Research

Learning, Signaling, and Social Preferences in Public-Good Games

Marco A. Janssen[1] and T. K. Ahn[2]

ABSTRACT. This study compares the empirical performance of a variety of learning models and theories of social preferences in the context of experimental games involving the provision of public goods. Parameters are estimated via maximum likelihood estimation. We also performed estimations to identify different types of agents and distributions of parameters. The estimated models suggest that the players of such games take into account the learning of others and are belief learners. Despite these interesting findings, we conclude that a powerful method of model selection of agent-based models on dynamic social dilemma experiments is still lacking.

Key Words: *laboratory experiments; public goods; agent-based model; learning; social preferences*

INTRODUCTION

Social dilemmas are situations in which behavior that is rational for and in the self-interest of individuals results in socially suboptimal outcomes. Most environmental problems, such as clean air, the management of common-pool resources, recycling, etc., involve social dilemmas.

Experimental research has contributed a great deal to the understanding of the factors that affect the level of cooperation in repeated social-dilemma games, such as games that provide public goods and common-pool resources (CPR). Many scholars now agree that some players do not seem to be interested in maximizing their own incomes, that players are heterogeneous on several dimensions, and that rates of cooperation are affected by the pay-off functions, matching protocols, and other institutional features of experimental treatments (Ledyard 1995, Ahn et al. 2003, Ones and Putterman 2004).

To date, however, no widely accepted models of individual decision making exist that provide the micro-foundations for such empirical regularities. What are the motivations and learning rules used by the players? Do players differ from one another in significant ways? If so, what are the key dimensions of such heterogeneity? To what extent, and in what manner, do players in repeated social-dilemma games learn from past experiences during the game?

Do some players behave strategically to increase their future pay offs? In sum, we need micro-level models of individual behavior that are open to heterogeneity across players to advance our knowledge of the dynamics in repeated social-dilemma games.

Significant progress has been made by behavioral economists, game theorists, and experimentalists who have developed rigorous models of behavior in game settings and tested these with controlled experiments with human subjects (see Camerer 2003 for a review). These models are often called models of learning in the sense that they explain the emergence of equilibrium over time. Social-dilemma research can greatly benefit from taking these efforts seriously when providing micro-level explanations of macro-level regularities in *N*-person social-dilemma games.

On the other hand, the study of learning models can expand its horizons greatly by taking *N*-person social-dilemma games seriously. Most of the learning models have been applied to rather simple games, which is understandable given that the formulation and testing of those learning models are often very complicated tasks. The increasing level of sophistication in the formulation of the models and their tests now allows us to expand the horizon of learning models to more complicated game settings. *N*-person social-dilemma games can test

---

[1]Arizona State University, [2]Florida State University and Korea University

alternative learning models in more demanding contexts and provide an opportunity to develop more relevant models of behavior.

This paper attempts to expand the horizon of behavioral models to repeated *N*-person social-dilemma games. Specifically, this study compares the empirical performance of several alternative learning models that are constructed based on the social preferences model of Charness and Rabin (2002) and the behavioral learning model of Camerer and Ho (1999). The models are tested with experimental data drawn from the public-good experiments by Isaac and Walker (1988) and Isaac et al. (1994).

Motivated by experimental observations that are not consistent with equilibrium predictions, researchers have developed models of learning in which players learn to play the equilibria as a game repeats. Earlier efforts to model a learning process in repeated games include reinforcement learning or routine learning (Bush and Mosteller 1955, Cross 1983), fictitious play and its variants (Robinson 1951, Fudenberg and Kreps 1993, Young 1993, Fudenberg and Levine 1995, Kaniovski and Young 1995), and replicator dynamics (Fudenberg and Maskin 1990, Binmore and Samuelson 1992, 1997, Ellison and Fudenberg 1993, Schlag 1998). To test the explanatory power of these models more rigorously, many game theorists and experimentalists began to use specific experimental data (Crawford 1995, Crawford and Broseta 1998, Cheung and Friedman 1997, Broseta 2000). However, these studies tend to test a single model, usually by estimating the parameters of a specific model using a set of experimental data.

More recently, researchers began to compare the explanatory power of multiple models using data from multiple experimental games, which represented a step forward from the previous approaches. These studies include Boylan and El-Gamal (1993), Mookherjee and Sopher (1994, 1997), Ho and Weigelt (1996), Chen and Tang (1998), Cheung and Friedman (1998), Erev and Roth (1998), Camerer and Ho (1999), Camerer and Anderson (2000), Feltovich (2000), Battalio et al. (2001), Sarin and Vahid (2001), Tang (2001), Nyarko and Schotter (2002), Stahl and Haruvy (2002), and Haruvy and Stahl (2004). Another noticeable aspect of the current research is the careful examination of the testing methods themselves and the use of multiple criteria of goodness-of-fit (Feltovich 2000, Bracht and

Ichimura 2001, Salmon 2001). Thus, the research has progressed from parameter fitting of a single model to rigorous testing of alternative models in multiple game settings and to careful examination of testing methods.

We extend this comparative approach in several ways. First, the decision-making setting that we study involves the provision of public goods in which the predicted equilibrium of zero contribution has repeatedly been shown to misrepresent actual behavior. Thus, in framing our research, we find that "learning" is not necessarily the general theme. There are dynamics at individual and group levels. However, it is still an open question, especially in repeated social dilemma games, whether those dynamics result from learning or other mechanisms such as forward-looking rational and quasi-rational choices. In general, we entertain the hypothesis that heterogeneity across players on multiple continuous dimensions is the key aspect of the micro-foundations that generate the observed dynamics in repeated public-good games.

Second, several other factors posed challenges to estimating model parameters and developing goodness-of-fit measures. They included (1) the large number of players, which ranged from four to 40 in our data; (2) the number of the stage game strategies for each player, which varied from 11 to 101 in our data; and (3) the variation in the number of rounds, which ranged from 10 to 60 in our data. Previous studies have used various estimation methods such as regression (Cheung and Friedman 1998), maximum-likelihood gradient search (Camerer and Ho 1999), and grid search (Erev and Roth 1998, Sarin and Vahid 2001). A number of recent studies show that structural estimation of the true parameters using regression methods is problematic for modestly complicated models (Bracht and Ichimura 2001, Salmon 2001, Wilcox 2006). Salmon shows that maximum-likelihood estimation of learning models is not capable of discriminating among contending learning models. Econometric approaches that assume a "representative player" lead to serious biases in the estimated parameters when there is structural heterogeneity across the players (Wilcox 2006). With such problems in mind, we can perform only some modest comparative analyses. In this study, maximum-likelihood estimation of representative agents is used as a starting point, but we also compare alternative models in terms of their performance of macro-level metrics.

Third, the experimental results of public-good games are multilevel. This poses the question of which aspects of the experimental data need to be explained. Using only average behavior as the target of calibration may severely distort empirical tests in the public-good games. This is because the same average can result from widely different combinations of strategies at the player level. In addition, players change their contributions over time, some quite frequently and dramatically, others not so often and in small steps. We develop multiple indicators that characterize behavior at individual and group levels and changes in behavior over time. These include average contribution level, variance across individual contribution in a given round, and variability of change in contribution between rounds at the individual level.

Fourth, the analyses performed in this paper provide a number of examples of how to develop and test agent-based models using experimental data. Although behavioral game theorists estimate their formal models in a similar fashion, we focus on heterogeneity within the player population and on determining how to estimate and formalize this heterogeneity. Agent-based modelers are also interested in macro-level results of agent-agent interactions. Therefore we also compare the macro-level patterns between our empirical data and the simulated data based on the tested models.

The remaining sections of this paper are organized as follows. In the second section, we discuss the experimental environment of linear public-good games. We use experimental data from Isaac and Walker (1988) and Isaac et al. (1994) and discuss the main stylized facts from that data set. In the third section, we present the formal model in detail. This model combines basic models of other studies such as the experience-weighted attraction model of Camerer and Ho (1999) and the hybrid utility formulation of Charness and Rabin (2002), and we formalized the signaling process suggested by Isaac et al. (1994). In the fourth section, we report parameter estimates using maximum-likelihood estimation. We applied maximum likelihood to different levels of scale, including the representative agent, different types of agents, and the level of the individual. We summarize our findings and suggest directions for further research in the final section.

LINEAR PUBLIC-GOOD PROVISION GAMES

This section introduces the notations related to *N*-person linear public-good games and reviews the most prominent features of behavior at both individual and group levels in such experiments. We will use experimental data from Isaac and Walker (1988) and Isaac et al. (1994) throughout this paper.

Public-good provision experiments

The standard linear public-good provision experiment (Marwell and Ames 1979, 1980, 1981, Isaac et al. 1984, 1985, 1994, Isaac and Walker 1988, to name only some of the pioneering researchers) can be characterized by the number of players (N), the marginal per capita return (r), the number of repetitions (T), and the initial endowment for each player (ω). An experimental linear public-good provision game involves a free-rider problem if $r < 1$ and $N * r > 1$.

