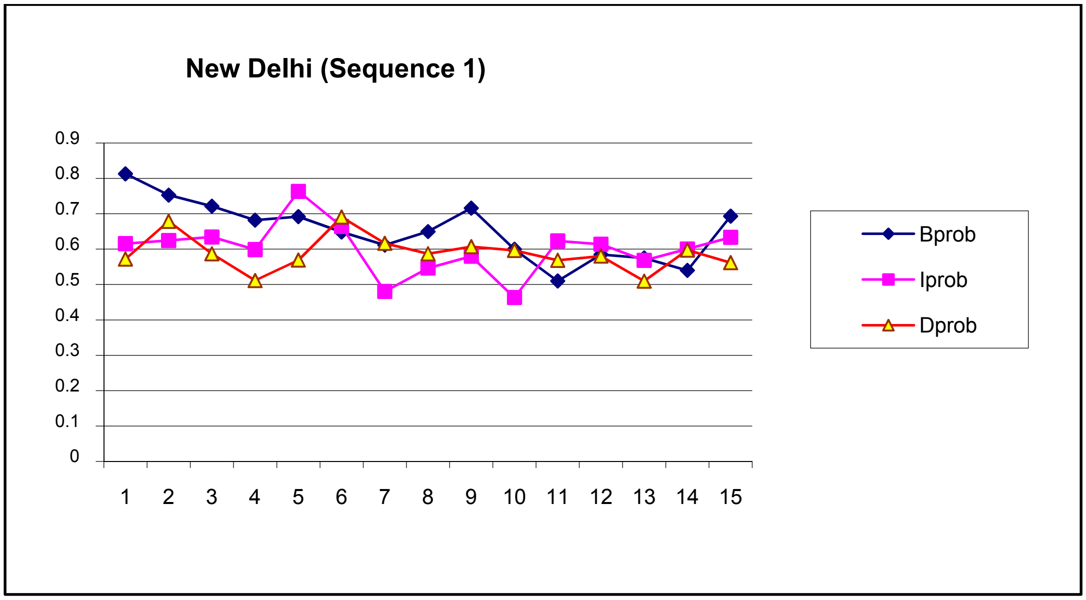

Figure 1. Delhi Sessions (pooled for periods 1-16)

