ELSEVIER

Journal of Economic Behavior & Organization 67 (2008) 782–793

JOURNAL OF
Economic Behavior & Organization

www.elsevier.com/locate/econbase

Strategic behavior and learning in repeated voluntary contribution experiments

Laurent Muller[a], Martin Sefton[b,*], Richard Steinberg[c], Lise Vesterlund[d]

[a] INRA-GAEL, University of Grenoble, BP 47, 38040 Grenoble Cedex 9, France
[b] University of Nottingham, School of Economics, University Park, Nottingham NG7 2RD, UK
[c] Indiana University-Purdue University at Indianapolis, Department of Economics, Indianapolis, IN 46202, USA
[d] University of Pittsburgh, Department of Economics, Pittsburgh, PA 15260, USA

Received 3 July 2006; received in revised form 10 September 2007; accepted 11 September 2007
Available online 14 September 2007

Abstract

Voluntary contribution experiments systematically find that contributions decline over time. We use a two-stage voluntary contribution game to investigate whether this decrease is caused by learning or strategic behavior. Using a strategy method we find a robust pattern of declining contributions: contributions in stage 2 are 45 percent lower than in stage 1. Repeating the game five times we find that experience generates a smaller decline in contributions: stage 1 contributions decrease by around 7 percent per game. Finally we find no significant differences between the strategy and direct-response method, which suggests that our results help explain behavior in the latter.

JEL classification: C92; C72; H41

Keywords: Voluntary contributions; Cooperation; Experiment

1. Introduction

A robust finding in experimental studies of voluntary contribution games is that cooperation decreases with repeated interaction. Two explanations have been provided for this result. One is that strategies initially are well-defined, but that they depend on the history of play and therefore may cause players to change their actions over the course of the game. Another is that participants slowly begin to understand the game and refine their strategies accordingly. While both of these explanations are convincing, limited work has been done to determine directly the role played by each. The purpose of this study is to provide such an examination.

To determine the extent to which decreasing contributions arise as a result of history-dependent strategies we need to identify the participants' strategies in a multi-stage voluntary contribution game. To do so we apply the strategy method (Selten, 1967) to a two-stage voluntary contribution game. Participants submit a plan of action for the two-stage game, specifying a stage 1 contribution and a stage 2 response to any possible number of aggregate stage 1

* Corresponding author. Tel.: +44 115 846 6130; fax: +44 115 951 4159.
E-mail addresses: laurent.muller@grenoble.inra.fr (L. Muller), martin.sefton@nottingham.ac.uk (M. Sefton), rsteinbe@iupui.edu (R. Steinberg), vester@pitt.edu (L. Vesterlund).

0167-2681/$ – see front matter
doi:10.1016/j.jebo.2007.09.001

contributions by other group members. The plan is then matched with that of two other participants, and actions are taken in accordance with the participant's strategy. An advantage of the strategy method is that participants make a decision for every possible history of aggregate contributions, not just those reached in the course of actual play. Thus participants indicate how they would respond to different contributions, making it straightforward to identify strategies. We find that the pattern of contributions is comparable to the results from previous repeated voluntary contribution experiments in that the elicited strategies imply declining average contributions. The elicited strategies cause contributions to decline by an average of 45 percent within the two-stage games.

To examine the potential role played by learning we have participants play a sequence of 5 two-stage games, playing against a new set of people in each game. This allows us to examine whether experience with the game leads to a modification of strategies and whether such modifications decrease contributions. We find that individuals do modify their strategies across games, and as a result contributions in stage 1 sometimes increase and sometimes decrease across successive games. On average stage 1 contributions decrease by 7 percent per game. Stage 2 contributions decrease more reliably across games, but the average decrease is still only 14 percent per game. Thus experience with the game leads to an erratic and less pronounced deterioration in contributions, compared with the systematic and substantial deterioration generated by submitted strategies.

We observe, as in previous studies, considerable heterogeneity in submitted strategies. Following Fischbacher et al. (2001), we classify participants on the basis of the strategies they submit, and in any game most participants are classified as "free riders" (participants who contribute nothing in stage 2) or "conditional cooperators" (contributors whose stage 2 contribution increases with other group members' stage 1 contribution). Individual strategies are not stable across games, as many participants switch from one class to another across successive games. While a potential interpretation is that these participants do not have stable preferences, we cannot rule out that participants nonetheless have well-defined preferences over the outcome of the two-stage game. The reason is that experience may help participants better understand the game and may cause them to update their prior on the strategies others use when playing the game. Thus participants may be uncertain of what strategies best serve their interests and revise strategies as they learn how their own strategies influence outcomes. While participants often change their strategies from game to game, the *distribution* of strategies is quite stable: the proportions of different types do not change much from game to game.