Suppose that, in a given round, player i contributes x_i of ω for the provision of the public good. His monetary pay off (π_i) is:

$$\pi_i = \alpha \cdot \left\{ \omega - x_i + r \sum_{i=1}^{N} x_i \right\} \tag{1}$$

in which α is the conversion rate by which monetary earnings are calculated from experimental endowment units such as "tokens" and experimental pay offs. The equilibrium prediction, assuming that players maximize their own monetary pay offs, is that the public good will not be provided at all. This prediction still holds when the situation is repeated for a known finite number of rounds. However, experimental studies regularly find that, in such experiments, public goods are provided at substantial, though usually suboptimal, levels. In addition, many aspects of the experimental results seem to vary systematically depending on the aforementioned experimental parameters, such as the size of group and the marginal per capita return.

Stylized facts from the public-good games: What needs to be explained?

We present three observations or stylized facts from linear public-good provision that any attempt to offer coherent theoretical explanations should address. The stylized facts are illustrated with data on the six experimental treatments, defined by the marginal per capita return (*MPCR* hereafter), and the group size, shown in Figs. 1 and 2.

Observation 1. The time course of the average contribution at the group level is a function of group size and the *MPCR*.

The average level of contribution for public-good provision and its change over time differs across experimental settings. Some extreme experimental conditions with low *MPCR* show a rapid convergence to an almost complete free-riding, whereas other treatments with relatively high *MPCR* show a pattern of stabilization of the contribution level at approximately 50% of the total endowment. Still other experimental conditions exhibit trends in between these two extremes, typically showing an overall decrease in contribution level. Experiments with longer durations of 40 or 60 rounds (Fig. 2) also show declining trends toward zero. Controlling for *MPCR,* it appears that, the larger the group size, the higher the contribution level. This can be seen most clearly in Fig. 1 when one compares three treatment conditions. For an *MPCR* of 0.3, groups of size 4 (filled diamond) show the lowest contribution, groups of size 10 (filled triangle) show a noticeable increase in contribution level compared to that of groups of size 4, and groups of size 40 show contribution levels of around 50% without a clear declining trend. However, this apparently benign effect of group size is not present for the *MPCR* value of 0.75. Both groups of size 4 and 10 show very similar trends of contribution when the *MPCR* is 0.75.

Observation 2. For a given level of average contribution in a round, there is a substantial variance in the level of contribution at the individual level.

Variance in contribution levels across players in a given round is another important factor characterizing public-good experimental results. In some rounds, all players contribute a similar proportion of their endowment; obviously, this is more likely when the average contribution is near zero. In other rounds, there is a diversity of contribution levels ranging from 100% to 0. An interesting observation comes from a session in Isaac et al. (1985), with *MPCR* = 0.3 and group size 40. The players in the session were all experienced. As Fig. 3 shows, there is a tendency for contribution levels to bifurcate toward the extremes of 0 and 100% over time. In the experimental session, about 20% of players contribute all of their endowments to the public-good account. This type of complete contributors increases to 40% by the final round of the experiment. At the same time, the proportion of complete free-riders also increases from 10% in the first round to more than 30% in the 10th. Thus, by the final round, the complete contributors and the complete free-riders together comprise more than 70% of the group. This micro-mechanism generates the stable group-level contribution shown in Series (40, 0.3), marked by hollow circles, in Fig. 1, with increasing variance shown in the corresponding series in Fig. 4.

Observation 3. Players change contribution levels between rounds. The extent and direction of such changes vary across players. Variability across players and between rounds for a player appears to be dependent on the experimental parameters and the number of rounds remaining.

Third, the variability of contribution across rounds differs from one player to another. Some players change their contribution levels rather dramatically between rounds; others maintain relatively stable levels of contribution across rounds. From the perspective of agent-based modeling, we are interested in seeing whether we can observe patterns and distributions at the population level. Figure 5 shows the relative change in contribution levels at the player level between rounds. We derived this figure by calculating for each observation the relative change in contribution between every two rounds. Thus, when a player invested 10 tokens in one round and six in the subsequent round, we registered 40% for this agent between these two rounds. This was done for all rounds and for all agents. We then calculated the relative frequency of the occurrence of different categories of change, e. g., -100% to -95%, -95% to -85%, -85% to -75%, etc. By doing this we derived a distribution. We plotted the relative frequencies on a logarithmic scale to emphasize the observed distribution. We saw a dominance of situations in which players did not change their investment, but also a relatively

Fig. 1. Average contribution over the round for the six different treatments. The number of participants and the marginal per capita return are given in parentheses.
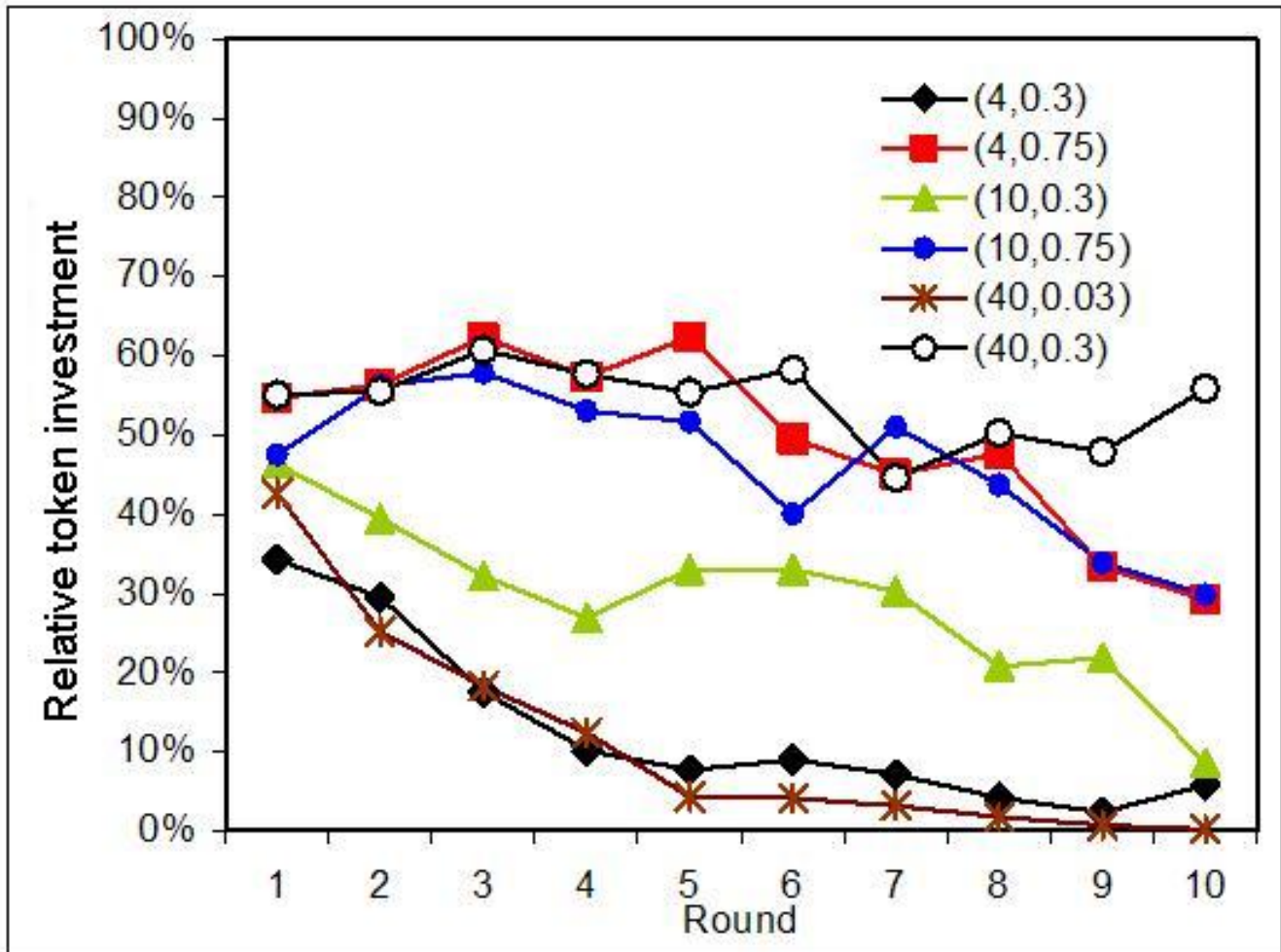

high frequency of situations in which the players changed their investments by 100%. The challenge for the model exercise is therefore not only to replicate the data at an aggregate level but also to generate results that incorporate between-player variability and variability over time at the player level.