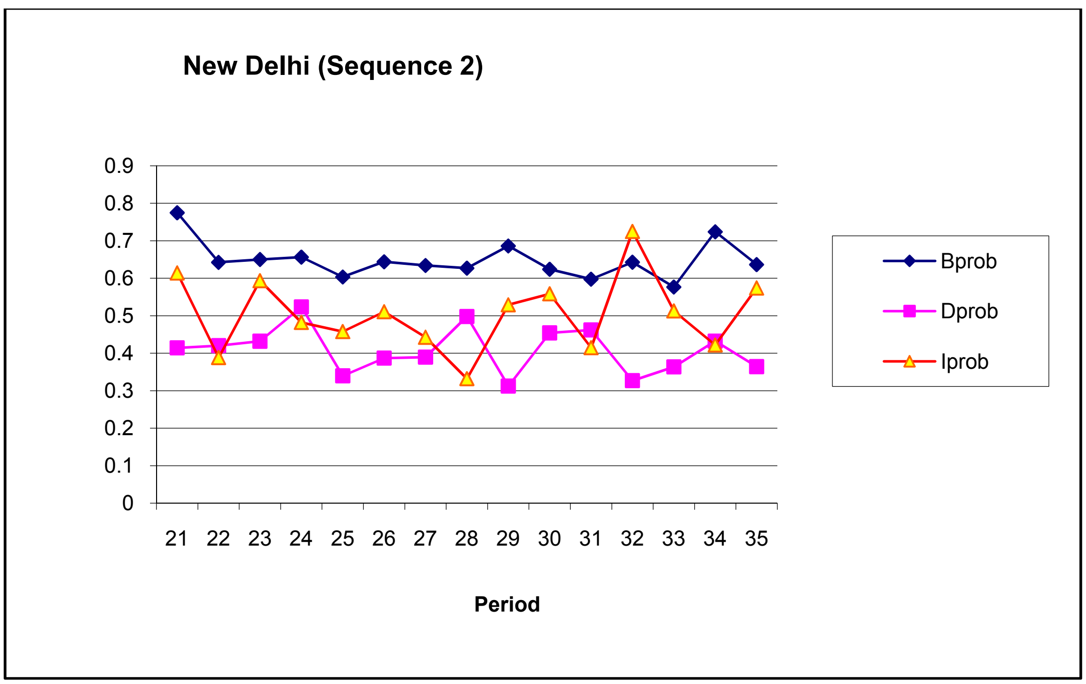

Figure 2. Delhi Sessions (pooled for periods 21-36)

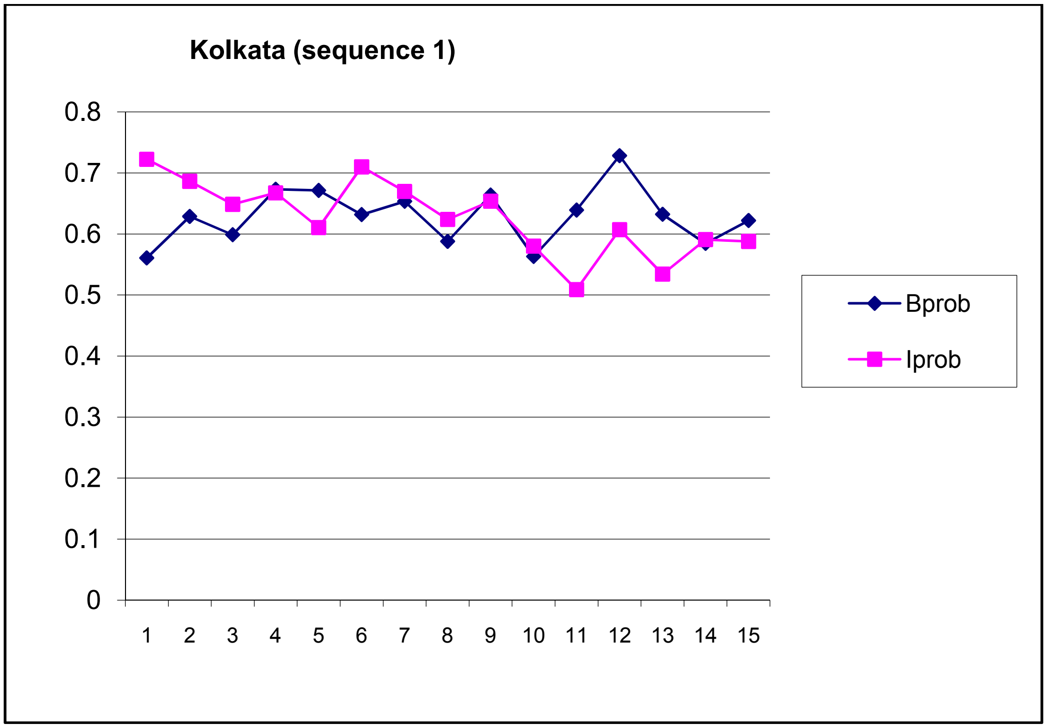

Figure 3. Kolkata Sessions (pooled for periods 1-16)