A potential disadvantage of the strategy method is that it may elicit strategies that differ from those used by participants in a traditional experiment based on a 'direct-response' method. The direct-response method, applied to this game, would have participants make simultaneous stage 1 contributions, and then, after observing aggregate stage 1 contributions, participants would make simultaneous stage 2 contributions. For our strategy-method examination to shed light on the decrease in contributions observed in direct-response games it is necessary that the elicited behavior is similar across the two elicitations. Although the two methods are equivalent from a standard theoretical point of view, there are plausible behavioral reasons why differences may emerge (see the discussions in Roth, 1995 or Brandts and Charness, 2000). For example, eliciting a strategy for the two-stage game encourages subjects to consider stage 2 when making their stage 1 decision; if the two stages were played out in sequence a subject might make a stage 1 decision without giving any consideration to stage 2. If behavior under the direct-response method differs from that under the strategy method, then it is unreasonable to argue that we can use elicited strategies to shed light on the decrease in contributions observed in the direct-response interaction.

Previous research delivers mixed evidence on whether people behave differently according to whether the strategy or direct-response method is used. Brandts and Charness (2000) examined Sequential Prisoners' Dilemma and Sequential Chicken games and found no significant differences between outcomes elicited by direct-response and strategy-method approaches. They carefully conclude (p. 234): "in games of low complexity, the strategy method may be a valid technique for collecting a rich data set without affecting participants' decisions significantly." However, other studies provide evidence of significant differences between direct-response and strategy-method approaches. For example, Güth et al. (2001) compare alternative versions of a mini-ultimatum game and find significant differences when the direct-response method is used, but not when a strategy method is used, Brandts and Charness (2003) study a game with pre-play communication in which players can punish opponents who send misleading messages and find that under a direct-response method punishment rates are significantly higher than under a strategy method, and Burton et al. (2005) study coordination games with pre-play communication, finding that under a direct-response method coordination on the efficient equilibrium was significantly lower than under a strategy method.

Thus, since equivalence between the two elicitation procedures cannot be taken for granted, to test the validity of the strategy-method implementation of our game we also run a direct-response version of the game where participants simultaneously choose contributions for the first stage, are informed of stage 1 contributions, and then play the second stage. Our results reveal that the contributions elicited by the strategy method are no different from those of our control, direct-response treatment. Thus the behavior that results from the elicited strategies is consistent with that of the direct-response games.

The remainder of the paper is organized as follows. In Section 2 we review the related literature. We then describe our experimental design and procedures in Section 3. The results are presented in Section 4, and we conclude in Section 5.

2. Literature review

A number of studies have examined contributions in a repeated voluntary contribution game. Among the first is Isaac and Walker (1988) who, in a series of repeated linear public goods games, find that repetition causes free riding to increase and contributions to 'decay'. With 10 repetitions of the game contributions tend to start out at about half the endowment and decrease to about 15–25 percent of the endowment in round 10. The two most common arguments for this decrease in contributions are that as the game is repeated, errors diminish as participants learn to free ride and that with repetitions, participants can play history-dependent strategies, potentially causing both conditional cooperators and free riders to decrease their contributions over the course of the game.

A number of studies have used indirect tests to determine whether the 'learning' or the 'strategies' hypothesis is the likely cause of the decay in contributions. To examine whether participants at the end of the repeated game have learned to free ride, Isaac and Walker invited past participants back to the laboratory to determine if their contribution rates differed from that of inexperienced participants. In contrast to the learning hypothesis and in support of the strategies hypothesis, behavior of experienced and inexperienced participants is practically the same. Initial contributions are high and then decrease as the game is repeated.

Andreoni (1988) uses a different approach to examine the two hypotheses. He argues that if the decay in cooperation is caused by state-contingent strategies then contributions should be larger when participants play in a finitely repeated game (partners) than when they are randomly paired with new participants after every repetition of the game (strangers). To test the learning hypothesis further he also examines the effect of restarting the game after 10 periods. If the decay in contributions is caused by participants learning to free ride, then contributions should be unaffected by the restart. His study provides evidence against both hypotheses. First, in contrast to the learning hypothesis, restarting the experiment causes a significant increase in contributions in the partner treatment. Second, in every round of the game contributions are larger in the stranger than partner treatment, suggesting that the decay in contributions cannot be caused by strategic play alone.[1]

Other researchers have subsequently used Andreoni's design to compare the behavior of partners versus strangers. While these have succeeded in replicating the restart effect, the evidence is mixed when replicating higher contributions in the stranger treatment.[2] Andreoni and Croson (in press) summarize these mixed findings and conclude, along the lines of Palfrey and Prisbrey (1996), that the reason why contributions in the stranger treatments sometimes are larger and sometimes smaller than in the partner treatment is that confusion or errors seem to be larger in the stranger treatment. In addition to errors differing between the two treatments, it has also been noted that both the decision-making process and the manner in which one learns across games can differ substantially between the two treatments. Thus it is not clear that by comparing contributions in the partners versus strangers treatments we can determine what causes the decay in contributions in the partner treatments.[3]