MODEL

A general formal model is presented in this section that represents the decision making of agents in social dilemmas. The model will be tested on experimental data. The model is built on three components: (1) the probabilistic choice model to define the decision, (2) the learning model that captures the change in behavior over time at the player level, and (3) the social utility function by which a player evaluates outcomes of the game. The social utility function is embedded in the learning model, which in turn is embedded in the probabilistic choice model that determines the relative probabilities of choosing different levels of contribution.

Fig. 2. Average contribution per round for two different treatments. The number of participants, the marginal per capita return, and the number of rounds are given in parentheses.
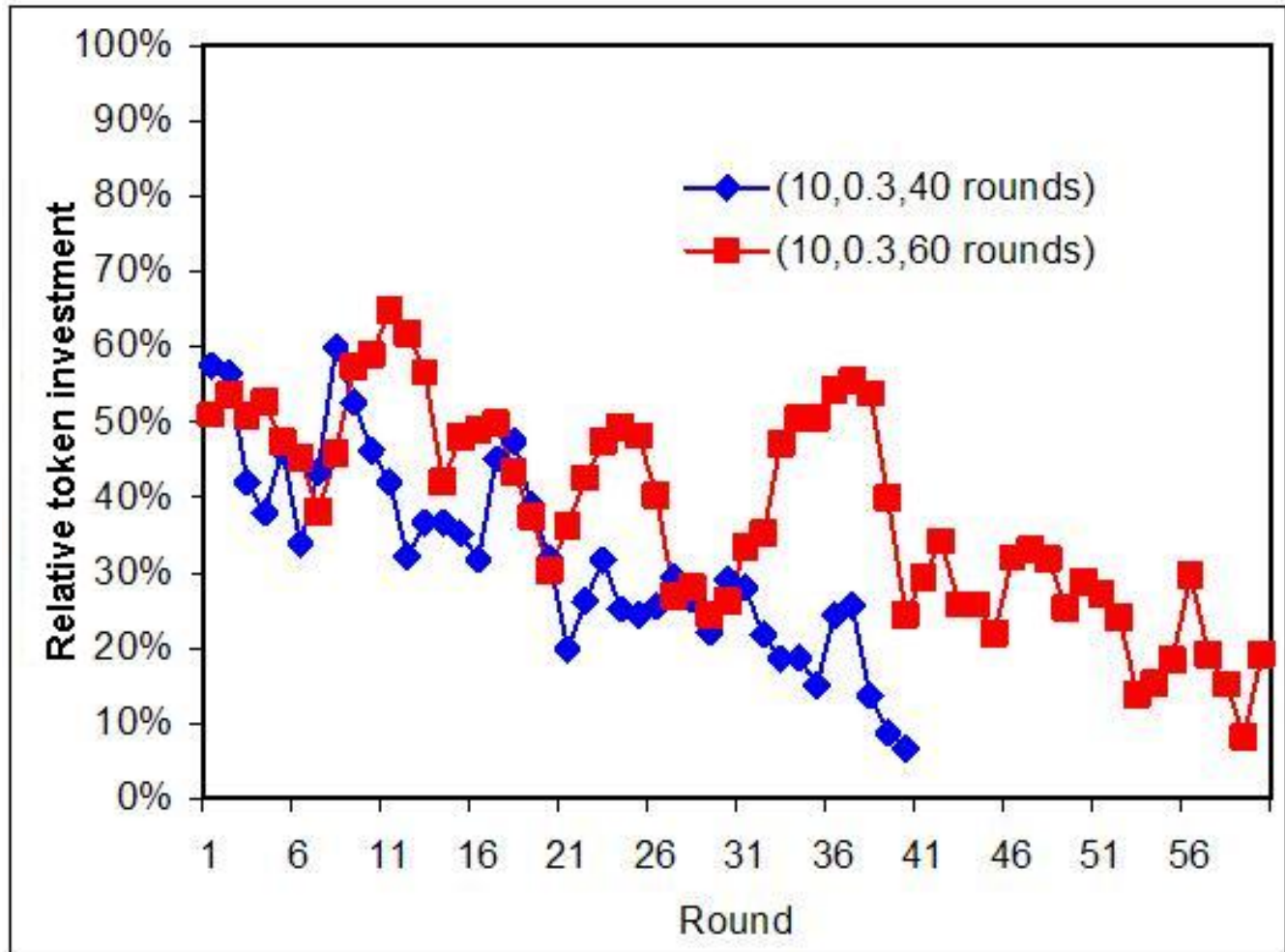
Probabilistic choice

The general structure of probabilistic choice is the same across several models that we test. Let P_i^x denote the probability that agent i contributes x units of total endowment ω for the public-good provision. Then,

$$P_i^x = \frac{e^{\varphi_i \cdot A_i^x}}{\displaystyle\sum_{j=0}^{\omega} e^{\varphi_i \cdot A_i^x}} \tag{2}$$

in which the parameter φ_i is called the response sensitivity and A_i^x is the attraction of choice x to agent i. A high value of φ_i leads to a sharper discrimination among strategies, and a high value of A_i^x implies that contribution level x has a large chance of getting chosen by agent i.

Learning behavior

The way players learn in repeated public-good games is modeled as the updating of attraction parameter A_i^x. The learning model is based on the experience-weighted attraction (EWA) model of Camerer and Ho (1999). This model assumes that

Fig. 3. Increasing variance across players over time in a public-good experiment with marginal per capita return equal to 0.3 and group size equal to 40. There seem to be a bifurcation within the population of those who invested everything and those who invested nothing in the public good.
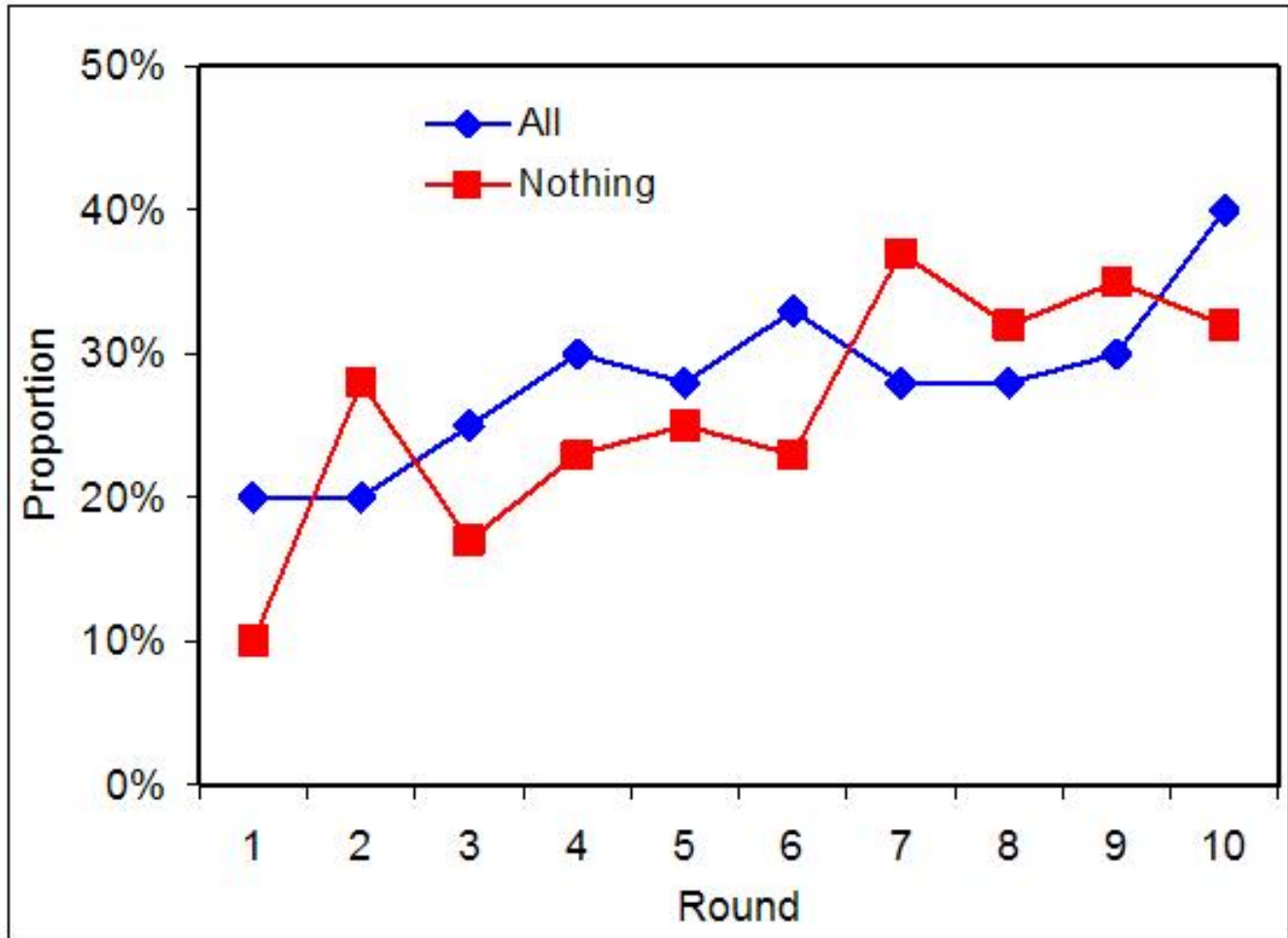