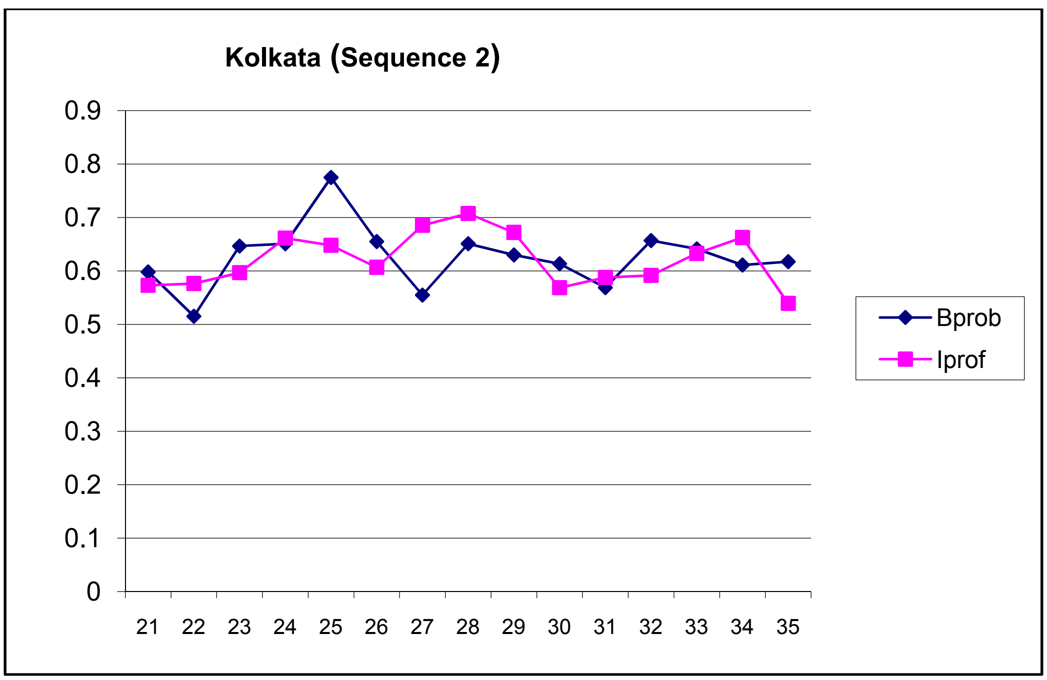

Figure 4. Kolkata Sessions (pooled for periods 21-36)

Table 4. Order Effects

Treatment	*Sequence 1*	*Sequence 2*	*Mean Difference*	*Std. Error*	*p-value (two-sided)*
Baseline (*B*)	0.6431	0.6341	0.0091	0.014	t-test 0.5045 Wilcoxon 0.6167
Decrease (*D*)	0.5827	0.3974	0.1852	0.002	t-test 0.0000 Wilcoxon 0.0000
Increase (*I*)	0.6235	0.5957	0.0278	0.017	t-test 0.0956 Wilcoxon 0.0815

Table 5. Pooled Treatment Effects

Treatment	*Decrease(Ave. = 0.4955)*	*Increase (Ave. = 0.6112)*
Baseline (Ave. = 0.6383)	0.1428 (0.0135)***	0.0270 (0.0107)**
Decrease (Ave. = 0.4955)		-0.1157 (0.0143)***

(*** = 1%, ** = 5%, standard errors in parentheses)

Pooling observations with no regard for location and order we calculate the mean differences in average probability over treatment groups in table 5. In the aggregate, notice that the baseline average is slightly greater than the average in the *I* treatment and is significantly different using a standard two-sided t-test. *On average, the imposition of a cut-off seems to lower contribution probability regardless of whether the treatment theoretically predicts it or not.* The presence of moderate to high levels of risk aversion in subjects may potentially lead to such an outcome. However, as our experiment does not provide a concavity calibration for an agent's utility, it is not possible for us to compare our observed investment probabilities with the appropriate risk-adjusted Nash equilibrium benchmarks. In addition to t-tests, we also perform a non-parametric Kruskal-Wallis test (a generalized form of the Wilcoxon-Mann-Whitney test) that checks if the mean probabilities for the three treatments are statistically different. The Chi-squared statistic (allowing for ties in ranks) is 98.5 with a probability value of 0.0001, indicating that the probabilities over the three treatments are significantly different. This compliments the results from the t-tests reported in table 5.

We explore the differences in the means of observed probabilities in different locations. As mentioned earlier B pools the whole treatment group B0, B1 and B2, while I includes both I0 and I1.Only treatments *B* and *I* can be compared across locations as there were no *D* sessions run in Kolkata. Notice also for the Delhi sessions, (which employed a higher quality cut-off) the *B* sessions have an average probability 7 percentage points higher than the *I* sessions (with a difference significant at the 1 % level) whereas in Kolkata the *B* probability is about a percentage point higher than the *I* probability and not statistically significant at the 1 % and 5 % levels.