This paper uses a different approach to distinguish between the two hypotheses. First our learning hypothesis is broader, as we do not focus solely on learning as a process that teaches participants to free ride over time; rather we examine a repeated public good game multiple times, and we attribute any modification of strategies and resulting

[1] Andreoni and Croson (in press) argue that "...with plenty of experience in a number of finitely repeated games, subjects will learn the benefits of reputation building. In a single finitely repeated game, such as these public goods experiments, these results indicate that subjects are unlikely to have learned the sophisticated strategy of reputation building" (footnote 6).

[2] See for example Croson (1996).

[3] Along the lines of the learning hypothesis a series of other studies have argued that the decay in contributions in the strangers environment primarily is caused by a decrease in errors (see e.g. Andreoni, 1995; Palfrey and Prisbrey, 1996; Houser and Kurzban, 2002).

changes in contributions between supergames as being due to learning.[4] Second, in determining the role of strategic play we directly elicit the participants' strategies in the repeated public good game. Thus we do not try to infer the strategies by comparing behavior in two different treatments. While others previously have applied the strategy method to public good games, our approach is unique in that we elicit strategies over the entire repeated public good game and use these strategies to determine whether the decrease in contributions is due to learning or state-contingent strategies.

For example, Fischbacher et al. apply a strategy method to a sequential voluntary contributions game where three group members first make simultaneous contributions, and then after observing their total contribution, a fourth group member makes her contribution. In their experiment the fourth group member indicates a contribution for each possible average contribution by the other three. Thus they can determine whether behavior in sequential public goods games depends on the contributions made by others and classify the strategies that participants use. Fifty percent of their participants are classified as "conditional cooperators" (participants whose contribution in the role of fourth group member is increasing in others' contributions) and 30 percent are classified as free riders (participants whose contribution in the role of fourth group member always is zero). Fischbacher and Gächter (2006) extend this elicitation procedure to a two-part experiment. In one part they classify people in the sequential public good game, and in the second part participants play a 10-period repeated simultaneous-move public good game in a stranger environment with random rematching after every game. The distribution of types is similar to that of Fischbacher et al. (2001), and these type classifications are found to be a good predictor of behavior in the repeated public good experiment. They argue that preference heterogeneity in the sequential public good game can be used to understand better the dynamics of simultaneous-move voluntary contributions games in the strangers environment.

To the best of our knowledge Keser (2000) is the only study that elicits strategies for a repeated simultaneous-move public good game and allows participants to experience their submitted strategies. Using a rather special participant pool she asked 50 academic economists to submit strategies for playing a 25-period repetition of a voluntary contributions game. Strategies were then translated into computer code and matched with one another in every possible group of four, generating average payoffs for each strategy. Participants received feedback about payoffs and could then revise and resubmit strategies for a second, then a third simulation. She found that in all three simulations, strategies typically consisted of high initial contributions, followed by a phase of reciprocation, and then by a very pronounced endgame effect. With each simulation initial contributions increased, and endgame effects occurred earlier. The endgame effects she observes are stronger than in most standard experiments, perhaps because her participants are more sophisticated or more familiar with public goods experiments or because their motivations differed from those of typical experimental participants.[5]

As we focus specifically on the role played by learning and state-contingent strategies in decreasing contributions in the standard repeated simultaneous-move public good game, our approach differs from that of the earlier studies. Using a standard experimental subject pool we eliminate all possibilities of repeated interaction across games as participants are matched with another participant for one supergame only. Furthermore, to check that the elicited strategies match those of actual play of a repeated simultaneous-move game we conducted some control sessions with new subjects where subjects played an identical direct-response version of the game.

3. Experimental procedures

The experiment was conducted at the University of Nottingham in Spring 2001. We ran four sessions of a direct-response treatment and four sessions of a strategy-method treatment, with 15 participants in each session, for a total of 120 participants. Participants were recruited from a pool of undergraduate students at the university and randomly assigned to a treatment. No participant participated in more than one session of the experiment.

All sessions used an identical protocol. Upon arrival, participants were seated at computer terminals and given a set of instructions, which the experimenter then read aloud (Appendix A of the supplementary material). As part of

[4] Note that this can include learning in terms of developing an improved understanding of the game as well as learning in terms of updating priors on the anticipated strategies of opponents.

[5] Her participants were informed that after each simulation one participant would be randomly selected to be paid their earnings, and also that after the experiment participants' strategies, payoffs, and identities would be publicly revealed. In this latter respect her study is somewhat similar to the repeated prisoner's dilemma tournaments reported in Axelrod (1984).

the instructional phase of the session participants completed a computerized quiz that tested their understanding of the game. When all participants had completed the quiz correctly, the experimenter continued with the instructions.