each strategy has a numerical attraction that affects the probability that it will be chosen. Agent i's attraction to strategy x, i.e., contribution of x units, in round t is denoted as $A^x_i(t)$. The initial attraction of each strategy is updated based on experience. The variable $H(t)$ in the experience-weighted attraction (EWA) model captures the extent to which past experience affects an agent's choice. The variables $H(t)$ and $A^x_i(t)$ begin with initial values of $H(0)$ and $A^x_i(t)$. The value of $H(0)$ is an agent-specific parameter to be calibrated. Updating is given by two rules. First,

$$H(t) = H(t-1)\lambda_i(1-\kappa_i)+1 . \qquad (3)$$

The parameter λ_i represents forgetting or discounting of the past experience, and κ_i determines the growth rate of attractions. Together they determine the fractional impact of previous experience. The second rule updates the level of attraction as follows. The model weighs hypothetical pay offs that unchosen strategies would have earned by parameter δ_i and weighs pay offs actually received by an additional $1 - \delta_i$. Define an indicator function $I(x, y)$ to be 0 if $x \neq y$ and 1 if $x = y$. The EWA attraction updating equation is the

Fig. 4. Average standard deviation of the contribution level within a group for the six different treatments. The number of participants and the marginal per capita return are given in parentheses.
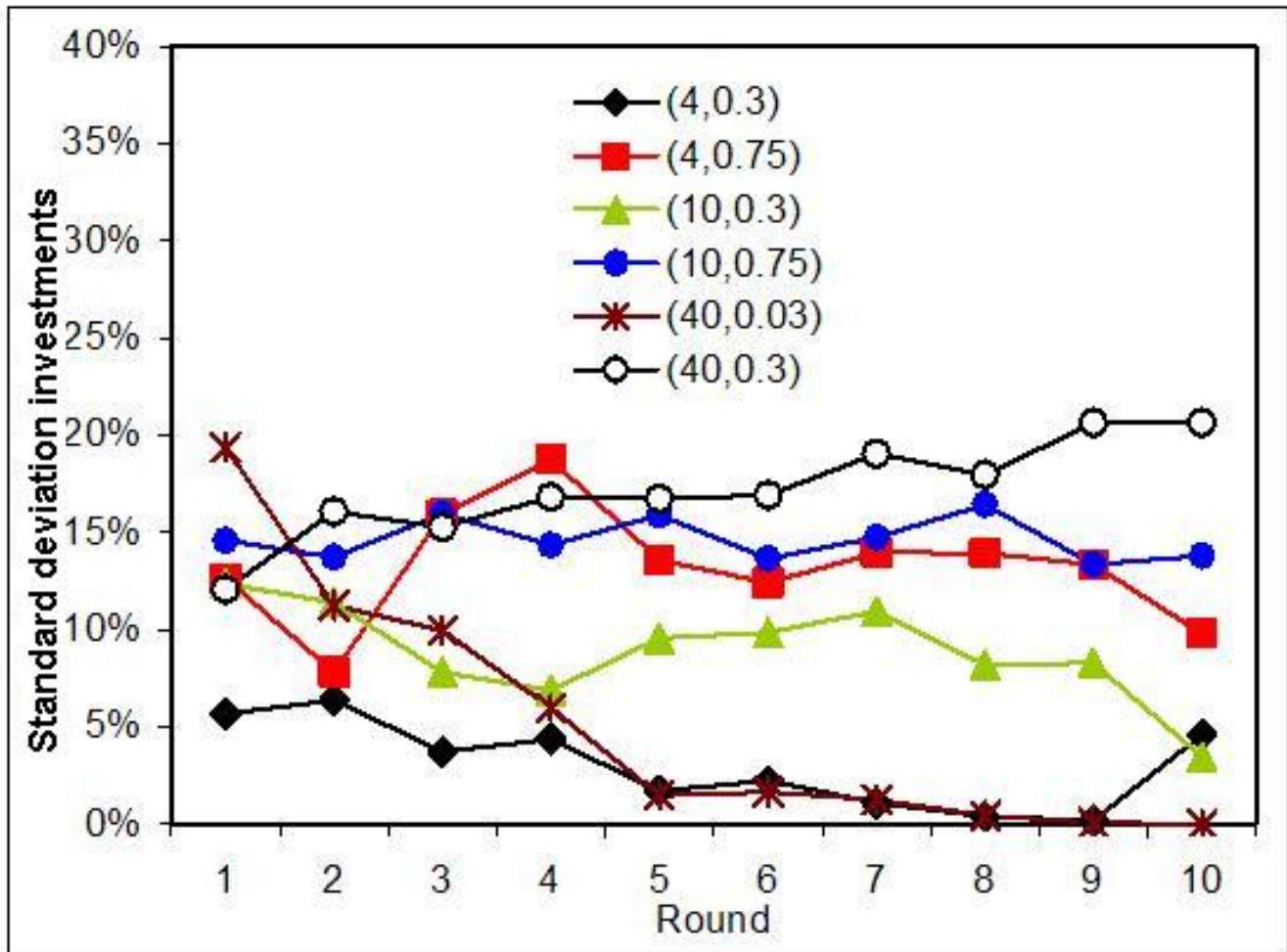
sum of a depreciated experience-weighted previous attraction plus the weighted pay off from period t, normalized by the updated experience weight,

$$A_i^x(t) = \frac{\lambda_i H(t-1) A_i^x(t-1) + [\delta_i + (1-\delta_i) I(x_i, x_i(t))] u_i(x_i, x_{-i}(t))}{H(t)} \quad (4)$$

The parameter λ_i is a discount factor that depreciates previous attraction. When δ_i is equal to 0, EWA mimics reinforcement learning as used by Erev and Roth . When δ_i is equal to 1, the model mimics belief learning as used by Sarin and Vahid (1999).

Following Wilcox (2006), we assume that the initial value of H is

$$\frac{0.5}{1-(1-\kappa_i)\lambda_i} \qquad (5)$$

which means that agents do not have much previous experience. The term u_i represents the utility of player i, which is a function of his own pay off as well as pay offs to others. The details of this social utility function are explained below.

Social preferences

The fact that many players in public-good games do contribute to the provision of a public good at a substantial level, even in the final rounds, indicates that their preferences are not entirely dictated by the monetary pay offs they receive in the experiments. Thus, allowing for social preferences is crucial in explaining the dynamics of these games. In addition, the extent to which agents deviate from purely selfish motivation differs from one agent to the next. There are multiple ways of representing these heterogeneity preferences (Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Charness and Rabin 2002, Cox and Friedman 2002, for example).

The utility functions are modified to reflect the specifics of the repeated N-person public-goods provision experiments. That is, instead of the exact distribution of the pay offs to others, an agent is assumed to consider the average of the pay offs of others: $\overline{\pi}_{-i}$. We use the average because, in the experiments that generated the data being used, the players did not have information about the exact distribution of pay offs to other group members; they could only infer the average pay off to others.

Charness and Rabin (2002) developed a general model for social preferences that embeds other models. The utility function is defined as

$$U_i = \rho \cdot \overline{\pi}_{-i} + (1-\rho) \cdot \pi_i, \text{ if } \pi_i \geq \overline{\pi}_{-i} \text{ and}$$
$$U_i = \chi \cdot \overline{\pi}_{-i} + (1-\chi) \cdot \pi_i, \text{ if } \pi_i < \overline{\pi}_{-i} \text{ and} \qquad (6)$$

where $\chi \leq \rho \leq 1$. A lower value of χ compared to ρ implies that a player gives a larger weight to his own pay off when his pay off is smaller than the average pay off of others than when it is larger. When $\chi \leq \rho \leq 0$ the player is highly competitive. The players like to have their pay offs higher than those of the other players. An alternative model is that players prefer the pay offs among the players to be equal. This so-called inequity aversion holds when $\chi < 0 < \rho < 1$ (see Fehr and Schmidt 1999). The third model is the so-called social welfare consideration, which holds when $0 < \chi \leq \rho \leq 1$. The parameter ρ captures the extent to which a player weighs the average pay offs of the other N-1 agents compared to his own pay off when his own pay off is higher than the average payoff of the others. If $\rho = \chi = 0$, we have the condition that a player cares only about his own welfare.