Based on the above analysis, our main findings are given below:

Finding 1: The average per period mixed strategy probabilities in the laboratory are significantly lower than the MSNE predictions.

Finding 2: Both the decrease (*D*) as well as Increase (*I*) treatments yield contribution probabilities which are significantly lower than the baseline (*B*).

Table 6. Treatment Effects by Location

|Difference of means| (std. error)	*Kolkata Baseline (Ave. = 0.6321)*	*Delhi Decrease (Ave. = 0.4955)*	*Delhi Increase (Ave. = 0.5669)*	*Kolkata Increase (Ave. = 0.6254)*
Delhi Baseline (Ave. = 0.6442)	0.0121 (0.0135)	0.1486*** (0.0152)	0.0773*** (0.0206)	
Kolkata Baseline (Ave. = 0.6321)				0.0067 (.0132)
Delhi Decrease (Ave. = 0.4955)			0.0714** (0.0216)	
Delhi Increase (Ave. = 0.5669)				0.0585*** (0.0203)

(*** = 1%, ** = 5%, standard errors in parentheses)

6. Noisy Equilibria: Quantal Response

We see from the previous section that play in the experiment does not approach the MSNE prediction. In fact particularly for treatment *I*, behavior in the experiment is by and large in the opposite direction to what is predicted by theory. In this section we fit a noisy equilibrium model, namely the Quantal Response Equilibrium (QRE), given in McKelvey and Palfrey (1995), Goeree et al. (1995) and Anderson et al. (2002) to the data.

In game theory experiments, observed laboratory behavior is often responsive to costs associated with deviations from the Nash equilibrium. These effects are incorporated in the QRE in an approach that introduces noisy behavior via probabilistic choice models. The source of this noise could be preference shocks, experimentation, or actual mistakes in judgment. We adopt the probabilistic choice framework due to McKelvey and Palfrey (1995).

Here for each player *i* and each strategy $j \in \{1, 2, ..., J_i\}$ let $\sigma \in \sum$ denote a strategy profile. Accordingly, $\Pi_{ij}(\sigma)$ is the expected payoff to i of adopting the pure strategy s_{ij} when the other players use σ_{-i}. It is assumed that for each pure strategy s_{ij} there is an additional privately obtained payoff disturbance ε_{ij}. Thus the player chooses his strategy not on $\Pi_{ij}(\sigma)$ but on $\Pi'_{ij}(\sigma) = \Pi_{ij}(\sigma) + \mu_i \varepsilon_{ij}$, where $\mu_i \in \mathbf{R}^{++}$ is an error parameter that measures the importance of the epsilon shock and $\varepsilon_i = (\varepsilon i_1, \varepsilon i_2, ... \varepsilon_{iJi})$ has an *admissible* joint distribution.[14] The equilibrium condition has the consistency property (akin to rational expectations), that beliefs which determine expected payoffs match the choice probabilities that result from applying a probabilistic choice rule to those expected payoffs. For a 2x2 game, assuming identical and extreme value disturbances implies that P_{ij} is given by the logit formula:

$$P_{ij} = \frac{\exp(\pi_{ij})}{\exp(\pi_{ij}) + \exp(\pi_{ik})} \quad ... (3)$$

[14] An *admissible* joint distribution function is characterized by the properties of [1] absolute continuity [2] unbiasedness [3] independence across players (not necessarily across strategies). For a detailed discussion on the properties that constitute admissibility refer to McKelvey and Palfrey (1995).

With 2 players (players 1 and 2) and 2 strategies (*I* and *N*) for our specific game equation 2 expands out to four equations given below:

$$P_{1I} = \frac{\exp(\pi_{1I})}{\exp(\pi_{1I}) + \exp(\pi_{1N})} \quad \text{... (3')}$$

$$P_{1N} = \frac{\exp(\pi_{1N})}{\exp(\pi_{1I}) + \exp(\pi_{1N})} \quad \text{... (4')}$$

$$P_{2I} = \frac{\exp(\pi_{2I})}{\exp(\pi_{2I}) + \exp(\pi_{2N})} \quad \text{... (5')}$$

$$P_{2N} = \frac{\exp(\pi_{2N})}{\exp(\pi_{2I}) + \exp(\pi_{2N})} \quad \text{... (6')}$$