The decision-making phase of the session then began with a practice two-stage game in which the participants played against the computer for hypothetical payoffs. The practice game was used simply to familiarize participants with their task and the computer screens. Participants then played 5 two-stage games in which they were randomly and anonymously matched into groups of three, no participant ever being matched with another participant more than once. Participants were endowed with two tokens in each stage of the game and had to decide how many tokens to place in a group account and how many to place in a private account. For each token an individual placed in the private account, that individual received £1.50, and for each token placed in the group account by any individual all three group members received £1. Denoting subject *j*'s contribution in stage *t* by x_{jt}, participant *i*'s payoff from a two-stage game is

$$\pi_i = 1.5(2 - x_{i1}) + \sum x_{j1} + 1.5(2 - x_{i2}) + \sum x_{j2}, \quad i = 1, \ldots, 3$$

where summations are taken over all three group members. At the end of the experiment one of the 5 two-stage games was selected at random, and the participant was paid her earnings from that game. These monetary incentives imply that to maximize group earnings participants should contribute all tokens to the group account in all stages, whereas the standard game-theoretic prediction, based on own-earnings maximization, dictates zero contributions.

In our direct-response treatment each participant was informed of the total contributions to the group account in stage 1 by the other two group members and then made a stage 2 decision. In our strategy-method treatment, each participant completed a strategy for the two-stage game, which consisted of a stage 1 contribution and a stage 2 contribution conditional on total stage 1 contributions by the other group members.[6] At the end of the game participants were informed of own-contributions, total contributions by other group members, and their payoff, and they recorded this information on a record sheet.

Throughout the session participants were only allowed to ask questions by raising their hands and speaking to the experimenter in private. Participants were not allowed to communicate with one another throughout the session, except via the decisions they entered on their terminal. At the end of the experiment participants were paid their earnings in private. All sessions lasted less than an hour and participants earned an average of £7.83 (with a minimum of £4.50 and a maximum of £12.00).

4. Results

Our analysis of the data starts with a brief overview. We first confirm that the behavior derived from the strategy method is consistent with that of the direct-response game, and we determine the extent to which these strategies cause contributions to decrease over the repeated public good game. Second, we then classify the elicited strategies, examine the distribution of types, and how these change with repetition. These changes in strategies help us assess the extent to which learning is responsible for decreasing contributions in the repeated public good game.

4.1. Overview

Average contributions are similar in our direct-response and strategy-method treatments: averaging over all games of the direct-response (strategy-method) treatment, participants contribute 37 percent (41 percent) of their endowment in stage 1 and 23 percent (23 percent) in stage 2. Moreover, the time-series of contributions in the two treatments track each other closely, as shown in Fig. 1.[7]

Formal statistical tests fail to reject the hypothesis that the two elicitation methods induce the same contribution behavior. Using two-sided Wilcoxon rank-sum tests applied to session-level data from all games we find no significant

[6] With this strategy-method design the strategy space grows geometrically as stages are added. For example, a three-stage game would require participants to complete a matrix of conditional contributions. Thus, submitting strategies becomes much more cognitively demanding and time-consuming as the number of stages grows, and we therefore focus on the simple two-stage games. Likewise, we restricted the endowment space to 0, 1 or 2 tokens, and group sizes to three, to reproduce a voluntary contributions game structure while maintaining a relatively simple task.

[7] Average contributions for each stage of each session are tabulated in Appendix B of the supplementary material.

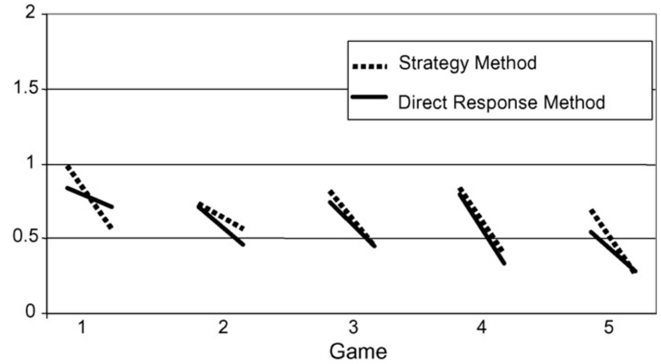

Fig. 1. Average contributions by game. *Note*: The left end of each line indicates the mean contribution in stage 1, and the right end indicates the mean in stage 2.

difference across treatments for either stage 1 contributions ($p = 0.561$) or stage 2 contributions ($p = 0.773$). We also find no significant differences in contributions if we focus on any given game.[8] All of these tests compare two sets of four observations, a rather conservative use of the data. We also made the same comparisons using individual-level data (i.e., comparing 2 sets of 60 observations). Although this procedure is biased toward finding significant differences because it overstates the number of independent observations, we still fail to find any significant differences between elicitation methods.[9]

We also examined changes in contributions between stages 1 and 2. Using a two-sided Wilcoxon rank-sum test applied to session-level data from all games, we again find no significant difference between the two elicitation methods ($p = 0.468$), and with one exception for the case of game 1, we also find no significant differences between treatments if we focus on a given game or use individual-level data.[10] The overall message is that our strategy-method treatment generates behavior remarkably consistent with that of a direct-response method.