Signaling

Another component in the utility function has to do with the forward-looking signaling behavior of the players in repeated games. Isaac et al. (1994) propose the hypothesis that these players are involved in a forward-looking intertemporal decision problem. Players may signal their willingness to contribute for a public good in the future by contributing at a high level in the current round. A player may benefit from this signaling if others respond positively in the following rounds. If this is the case, the potential benefit of signaling depends on the number of rounds that are left before the game ends. Therefore, one would expect less signaling toward the end of a game. This is consistent with their findings in experiments with more than 10 rounds (Figs. 1 and 2). That is, the decline of contribution level depends not so much on the number of rounds played as it does on the number of rounds remaining.

We assume that the attraction of strategy x_i as formulated in Eq. 2 is adapted to include the signaling component in the following way

$$A_i^x(t) = \frac{\lambda_i H(t-1)A_i^x(t-1) + [\delta_i + (1-\delta_i)I(x_i, x_i(t))](u_i(x_i, x_{-i}(t)) + x_i \cdot r \cdot \theta \cdot \left(\frac{T-t}{T}\right)^{\eta})}{H(t)} \qquad (7)$$

The added component indicates that a player thinks that his contribution level in the current round, *x,* positively affects others contribution in the future. In addition, the larger the marginal per capita return (*MPCR*) is, the more positive a player's assessment of the effect of his own contribution on the future contributions of others. The two individualized parameters, θ_i and η_i, also affect the signaling strength of *i,* generating another dimension of heterogeneity across agents. Specifically, θ_i represents player *i*'s belief about how large the positive effect of his contribution will be on the future contributions of others. The parameter η_i models player *i*'s end behavior, given that $(T - t)/T$ is smaller than 1, a larger (smaller) η_i.

MODEL ESTIMATION

For the eight treatments shown in Table 1, which contains 278 players, we have estimated the parameters listed in Table 2. Three types of estimations were conducted: (1) representative agent estimation, (2) multiple-type estimation, and (3) individual estimation. In the representative agent estimation, we assume that all the players are of the same type and estimate the parameters of the model. In the multiple-type estimation, we use the methodology of El-Gamal and Grether (1995) that divides the players into multiple segments to find the best fit. In the individual-level estimation, the parameters are estimated for each individual player.

Because of the stochastic nature of the model, we use conventional maximum likelihood (*L*) estimation to estimate the parameters. Fitting the model, however, is not an adequate approach for evaluating model performance (Pitt and Myung 2002). The main problem is that more complicated models have more degrees of freedom to fit the data. The trade-off is between the fit of the data and the complexity of the model. We use two criteria to evaluate the different estimated model versions.

The first criterion is the Akaike Information Criterion (AIC), which is defined as

$$AIC = -2\ln L + 2k \tag{8}$$

where *k* is the number of parameters of the model. Thus, for each parameter added to the model, the maximum likelihood needs to increase more than one unit to justify this extra parameter. The Bayesian

Information Criterion (BIC) also includes the number of observations *N* used in the equation:

$$BIC = -2\ln L = k\ln(N). \tag{9}$$

This means that, the more observations are used, the more an extra parameter must contribute to improving the maximum likelihood to justify this extra parameter. For example, when *N* is 8, the improvement of the maximum likelihood must be slightly more than one unit, but, when *N* is 80, the improvement must be more than 2.2 units. Both AIC and BIC are ways to strike a balance between the fitness and complexity of models and favor the models with lower AIC/BIC values.

Representative agent estimation

Here, we estimated four variants of the general model. In each of the estimated models, agents are assumed to be homogeneous, i.e., they have the same set of parameters. The four models include different elements of the general model denoted "SP" (social preference according to the Charness-Rabin social welfare utility function), "L" (experience-weighted attraction learning model of Camerer and Ho), and "S" (signaling). They are listed below:

1. Model SP: Probabilistic choice with social preferences, without learning and signaling.

2. Model SP+L: Probabilistic choice with social preferences with learning and without signaling.

3. Model SP+L+S: Probabilistic choice with social preferences with learning and with signaling.

4. Model L: Probabilistic choice of income maximizers with learning and without signaling.

Tables 3 and 4 show the results of the maximum likelihood estimation. In both the 10- and 40-round data, the most comprehensive model (SP+L+S) gives the lowest AIC and BIC, and the simple SP

Table 1. Experimental conditions and parameters. The parameters are defined in Table 2. IW88 signifies that the data are from Isaac and Walker (1988); IWW94 signifies that the data are from Isaac et al. (1994). In all of these eight experimental conditions, the participating players were experienced in the sense that they had previously participated in another public-good game. Also, cash payments were made in all of the eight conditions. The eight conditions are divided into two subsets: six experimental conditions with 10 rounds and two experimental conditions with 40 and 60 rounds.

Condition	Source	N	R	ω	α	No. of rounds	No. of sessions
1	IW88	4	0.3	62	1	10	6
2	IW88	4	0.75	25	1	10	6
3	IW88	10	0.3	25	1	10	6
4	IW88	10	0.75	10	1	10	6
5	IWW94	40	0.03	50	2	10	1
6	IWW94	40	0.3	50	0.5	10	1
7	IWW94	10	0.3	100	0.5	40	2
8	IWW94	10	0.3	100	0.5	60	1

model performs the worst. The differences in AIC and BIC values are relatively small between Model SP+L and Model SP+L+S and, thus, one might wonder whether the added complexity of SP+L+S is worth the trouble. However, recall that both the AIC and BIC values already account for the degree of complexity in evaluating performance of the models. Thus, we consider that the three features of social preference, learning, and signaling are all essential in explaining behavior in repeated public-good games.

In the 10-round data, the estimated parameters are quite similar between Models SP+L and SP+L+S (Table 3). The positive values of ρ and χ suggest that the players on average had a social-welfare utility function in which utility is a weighted average of one's own pay off and the average of the pay offs of others. Admittedly, this is somewhat different from the more widely accepted wisdom that the players in the social-dilemma games typically exhibit a conditionally cooperative behavior as suggested by Fehr and Schmidt (1999) or Bolton

and Ockenfels (2000). The utility function of Charness and Rabin (2003) that we used in this study embeds the inequality aversion as a special case. That is, if the estimation results were a positive ρ and a negative χ, that would be consistent with a preference for inequality aversion. It is possible that, if we used either Fehr and Schmidt's or Bolton and Ockenfel's utility functions, we could have found estimates that are consistent with an inequality aversion. However, because the main focus of our study is to test the significance of some broad features such as learning, signaling, and social preference, we did not conduct a separate estimation using an inequality aversion function. Instead, we limit the result as suggesting that some level of other-regarding preferences is present among the players, not necessarily that a social-welfare utility function is superior to an inequality-aversion utility function.

Also, notice that in Model SP without learning or signaling, the representative agent appears to be competitive, i.e., a difference maximizer, as

Table 2. Parameters of the model.

Parameter	Interpretation
ρ	Weight to others' payoff when $\pi_i > \overline{\pi}_{-i}$
χ	Weight to others' payoff when $\pi_i < \overline{\pi}_{-i}$
φ	Response sensitivity
λ	Forgetting rate
δ	Weight to forgone payoffs
κ	Rate of attraction growth
θ	Signal weight
η	Weight of future rounds in signaling

suggested by negative ρ and χ. However, because Model SP has a significantly poorer fit compared to Models SP+L and SP+L+S, and the estimates are quite similar between Models SP+L and SP+L+S, we consider the results of Model SP estimation to be invalid.

The discount of the past, parameter λ, is approximately 0.85 in both the 10- and 40/60-round data. The weights of forgone pay offs δ are 0.55 and 0.72, respectively, which suggests that the players are more belief learners than reinforcement learners. The rates of attraction growth are 0.06 and 0.03, which represent a rapid attraction to particular choices.

The estimated signaling parameters differ between the two data sets. The 10-round experiments lead to a short and strong effect of signaling, with θ equal to 2.05 and η equal to 10. However, the 40- and 60-round experiments lead to a weaker effect of signaling, although it does have an effect over a relative longer period than the 10-round experiments. This might indicate that the relative effect of signaling differs when the duration of the game changes.

Estimation and model evaluation with multiple types of agents

Now that we have estimated the representative agent, we will perform maximum likelihood estimation with different types of agents. Using the methodology of El-Gamal and Grether (1995), we maximize the likelihood function and at the same time classify different types of agents. Because the full SP+L+S model came out the strongest in our representative agent estimation, we used it in the estimation of multiple types. Because the model specification is identical, the only difference among the estimated models is the number of types allowed.

Table 3. Estimated parameter values for different model versions using the data from experiments with 10 rounds. The model versions are defined as follows: SP = probabilistic choice with social preferences, without learning and signaling; SP+L = probabilistic choice with social preferences with learning and without signaling; SP+L+S = probabilistic choice with social preferences with learning and with signaling; and L = probabilistic choice of income maximizers with learning and without signaling. See Table 2 for a list of parameters and their definitions. InL stands for maximum likelihood, AIC for Akaike Information Criterion, and BIC for Bayesian Information Criterion.