The fixed point of these four equations (3′ – 6′) yields the Quantal Response Equilibrium for the game. We estimate the value of μ_i by minimizing the sum of squared deviations of P_{ij}s from the observed values of probabilities from the laboratory. The QRE contribution probabilities thus estimated from our sessions are presented in table 3.[15] The average error level μ, estimated from the data varied between sessions and was around 50, which is high compared to the estimates obtained by Anderson et al. (2002) who obtain it for Normal Form games such as the Traveler's Dilemma and the Minimum Effort Coordination Game.

7. Learning Models

From section 5 it is clear there is no adherence to Nash play in the laboratory when the benchmark of comparison is the standard MSNE. When a noisy measure such as the QRE is used in section 6, we see that laboratory behavior implies that subjects are prone to large decision errors. Indeed we see that there is no tendency for individual strategies to converge to any steady state probabilities and thus we need to look for other explanations to explain observed play in the laboratory. Theories of learning may provide us with an approach to model such situations.

Reinforcement Learning constitutes an important class of models. In these, actions or strategies that have yielded relatively higher (lower) payoffs in the past are more (less) likely to be played in the future.[16] Reinforcement learning does not require any information on the part of a player about the play of other participants (except what is deductively revealed from the payoffs) or the adherence to game theoretic strategy. Thus, reinforcement learning involves a very minimal level of rationality that is only somewhat greater than that possessed by zero intelligence (ZI) agents. For our experiment we use Roth and Erev's (RE henceforth; Erev and Roth, 1998) reinforcement learning model.[17]

---

15 In one of the Baseline sessions (4/15/09) the solver algorithm did not find a fixed point and thus we have no converging (NC) logit equilibrium probability.

16 "Reinforcement," "stimulus-response" or "rote" learning follows closely from Thorndike's (1911) 'law of effect'. i.e. - actions or strategies that have yielded relatively higher (lower) payoffs in the past are more (less) likely to be played in the future.

17 Another important class comprises the belief based learning models. The reason why we cannot use a belief based approach is that we do not provide a player with any information regarding the likelihood of

7.1 The Roth-Erev (RE) One and Three Parameter Models

We discuss the RE one and three parameter models in this section. Let us first describe the one parameter model. Suppose there are N actions/pure strategies in a Normal Form game. In round t, player i has a propensity $q_{ij}(t)$ to play the j^{th} pure strategy.

Initial (round 1) propensities among players who are in the same role are equal, i.e. -

$$q_{ij}(1) = q_{ik}(1) \quad \text{for all } j, k \qquad \ldots(7)$$

and

$$\sum_j q_{ij}(1) = S_i(1),$$

where $S_i(1)$ is an initial strength parameter, equal to a constant that is the same for all players, $S_i(1) = S(1)$; the higher (lower) is S(1) the slower (faster) is learning.

If player i plays his k^{th} pure strategy at time t and receives a reinforcement of R(x), where $R(x) = x - x_{min}$, then the propensity to play strategy j is updated by setting

$$q_{ij}(t+1) = \begin{cases} q_{ij}(t) + R(x) & \text{if } j = k \\ q_{ij}(t) & \text{otherwise} \end{cases} \qquad \ldots(8)$$

The probability that agent i plays strategy j in period t is made according to the linear choice rule,

$$p_{ij}(t) = \frac{q_{ij}(t)}{\sum_{j=1}^{N} q_{ij}(t)} \qquad \ldots(9)$$

For the three-parameter model, suppose that, in round t, player i plays strategy k and receives payoff of x. Then i updates his propensity to play action j according to the rule:

$$q_{ij}(t+1) = (1-\varphi)\, q_{ij}(t) + E_k(j, R(x)) \qquad \ldots(10)$$

$$E_k(j, R(x)) = \begin{cases} (1-\varepsilon)R(x) & \text{if } j=k \\ (\varepsilon/(N-1))R(x) & \text{otherwise} \end{cases} \qquad \ldots(11)$$

Here the parameters are:

(a) The initial strength parameter S(1) which determines the speed of learning as in the one-parameter model.