The time-series displayed in Fig. 1 are similar to those from previous experiments in that contributions decline within each two-stage game. In game 1 for example, contributions decrease from 45 percent to 32 percent of token endowments (averaging over all eight sessions). Moreover, the same pattern is seen in subsequent games: contributions fall from 36 percent to 26 percent, 39 percent to 23 percent, 41 percent to 18 percent, and 31 percent to 14 percent in games 2–5, respectively.[11] Thus, on average, contribution rates decrease by 16 percentage points between stages 1 and 2, corresponding to an average decrease of contributions of 41 percent. Similar decreases are observed in each treatment (38 percent and 45 percent for the direct-response and strategy-method treatments, respectively). Indeed, average stage 2 contributions are lower than average stage 1 contributions in all eight sessions; thus stage 2 contributions are significantly lower than stage 1 contributions (using a one-sided Wilcoxon matched-pairs signed-ranks test, $n = 8$, $p < 0.005$).

Fig. 1 exhibits substantial 'restart effects'; that is, after contributions decline within a game, they restart at a higher level at the beginning of the next game. While this effect is not significant for the first restart, that is, the transition from game 1 to 2 (p-value = 0.6367 using a one-sided Binomial test), it is significant for the other three (p-value = 0.035 for the second restart, and p-value = 0.0039 for the third and fourth restarts).[12]

[8] The p-values are 0.363 (game 1), 0.883 (game 2), 0.307 (game 3), 0.549 (game 4) and 0.309 (game 5) for comparisons using stage 1 contributions. The corresponding p-values using stage 2 contributions are 0.243, 0.468, 0.773, 0.554, and 0.885.

[9] The p-values are 0.280 (game 1), 0.966 (game 2), 0.575 (game 3), 0.915 (game 4), 0.335 (game 5) and 0.448 (overall) for comparisons using stage 1 contributions. The corresponding p-values using stage 2 contributions are 0.281, 0.441, 1.000, 0.628 and 0.837.

[10] Based on session-level data the p-values are 0.020 (game 1), 0.307 (game 2), 0.884 (game 3), 0.559 (game 4), 0.147 (game 5) and 0.468 (overall). The corresponding p-values using individual-level data are 0.056, 0.577, 0.567, 0.661, 0.245 and 0.433.

[11] While endgame contributions are quite similar to those of Isaac and Walker, stage 1 contributions of the two-stage game tend to lie below stage 1 contributions of the 10-stage game.

[12] The restart observed here differs slightly from that of Andreoni (1988) and Croson (1996), as they focus on the cases where participants either continue with the same group of people after the restart or are rematched with new people after every decision in the game.