	SP	SP+L	SP+L+S	L
ϕ	0.033	0.260	0.276	0.231
ρ	-0.836	0.398	0.380	-(0)
χ	-0.837	0.395	0.374	-(0)
λ	-(1)	0.829	0.818	0.900
δ	-(1)	0.538	0.545	0.187
κ	-(1)	0.077	0.057	0.132
θ	-(0)	-(0)	2.05	-(0)
η	-(0)	-(0)	10	-(0)
lnL	-7038.2	-5505.7	-5495.1	-5933.1
AIC	14082.5	11023.4	**11006.2**	11874.2
BIC	14099.9	11058.3	**11052.7**	11897.4

Once the number of types is exogenously given in an estimation, the maximum likelihood estimation endogenously distributes the 248 players into different categories until the likelihood is maximized. Starting from the two-types model, we increased the number of types until the model started to perform more poorly than a model with a smaller number of types. Here the focus is on whether allowing for multiple types improves the fit. Thus, the substantive details of the estimation results are suppressed. For comparison purposes, the AIC and BIC values of the representative agent model estimation and the individual estimation, i.e., 248-types model, which will be discussed in the next subsection, are included in Tables 5 and 6.

We find that eight different types of agents best explain the data on the 248 players in the 10-round

experiments when we take into account the increasing complexity of the model with a larger number of parameters. We also find that two types of agents provide the best explanation for the 30 players in the 40/60-round experiments.

Table 5 and 6 show how the indicators of goodness of fit and the generalization indicators are affected by the number of agent types. Table 5 shows that, up to eight different types of agents, the performance of the model improves. From Table 6 it can be seen that two distinct types of agents improve the performance of the model, whereas it performs less well when we add more types of agents, i.e., an increase in BIC. In both 10- and 40/60-round data sets, the best multiple-types models (8-types in 10 rounds and 2-types in 40/60 rounds) perform much better than either the representative agent model or

Table 4. Estimated parameter values for different model versions using the data of experiments with 40 rounds and 60 rounds. The model versions are defined as follows: SP = probabilistic choice with social preferences, without learning and signaling; SP+L = probabilistic choice with social preferences with learning and without signaling; SP+L+S = probabilistic choice with social preferences with learning and with signaling; and L = probabilistic choice of income maximizers with learning and without signaling. See Table 2 for a list of parameters and their definitions. InL stands for maximum likelihood, AIC for Akaike Information Criterion, and BIC for Bayesian Information Criterion.

	SP	SP+L	SP+L+S	L
φ	0.0069	0.485	0.440	0.260
ρ	0.639	0.909	0.607	-
χ	-0.987	0.898	0.587	-
λ	-(1)	0.888	0.887	0.875
δ	-(1)	0.773	0.722	0.082
κ	-(1)	0.021	0.027	0.005
θ	-(0)	-(0)	0.472	-
η	-(0)	-(0)	0.473	-
lnL	-6062.6	-4754.6	-4730.4	-4970.8
AIC	12131.2	9521.2	**9476.8**	9949.6
BIC	12146.9	9552.6	**9518.7**	9970.6

the fully heterogeneous model. The optimal number of types is rather large, probably because of the complexity of the general model. Again, however, given that the AIC and BIC scores take into account model complexity, including the number of types, we cautiously conclude that it is essential to incorporate multiple types of agents defined on multiple and continuous dimensions of heterogeneity to understand the results from repeated experiments involving the provision of public goods.

Individual level estimation

Finally, we estimated the parameters for each player. This leads to a distribution of parameter values. Figure 6 provides the cumulative distribution of the estimated parameter values, which gives an indication of distributions. For most parameters, these distributions are remarkably similar among the 10- and 40/60-round data sets. Besides the distributions of the estimated parameters of the two data sets, we defined general distribution functions (Table 7) that mimic the observed distributions. This is the third line, i.e., the one with triangle legends, in each of the parameter figures in Fig. 6. We did not do a formal estimation of the general distributions in Table 7, but defined some simple forms that mimic the general features so that we could use this in simulation models as a general model that represents the statistics. Note that one of our aims is to derive agent-based models based on empirical data, and therefore a more general description is preferred. The generalized distributions might provide educated information for other agent-based models when heterogeneity of agents is assumed.

Table 5. The value of maximum likelihood (InL), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) for different numbers of types distinguished by maximum likelihood estimation using the method of El-Gamal and Grether (1995) for 10-round data.

| | Number of types | | | | | | | | | |
	1	2	3	4	5	6	7	8	9	248
lnL	-5495.1	-5191.5	-5055.3	-4994.9	-4926.5	-4885.2	-4827.1	-4780.5	-4775.4	-4028.8
AIC	11006.2	10415.0	10158.7	10053.8	9933.0	9866.4	9766.1	**9688.9**	9694.7	12025.7
BIC	11052.7	10508.0	10298.2	10239.9	10165.6	10145.6	10091.8	**10061.1**	10113.5	23564.6

Based on the derived parameter values of the individual players, we can perform an analysis of the characteristics of the various players. For each estimated agent, we determined what kind of utility model is most appropriate, and what kind of learning model is implied from the estimated parameter values. Table 8 shows the classified agents. Note that 16 possible types are presented; these represent the possible combinations of the four learning types and the four preference types. Also note that some of the types contain only a few players. Most of the players belong to the two upper rows, which correspond to either an inequity-aversion preference or a social-welfare preference with various learning types. Recall that in our estimation of multiple-type models, the model with eight types performed the best in the 10-round data. The classification of individuals based on the individual-level estimation is quite consistent with the multiple-type estimation result. Consequently, eight types in Table 8 contain 226 out of 248 players.

In terms of style of learning, most of the players are identified as belief learners, including Cournot-type learners, who take into account not only their experienced pay offs but also the pay offs the agents could have gotten had they made other decisions. Given the large number of decision options, i.e., 11 to 101 possible token investment options, the fact that most players are identified as belief learners is not a surprise because learning from only experienced observations, i.e., reinforcement learning, would take much longer. Also interesting is the fact that most of the players identified as

reinforcement learners have short memory spans as indicated by large λ parameters. This seems to suggest that they are not, in fact, learning systematically from their past experiences. With regard to social preferences, the inequity-aversion preference is the most frequently identified utility function. Note that 216 out of 248 players are identified as having either inequity-aversion or social-welfare preferences, again suggesting that incorporating social preference is essential in understanding the results of repeated social-dilemma experiments. Fewer than 10% of the agents are identified as interested only in maximizing their own pay offs.

In Appendix 1 we provide a more in-depth analysis of the models generated by the three different estimation techniques. In particular, macro-level statistics generated by the models are compared with the same statistics obtained from the data. Some of the macro-level statistics, such as those from Fig. 5, are not produced with great accuracy by the simulation models.

DISCUSSION

In this paper we evaluated versions of a hybrid model of decision making and learning in repeated public-good experiments. Our analyses show that most players have other-regarding preferences, and that the types of other-regarding preferences differ among the players. The players learn in different ways from their experience, but the most dominant

Table 6. The value of maximum likelihood (InL), Akaike Information Criterion (AIC), and Bayersian Information Criterion (BIC) for different numbers of types distinguished by maximum likelihood estimation using the method of El-Gamal and Grether (1995) for 40/60-round data.

| | Number of types | | | | | |
	1	2	3	4	5	30
lnL	-4730.4	-4621.0	-4602.3	-4581.8	-4571.9	-4367.7
AIC	9476.8	9274.0	9252.6	9227.7	9223.7	9215.4
BIC	9518.7	9358.0	9378.4	9395.5	9433.7	10474.0

result from our analysis is a belief learning process in which players take into account the potential benefit they could have derived if they had made different choices. Some players signal their willingness to invest in public goods in the hope that others will increase their investments too.

In sum, even in the baseline public-good experiments without additional institutional features such as punishment (Ostrom et al. 1992, Fehr and Gächter 2000, Anderson and Putterman 2006) or endogenous group formation (Coricelli et al. 2003, Ahn et al. 2005, Cinyabuguma et al. 2005), it is essential that the dynamics at the individual and group levels be explained as interactions among multiple types of players defined on multiple dimensions of heterogeneity. In this sense, as Ones and Putterman (2004) suggest, repeated *N*-person dilemmas need to be studied from the viewpoint of an ecology of interacting types. Consistent with experimental studies that specifically address the problem of heterogeneous preference types in repeated public-good games (Fischbacher et al. 2001, Kurzban and Houser 2001, Fischbacher and Gächter 2006), we find that most of the subjects have other-regarding preferences of inequality aversion or conditionally cooperative preferences.