(b) A “forgetting” parameter φ that gradually reduces the role of past experience.

(c) An experimentation parameter ε that allows for some experimentation among strategies.

Our data is slightly different from RE, as unlike the Normal Form games they study where only pure strategies are permissible, our game actually allows players to specify mixed strategies.

---

investment choice of the other player in any specific period. We simply provide to him the investment outcome of the other player. This impedes the formation of beliefs necessary to implement the belief-based approach. For a detailed survey of learning in human subject experiments see Duffy (2004).

Hence contribution to the public good is not a 0/1 choice but a probability. For degenerate probabilities, the propensity for only the strategy that is played in a period gets the full reinforcement (full reinforcement of either the Investment (Inv) or the No-Investment (NoInv) strategy), but for an interior probability, both strategies get reinforced in the proportion specified in the mixed strategy. For example in the one-parameter model, if $q_{iI}(t) = 100$ and $q_{iN}(t+1) = 250$ units of experimental currency. Suppose in period t, the player specifies the likelihood of investment to be p(t) = 0.60. Also suppose his payoff for period t is 100 units of experimental currency. Then in the next period,

$q_{iI}(t+1) = 100 + 0.60 \text{ X } 100 = 160$ and,

$q_{iN}(t+1) = 250 + (1\text{-}0.60) \text{ X } 100 = 290$

We use the same principle for updating the propensities in the three-parameter model.[18] Thus the updating rule for propensities in the one and three parameter models become:

$$q_{ij}(t+1) = \begin{cases} q_{ij}(t) + p(t)R(x) & \text{if } j = I \\ q_{ij}(t) + (1-p(t))R(x) & \text{if } j = N \end{cases}$$

This is a natural extension from the original pure strategy updating rule and reduces to that in the RE model when p(t) = 0 or 1.We use the data from both locations, which comprises between 36 and 40 periods of play for 100 subjects. Like RE, we fit the one and three parameter models to the data and use the Minimum Mean Squared Deviation (MSD) criterion to select parameters which best fit our data. Using a grid search for the one-parameter and three parameter models we obtain the values of S(1), φ and ε, which minimize the mean square deviation (MSD) from the observed probabilities (See table 7). The fit for these models are shown graphically in figures 5 (Delhi sessions) and 6 (Kolkata sessions).

Table 7. Reinforcement Learning Parameters

Parameters	*RE 1 parameter Kolkata*	*RE 3 parameter Kolkata*	*RE 1 parameter New Delhi*	*RE 3 parameter New Delhi*
S(1)	152.06	66.94707	223.7032256	40.13907586
φ		0.155945		0.311041988
ε		0.094425		0.133002187

Notice that the three-parameter model fits the observed behavior much better than the one parameter model.

***Finding 3*:** The RE three-parameter model tracks observed behavior better than the RE one-parameter model and static Nash equilibrium.

[18] Notice we use profits and not profit less minimum profit from last period in order to update the propensities in the present period. This is because the minimum per period profit (10, 8 or 5 units of laboratory currency) is very negligible in our study and would make a small difference to the speed of learning only initially, but over 40 periods, the impact of subtracting out this small number would become more and more negligible.