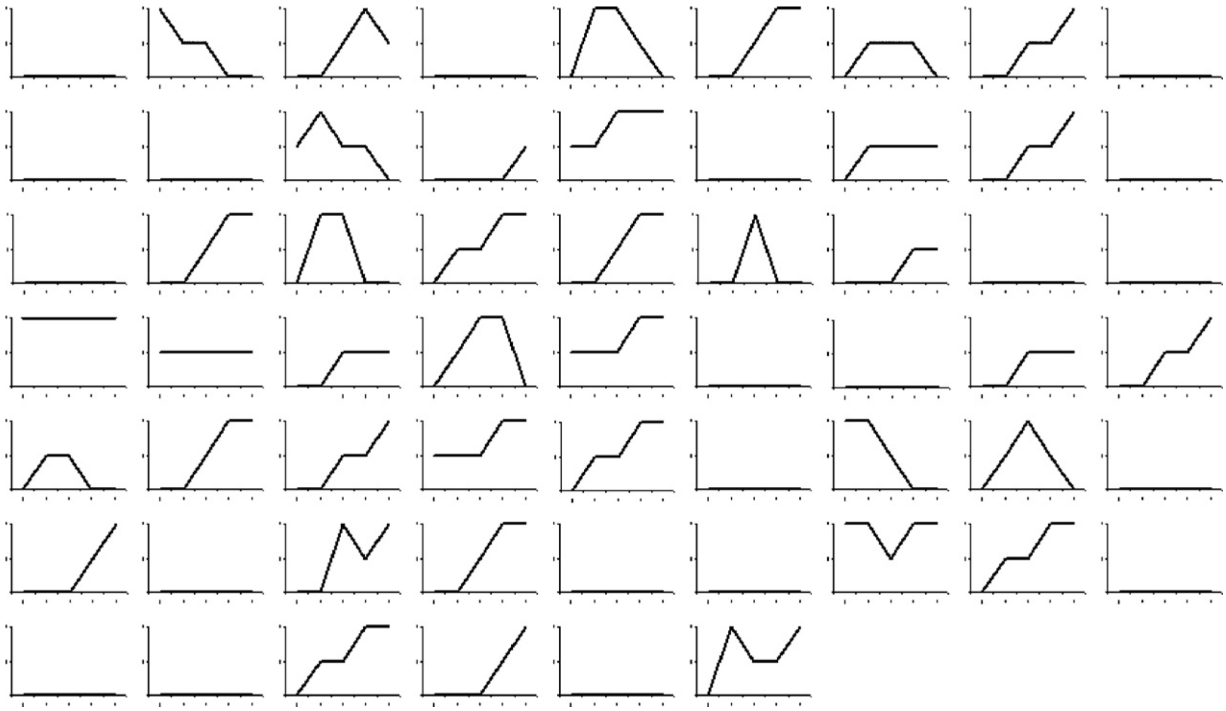

Fig. 2. Stage 2 strategies in game 1. *Note*: *x*-axis measures the aggregate stage 1 contribution by other group members, *y*-axis measures the participant's conditional stage 2 response. A panel represents one participant.

Table 1
Distribution of types by stage 1 contribution

Contribution in stage 1	Fraction (percentage) of			
	Conditional cooperators	Free riders	Hump-shaped	Other
0	3/18 (17 percent)	12/18 (67 percent)	2/18 (11 percent)	1/18 (5 percent)
1	11/25 (44 percent)	5/25 (20 percent)	4/25 (16 percent)	5/25 (20 percent)
2	9/17 (53 percent)	4/17 (24 percent)	3/17 (18 percent)	1/17 (6 percent)

4.2. Individual strategies: game 1

Next we analyze the strategies elicited in the first game. With 4 strategy-method sessions, each involving 15 participants, we elicited strategies from 60 participants. Fig. 2 presents the strategies initially submitted for the second stage by each participant, demonstrating substantial heterogeneity across participants. Following Fischbacher et al., we classify participants into four categories. Of our 60 participants, 35 percent contribute nothing independent of the contributions of others (free riders), while the contribution schedules are increasing for 38 percent (conditional cooperators), hump-shaped for 15 percent, and cannot easily be classified for the remaining 12 percent.[13]

In examining how our participants' stage 1 decisions relate to their stage 2 decisions, we observe some consistency. There is a systematic relation between stage 1 decision and the type classification based on stage 2 decisions, and we reject the hypothesis that the distribution of types is independent of stage 1 decisions ($\chi^2(6) = 13.46$, p-value = 0.036). As can be seen in Table 1, of the 18 participants who contributed zero in stage 1, 67 percent were classified as free riders and 22 percent were classified as conditional cooperators, while of the 17 participants who contributed two

[13] The corresponding percentages in the four person sequential public good game reported by Fischbacher et al. were 30 percent, 50 percent, 14 percent, and 7 percent, and by Fischbacher and Gächter were 23 percent, 55 percent, 12 percent and 10 percent. They use a slightly different classification procedure based on Spearman rank correlation statistics (see their paper for details).

tokens, 24 percent were classified as free riders and 53 percent as conditional cooperators. Thus, our results mirror Brandts and Charness (2000) and Fischbacher et al. (2001) in that stage 1 contributors are more likely to be conditional cooperators.[14] Table 1 also reveals that almost half of the participants who free ride in the second stage nonetheless made a positive contribution in the first stage. Thus free riders too cause contributions to decrease with repetition.

The notion of consistency addressed by these data is quite limited. An own-earnings maximizer may either contribute or not in stage 1, depending on his beliefs about how other participants respond in stage 2. Similarly, a participant who prefers to contribute, but only so long as others also contribute, may either contribute or not in stage 1, depending on his beliefs about other participants' contributions in stage 1. In the next subsection we will examine a different notion of consistency by studying how participants' strategies develop across games.