In addition, our simulation results suggest that most subjects, although they do have other-regarding preferences, are at the same time quite rational. They seem to form and update their beliefs about the

behavior of others and then choose their actions based on their beliefs and preferences. This finding may explain why punishment opportunity might encourage contribution even before punishment is exercised (Fehr and Gächter 2000) and why certain forms of endogenous group formation, especially expulsion, induce very high levels of contribution from the very beginning of an experiment (Cinyabuguma et al. 2005). Our results also suggest that the rationality of some, if not a majority of, subjects extends to signaling their intentions in an attempt to induce higher levels of contribution from others. An interesting venue for future research would be to derive the implications of the types identified in our study to richer institutional settings and test whether the results of such experiments can also be systematically explained in terms of the interaction of the types.

Methodologically, this paper is an attempt to expand the horizon of empirically grounded agent-based modeling practices. Our analysis combines rigorous tools from behavioral economics and cognitive science (maximum likelihood estimation) with agent-based models (emergent properties and macro-level metrics). For the empirical testing of agent-based models in laboratory experiments involving group dynamics, we derive a good starting point from statistical tools like maximum likelihood. Nevertheless, it is not sufficient to generate all the emerging properties from agent interactions. A problem with maximum likelihood

Fig. 6. Cumulated distribution of the estimated parameters for the two data sets. The parameters are defined in Table 2.
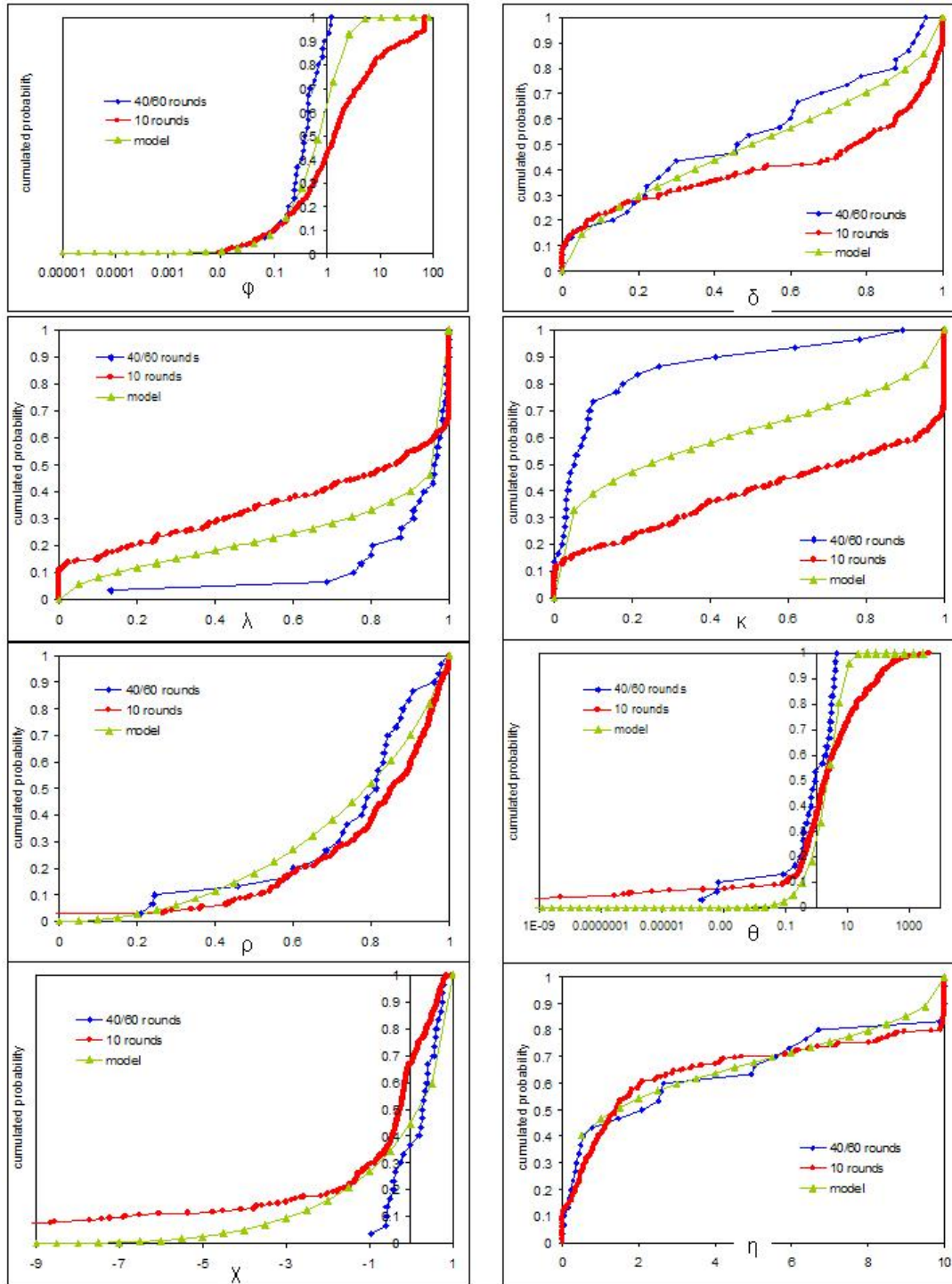
Table 7. Simple distribution functions of the parameters of the model.

Parameter	Distribution function
ρ	Beta(2,0.75)
χ	-9+10*Beta(3,0.5)
φ	Exponential(1.0)
λ	Beta(0.5,0.15)
δ	Beta(0.5,0.5)
κ	Beta(0.25,0.4)
θ	Exponential(0.3)
η	10*Beta(0.2,0.4)

estimation is the focus on the calibration of observations at the individual level. However, emergent patterns at the group level, such as the patterns in Fig. 5, are not necessarily generated when the model is calibrated at the individual level. Hence, agent-based models require methods for multilevel calibration and evaluation. The balance between fitting the data and generalizability remains another problem. Although we can include some penalties within the maximum likelihood estimation, such as the number of parameters, it is not clear whether this penalizes model complexity for agent-based models. For example, computational time might also be a consideration to be included in the penalty.

Despite the problems of model estimation and evaluation, we were able to develop a general model that mimics the most important elements of the experimental data. We found that other-regarding preferences, learning, and signaling all had to be included to explain the observations. Adding all these components was still beneficial after including penalties for model complexity. Assuming that there is agent heterogeneity improves the maximum likelihood estimation; this also occurs when additional complexity is penalized. Therefore, a representative agent model for public-good experiments is not justified based on our findings. We were able to derive parameter distributions based on the individual-level calibration of the experimental data. These parameter distributions can be used to inform applied agent-based models in which social dilemmas are involved.

Based on the distributions of parameter values, we found that players seem to use different learning models, namely belief learning and reinforcement learning, and other-regarding preferences, e.g., inequity aversion and social welfare. The largest group, about 25%, is classified as inequity-aversion players with reinforcement learning and forgetting. Only 10% of the players are classified as selfish.

The results of our model analysis depended on the specific functional forms we used. Although we based our model on hybrid model versions of experimental economics studies, we also considered some of the additional functional forms used in the literature. Nevertheless, our results showed the potential of using laboratory experiments to develop empirically tested agent-based models. Most notably, to explain the dynamics of social dilemmas, we had to incorporate multiple types of agents or distributions of parameter values. We have shown that we can detect this agent heterogeneity by various methods in analyzing the data. These empirically tested agent-based models might guide the parameterization of applied agent-based models.

Table 8. Classification of players based on the estimated parameter values. The parameters are defined in Table 2. The 248 players in the 10-round experiments are classified into 16 possible cases. The total number for each type of learning model is given in the last row, and the total number for each type of utility function is given in the last column. The 30 players of the 40/60-round experiments are also classified, and these numbers are given in parentheses.

	Cournot ($\lambda < 0.5$; $\delta > 0.5$)	Belief learning ($\lambda > 0.5$; $\delta > 0.5$)	Reinforcement learning ($\lambda < 0.5$; $\delta < 0.5$)	Reinforcement learning with forgetting ($\lambda > 0.5$; $\delta < 0.5$)	Total (sum of columns)
Inequity aversion ($\rho > 0.1$; $\chi < -0.1$)	32	45	11	60 (10)	143 (10)
Social welfare ($\rho > 0.1$; $\chi > 0.1$)	27 (1)	32 (12)	3	11 (6)	73 (19)
Competitive ($\rho < -0.1$; $\chi < -0.1$)	1	2	1	3	7 (0)
Selfish ($-0.1 < \rho < 0.1$; $-0.1 \chi < 0.1$)	6	5 (1)	1	8	20 (1)
Total (sum of rows)	66 (1)	84 (13)	16 (0)	82 (16)	248 (30)

LITERATURE CITED

Ahn, T. K., M. Isaac, and T. Salmon. 2005. *Endogenous group formation.* Florida State University, Gainesville, Florida, USA.