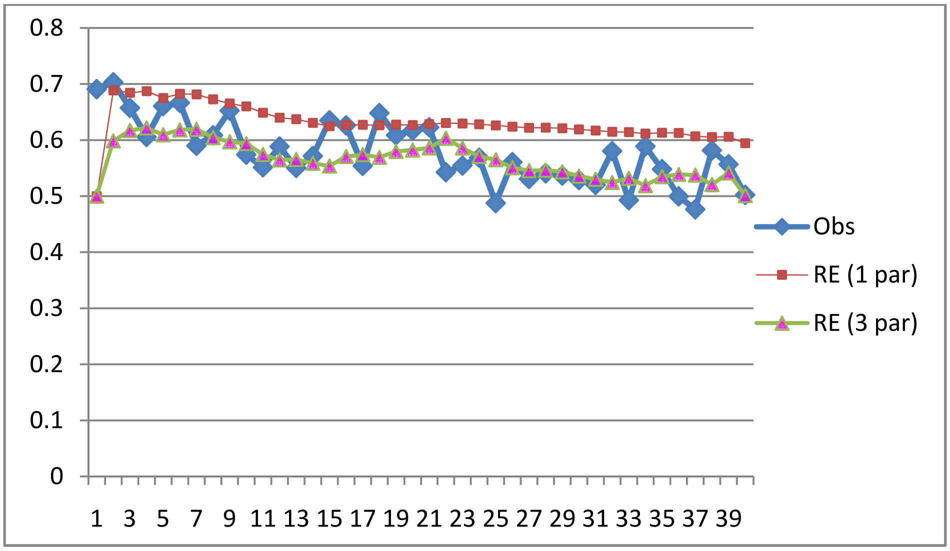

Figure 5. Delhi Sessions (Pooled)

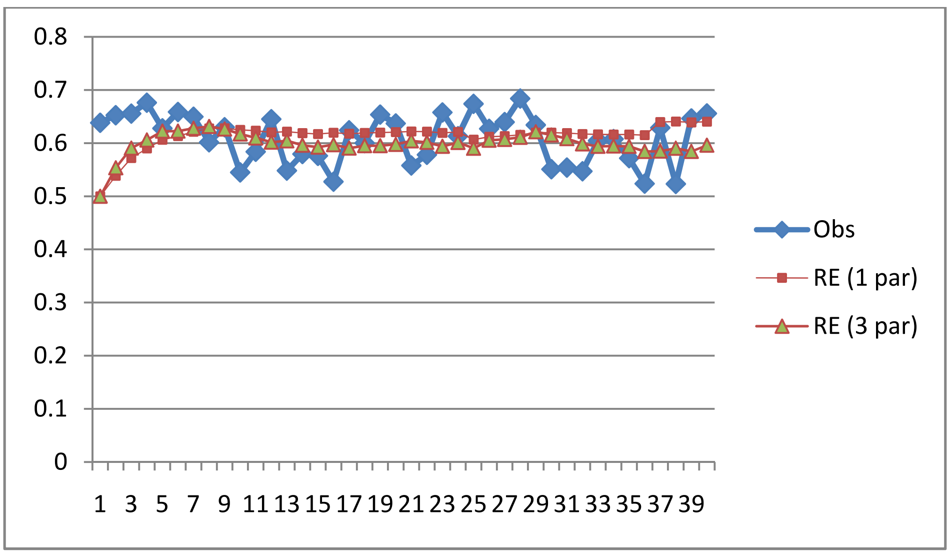

Figure 6. Kolkata Sessions (Pooled)

8. Conclusions and Discussion

We study a two-player stochastic public goods game with equilibrium contribution in mixed strategies. Our game has the feature that the quality of the good produced by at least one player has to be greater than a certain threshold or neither player makes more than a small level of benefit. Under such situations it is theoretically possible to show that the presence of this threshold increases the Nash equilibrium contribution level of players.

We test this game in the laboratory over two locations, and find that contrary to the levels of contribution predicted by Nash theory, the presence of this cut-off does not significantly raise contribution probabilities. A striking observation is that the observed behavior in the laboratory does not show convergence to any particular strategy. A noisy Nash equilibrium, provided by using a probabilistic choice (logit) model gives us equilibria close to that observed mixed strategy probabilities with a significant amount of decision error. The absence of any steady state behavior coupled with very noisy decision making may indicate that subjects are adaptively learning from their history of play in the previous periods. We fit a one-parameter and a three-parameter reinforcement learning model (RE) and find that the three-parameter RE model tracks observed behavior in the laboratory better than either the Mixed Strategy Nash prediction (MSNE), the Quantal Response Equilibrium (QRE) or the one-parameter RE model.

References

Anderson, S., J. Goeree and C. Holt (2002) "The Logit Equilibrium: A Perspective on Intuitive Behavioral Anomalies", *Southern Economic Journal*, 69: 21-47.