The predominance of free riding and conditionally cooperative strategies implies that on average stage 2 contributions increase with stage 1 contributions, albeit with a rather shallow slope. At best contributing one additional token in stage 1 increases the stage 2 contribution by each of the other group members by 0.25 tokens.[15] Fitting a line through the conditional contribution schedules using a simple OLS approximation for game 1 yields

$$\text{stage 2 contribution} = 0.29 + 0.16(\text{aggregate stage 1 contributions by other group members}).$$

Thus, by contributing an additional token in stage 1 a participant induces other group members to increase their stage 2 contributions by 0.16 tokens each, on average; as this is not sufficient to cover the £0.50 contribution cost, it is not worthwhile for an own-earnings maximizer to contribute in stage 1.[16] Given this relationship between stage 2 contributions and stage 1 contributions, stage 2 contributions will be below stage 1 contributions as long as stage 1 contributions exceed 21 percent of endowments; in fact stage 1 contribution rates were 45 percent, easily exceeding this threshold.[17]

4.3. Individual strategies: development across games

Next we examine how learning influences behavior, that is how behavior changes across the five repetitions of the games. Twenty-two of our 60 participants submitted the same type of strategy in every game. Thirteen of these submitted free-riding strategies in every game, and seven submitted conditionally cooperative strategies in every game. This leaves almost two-thirds of our participants whose strategy classification varies from game to game.

With 60 participants, each having four opportunities to revise their strategies, there are 240 opportunities for participant strategies to change. Table 2 presents a transition matrix indicating the transition rates from each class. For example, 92 free-riding strategies were observed in the first four games, and in 80 percent of these cases the participant remained in the free-riding category in the next game. Similarly, participants classified as conditional cooperators usually remain in the same category in the next game. In contrast, participants in the "hump-shaped" or "other" category usually transit to a different category.

Transitions from one category to another occur throughout the experiment: of 60 subjects 20 are classified differently in games 1 and 2; 24 in games 2 and 3; 21 in games 3 and 4; and 15 in games 4 and 5. A variety of explanations may account for these transitions. One possibility is that a participant whose behavior changes across games does not have stable preferences, while another possibility is that the participant is aware of what outcomes they prefer, but is learning how best to achieve them. This suggests caution should be used when interpreting strategies and their development over time. First, if participants do have stable preferences, elicited strategies do not map onto these in a direct way.

[14] Brandts and Charness had participants play two games, once as first-mover and once as second-mover (against different opponents and with no feedback between the two games). This allowed them to note a certain consistency in participants' decision rules in the two roles: when making decisions as second-movers, first-mover cooperators were much more likely to respond positively to a cooperative move than first-mover defectors. Fischbacher et al. also require participants to make decisions in two roles and note a similar consistency across decisions in each role. Conditional cooperators contribute 42 percent of their endowment, on average, when they make an unconditional decision, while free riders contribute 10 percent.

[15] The marginal increase in tokens by other group members depends on the initial contribution level. Increasing the contribution by other group members from 0 to 1 or from 1 to 2 tokens will on average increase the contribution by each of the other group members by 0.25, going from 2 to 3 it increases 0.1, and from 3 to 4 it increases 0.03. The average stage 1 contribution by other members is 1.97 in game 1.

[16] This result is based on the submitted strategies; if we instead look at the actual responses that were carried out, the corresponding coefficient on contributing would be 0.18.

[17] That is, denoting average stage 2 contribution by s_2 and average stage 1 contribution by s_1, since aggregate stage 1 contribution by others equals $2s_1$ we can see that: $s_2 < s_1 \Leftrightarrow 0.29 + 0.16 \times 2 \times s_1 < s_1 \Leftrightarrow s_1 > 0.42$. 0.42 tokens represent 21 percent of the endowment.

Table 2
Transitions between types across games

Category	Number observed in any of the first four games	Percentage categorized in the following game as			
		Free rider	Conditional cooperator	Hump-shaped	Other
Free rider	92	80	12	2	6
Conditional cooperator	91	17	69	2	12
Hump-shaped	24	12	25	42	21
Other	33	21	33	6	40

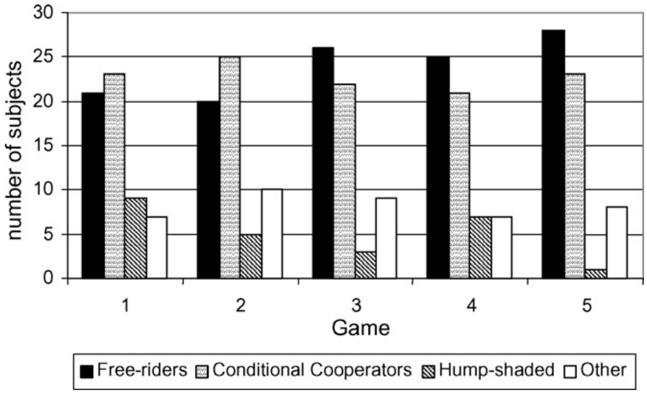

Fig. 3. Distribution of strategy types across games.

Second, if we interpret changes in behavior from game to game as 'learning', many of the observed transitions suggest that this goes beyond simply learning how best to maximize own-payoff.