Ahn, T. K., E. Ostrom, and J. M. Walker. 2003. Heterogeneous preferences and collective action. *Public Choice* **117**:295-314.

Anderson, C. M., and L. Putterman. 2006. Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior* **54**:1-24.

Battalio, R., L. Samuelson, and J. Van Huyck. 2001. Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica* **69** (3):749-764.

Binmore, K., and L. Samuelson. 1992. Evolutionary stability in repeated games played by finite automata. *Journal of Economic Theory* **57**:278-305.

Binmore, K., and L. Samuelson. 1997. Muddling through: noisy equilibrium selection. *Journal of Economic Theory* **74**:235-265.

Bolton, G. E., and A. Ockenfels. 2000. ERC: a theory of equity, reciprocity and competition. *American Economic Review* **90**:166-193.

Boylan, R. T., and M. El-Gamal. 1993. A fictitious play: a statistical study of multiple economic experiments. *Games and Economic Behavior* **5**:205-222.

Bracht, J., and H. Ichimura. 2002. *Identification of a general learning model on experimental game data.* Hebrew University of Jerusalem, Jerusalem, Israel.

Broseta, B. 2000. Adaptive learning and equilibrium in experimental coordination games; an ARCH(1) approach. *Games and Economic Behavior* **32**(1):25-30.

Bush, R., and F. Mosteller. 1955. *Stochastic models of learning.* Wiley, New York, New York, USA.

Camerer, C. F. 2003. *Behavioral game theory: experiments in strategic interaction.* Princeton University Press, Princeton, New Jersey, USA.

Camerer, C. F., and C. M. Anderson. 2000. Experience-weighted attraction learning in sender-receiver signaling games. *Economic Theory* **16**:689-718.

Camerer, C. F., and T. H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* **67**(4):827-874.

Charness, G., and M. Rabin. 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* **117**(3):817-869.

Chen, Y., and F.-F. Tang. 1998. Learning and incentive-compatible mechanisms for public goods provision: an experimental study. *Journal of Political Economy* **106**:633-662.

Cheung, Y.-W., and D. Friedman. 1997. Individual learning in normal form games: some laboratory results. *Games and Economic Behavior* **19**:46-76.

Cheung, Y.-W., and D. Friedman. 1998. A comparison of learning and replicator dynamics using experimental data. *Journal of Economic Behavior and Organization* **35**:263-280.

Cinyabuguma, M., T. Page, and L. Putterman. 2005. Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics* **89**(8):1421-1435.

Coricelli, G., D. Fehr., and G. Fellner. 2003. *Partner selection in public goods experiments.* Discussion Paper on Strategic Interaction 2003-13. Max Planck Institute of Economics, Jena, Germany.

Cox, J. C., and D. Friedman. 2002. *A tractable model of reciprocity and fairness.* University of Arizona, Tempe, Arizona, USA.

Crawford, V. 1995. Adaptive dynamics in coordination games. *Econometrica* **63**:103-144.

Crawford, V., and B. Broseta. 1998. What price coordination? Auctioning the right to play as a form of preplay communication. *American Economic Review* **88**:198-225.

Cross, J. G. 1983. *A theory of adaptive economic behavior.* Cambridge University Press, New York, New York, USA.

El-Gamal, M. A., and D. M. Grether. 1995. Are people Bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association* **90** (432):1137-1145.

Ellison, G., and D. Fudenberg. 1993. Rules of thumb for social learning. *Journal of Political Economy* **101**:612-643.

Erev, I., and A. E. Roth. 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review* **88**

(4):848-881.

Fehr, E. and S. Gächter. 2000. Cooperation and punishment. *American Economic Review* **90**:980-994.

Fehr, E., and K. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* **114**:817-868.

Feltovich, N. 2000. Reinforcement-based vs. belief-based learning models in experimental asymmetric-information games. *Econometrica* **68**:605-641.

Fischbacher, U., and S. Gächter. 2006. *Heterogeneous social preferences and the dynamics of free riding in public goods.* Institute for Empirical Research in Economics, Working Paper 261. University of Zurich, Zurich, Switzerland.

Fischbacher, U., S. Gächter, and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public good experiment. *Economic Letters* **71**:397-404.

Fudenberg, D., and D. M. Kreps. 1993. Learning mixed equilibria. *Games and Economic Behavior* **5**:320-367.

Fudenberg, D., and D. K. Levine. 1995. Consistency and cautious fictitious play. *Journal of Economic Dynamics and Control* **19**:1065-1090.

Fudenberg, D., and E. Maskin. 1990. Evolution and cooperation in noisy repeated games. *American Economic Review* **80**:274-279.

Haruvy, E., and D. Stahl. 2004. Deductive versus inductive equilibrium selection: experimental results. *Journal of Economic Behavior and Organization* **53**:319-331.

Ho, T.-H., and K. Weigelt. 1996. Task complexity, equilibrium selection, and learning: an experimental study. *Management Science* **42**(5):659-679.

Isaac, R. M., K. F. McCue, and C. R. Plott. 1985. Public goods provision in an experimental environment. *Journal of Public Economics* **26**:51-74.

Isaac, R. M., and J. M. Walker. 1988. Group size effects in public goods provision: the voluntary contribution mechanism. *Quarterly Journal of Economics* **103**:179-200.

Isaac, R. M., J. M. Walker, and S. H. Thomas. 1984. Divergent evidence on free riding: an experimental examination of possible explanations. *Public Choice* **43**:113-149.

Isaac, R. M., J. M. Walker, and A. W. Williams. 1994. Group size and the voluntary provision of public goods: experimental evidence utilizing large groups. *Journal of Public Economics* **54**(1):1-36.

Kaniovski, Y. M., and P. Young. 1995. Learning dynamics in games with stochastic perturbations. *Games and Economic Behavior* **11**:330-363.

Kurzban, R., and D. Houser. 2001. Individual differences in cooperation in a circular public goods game. *European Journal of Personality* **15(S1)**:S37-S52.

Ledyard, J. 1995. Public goods: a survey of experimental research. Pages 111-1194 *in* J. Kagel and A. Roth, editors. *The handbook of experimental economics.* Princeton University Press, Princeton, New Jersey, USA.

Marwell, G., and R. E. Ames. 1979. Experiments on the provision of public goods. I. Resources, interest, group size, and the free rider problem. *American Journal of Sociology* **84**:1335-1360.

Marwell, G., and R. E. Ames. 1980. Experiments on the provision of public goods. II. Provision points, stakes, experience and the free rider problem. *American Journal of Sociology* **85**:926-937.

Marwell, G., and R. E. Ames. 1981. Economists ride free, does anyone else? *Journal of Public Economics* **15**:295-310.

Mookherjee, D., and B. Sopher. 1994. Learning behavior in an experimental matching pennies game. *Games and Economic Behavior* **7**:62-91.

Mookherjee, D., and B. Sopher. 1997. Learning and decision costs in experimental constant-sum games. *Games and Economic Behavior* **19**:97-132.

Nyarko, Y., and A. Schotter. 2002. An experimental study of belief learning using elicited beliefs. *Econometrica* **70**:971-1005.

Ones, U., and L. Putterman. 2004. *The ecology of*

collective action: a public goods and sanctions experiment with controlled group formation. Department of Economics Working Paper 2004-01. Brown University, Providence, Rhode Island, USA.

Ostrom, E., J. Walker, and R Gardner. 1992. Covenants with and without a sword: self-governance is possible. *American Political Science Review* **86**:404-417.

Pitt, M. A., and I. J. Myung. 2002. When a good fit can be bad. *Trends in Cognitive Sciences* **6** (10):421-425.

Robinson, J. 1951. An iterative method of solving a game. *Annals of Mathematics* **54**:296-301.

Salmon, T. 2001. An evaluation of econometric models of adaptive learning. *Econometrica* **69** (6):1597-1628.

Sarin, J., and F. Vahid. 1999. Payoff assessments without probabilities: a simple dynamic model of choice. *Games and Economic Behavior* **28**:294-309.

Sarin, J., and F. Vahid. 2001. Predicting how people play games: a simple dynamic model of choice. *Games and Economic Behavior* **34**:104-122.

Schlag, K. H. 1998. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory* **78**:130-156.

Stahl, D. O., and E. Haruvy. 2002. Aspiration-based and reciprocity-based rules in learning dynamics for symmetric normal-form games. *Journal of Mathematical Psychology* **46**(5):531-553.

Tang, F.-F. 2001. Anticipatory learning in two-person games: some experimental results. *Journal of Economic Behavior and Organization* **44** (2):221-232.

Wilcox, N. 2006. Theories of learning in games and heterogeneity bias. *Econometrica* **74**(5):1271-1292.

Young, P. 1993. The evolution of conventions. *Econometrica* **61**:57-84.

Appendix 1. More model analysis

Please click here to download file 'appendix1.pdf'.