Banerjee, P. (2007) "Collective Punishments: Incentives and Examinations in Organizations", *Berkeley Electronic Journal: Contributions to Theoretical Economics*, 7: 34.

Berger, L. and J. Hershey(1994) "Moral Hazard, Rent Seeking and Free Riding", *Journal of Risk and Uncertainty*, 9: 173-186.

Bliss, C. and B. Nalebuff (1984) "Dragon-slaying and Ballroom Dancing: The Private Supply of a Public Good", *Journal of Public Economics*, 25: 1-12.

Bloomfield, R. (1994) "Learning a Mixed Strategy Equilibrium in the Laboratory", *Journal of Economic Behavior and Organization*, 25: 411–436

Chowdhury, S., D. Lee and R. Sheremeta (2011) *Top Guns May Not Fire: Best-Shot Group Contests with Group-Specific Public Good Prizes*, University of East Anglia working paper.

Cornes, R. (1993) "Dyke Maintenance and Other Stories: Some Neglected Types of Public Goods", *Quarterly Journal of Economics*, 108: 259–271.

Dawes, R. (1980), "Social Demmas"" *Annual Review of Psychology*, 31: 169-193.

Duffy, J. (2004) "Agent-Based Models and Human Subject Experiments", in K. Judd and L. Tesfatsion, Eds., *Handbook of Computational Economics vol. 2*, Amsterdam: Elsevier.

Dufwenberg, M. and U. Gneezy (2000) "Price Competition and Market Concentration: An Experimental Study", *International Journal of Industrial Organization*, 18: 7 –22.

Erev, I. andA. Roth (1998), "Predicting How People Play Gmes: Reinforcement Learning in Games with Unique Mixed Strategy Equilibria", American Economic Review, 88: 848-881.

Fischbacher, U. (2007) "Z-tree 2.1: Zurich Toolbox for Readymade Economic Experiments", *Experimental Economics*,10: 171-178.

Goeree, J., C. Holt and T. Palfrey (2005) "Regular Quantal Response Equilibrium", *Experimental Economics*, 8: 347-367.

Harrison, G. and J. Hirshleifer (1989) "An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods", *Journal of Political Economy*, 97: 201-225.

Heston, A., R. Summers and B. Aten (2011) *Penn World Table Version 7.0*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, March 2011.

Hirshleifer, J. (1983) “From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods”, *Public Choice*, 41: 371-386.

Hirshleifer, J. (1985) “From Weakest-Link to Best-Shot: Correction”, *Public Choice*, 46: 221-223.

Isaac, R. and J. Walker (1988) “Group Size Effects of Public Goods Provision: An Experimental Examination”, *Quarterly Journal of Economics*, 103: 179-199.

Kim, O. and M. Walker (1984) “The Free Rider Problem: Experimental Evidence”, *Public Choice*, 43: 3–24.

Kollock, P. (1998) “Social Dilemmas: The Anatomy of Cooperation”, *Annual Review of Sociology*, 24: 183-214.

Lipnowski, I. and S. Maital (1983) “Voluntary Provision of a Pure Public Good as the Game of 'Chicken”, *Journal of Public Economics*, 20: 381-386.

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

Ochs, J. (1994) “Games with Unique Mixed Strategy Equilibria: An Experimental Study”, *Games and Economic Behavior,* 10: 202–217.

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

Ostrom, E., J. Walker and R. Gardner (1992) “Covenants With or Without a Sword: Self Governance is Possible”, *American Journal of Political Science,* 86: 404-417.

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

Ostrom, E., R. Gardner and J. Walker (1994b) “Institutional Analysis and Common-Pool Resources”, in Ostrom et al, Eds., Rules, Games and Common-Pool Resources, Ann Arbor: University of Michigan Press.

Shachat, J. (2002) “Mixed Strategy Play and the Minimax Hypothesis”, *Journal of Economic Theory*,104: 189-226.

Thorndike, E. (1911) *Animal Intelligence,* New York: Hafner Publishing.

Wooders, J and J. Shachat (2001) “On the Irrelevance of Risk Attitudes in Repeated Two-Outcome Games”, *Games and Economic Behavior,*34: 343–363.

♦♦♦