Fig. 3 shows the distribution of strategy types across the five games. Depending on the game, free-riding strategies account for between 33 percent (game 2) and 45 percent (game 5) of submitted strategies, while conditionally cooperative strategies account for between 35 percent (game 4) and 42 percent (game 2) of submitted strategies. What is clear from Fig. 3 is that together, free riding and conditionally cooperative strategies predominate throughout the experiment. We also see that the distribution of types is fairly stable over the course of the experiment.

The implication of these distributions of strategy types for the average response function is shown in Fig. 4. The presence of conditionally cooperative strategies generates a response function that increases in others' stage 1 contributions, but the rate of increase is dampened by the presence of free-riding strategies.

From Fig. 4 it appears that the average response function is approximately an increasing linear function that shifts downward from game to game. Table 3 reports the results of using simple OLS to approximate the average response function by a straight line. The slopes vary from 0.14 to 0.17, indicating that if a participant were to increase his

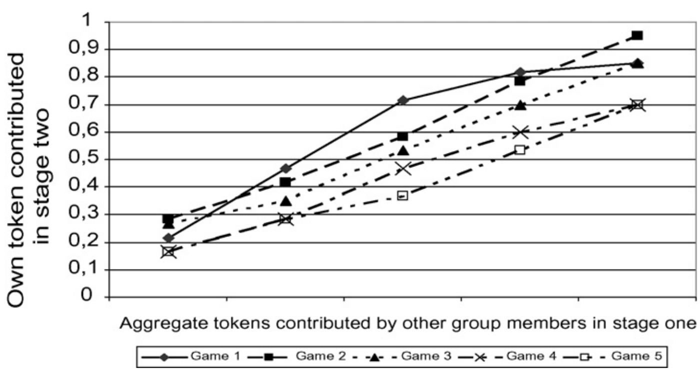

Fig. 4. Strategy method: stage 2 average contribution schedule.

Table 3
Linear approximations to stage 2 conditional response schedule

	Game				
	1	2	3	4	5
Intercept	0.29	0.26	0.24	0.17	0.15
Slope	0.16	0.17	0.15	0.14	0.13

stage 1 contribution by a token, the aggregate contributions made by others in stage 2 would, on average, increase by between 0.28 and 0.34 tokens. Thus, the degree of responsiveness, as measured by this slope, does not vary much and is never strong enough to make it worthwhile for an own-earnings maximizer to contribute in stage 1.[18] The more apparent change over time is with respect to the intercept of the response function, which falls steadily across games. Thus participants respond less generously to a given level of others' stage 1 contributions as games progress. A linear average response function with a slope less than 0.5 implies that if stage 1 contribution rates exceed some critical percentage of endowments, then average stage 2 contributions will be lower than average stage 1 contributions. Given the estimated response functions presented in Table 3, these critical percentages are 21 percent, 20 percent, 17 percent, 12 percent and 10 percent for games 1–5, respectively. Note that these critical percentages all are far below the contribution rates observed in the experiment.

The actual dynamic pattern of contributions depends on how stage 1 decisions and the stage 2 response function develop over games. As seen in Fig. 1, contributions in stages 1 and 2 appear to be lower in game 5 than 1. Wilcoxon matched-pairs signed-ranks tests support this impression: stage 1 contributions are significantly lower in game 5 than 1 ($n=8$, one-sided $p<0.01$), as are stage 2 contributions ($n=8$, one-sided $p<0.005$). However, the effect of experience is quite erratic from one game to the next; for example, stage 1 contributions actually increase between games 2 and 3 and between games 3 and 4. Stage 2 contributions show a more regular pattern of decrease across games. Overall, both stages 1 and 2 contribution rates vary quite widely around an average decrease of 4 percentage points per game. This corresponds to an average decrease in contributions in stage 1 of 7 percent and in stage 2 of 16 percent. Thus the decay in contributions from one game to the next (for a given stage) is lower and less stable than the decay from the first to the second stage (for a given game).

5. Conclusion

Previous experiments where participants repeatedly make voluntary contributions to public goods consistently result in a declining pattern of contributions. The existing literature has discussed two mechanisms that can lead to this pattern. One is that declining contributions are the result of reduced confusion as participants gain experience with the game. The other is that declining contributions result from the interaction of agents with heterogeneous motivations and reflect deliberate strategies whereby some participants condition their contributions on past contributions by others. In one of our experimental treatments participants submit strategies for playing a two-stage game and repeat this process five times, against new opponents each time. This treatment allows for a direct comparison of the effects of strategic responses and learning.

As in previous studies that elicit strategies in social dilemma experiments, most strategies can be classified as either free riding (contributing nothing in the second stage of our game, regardless of others' contributions in the first stage), or conditionally cooperative (where stage 2 contribution is increasing in others' stage 1 contributions). The observed distribution of submitted strategies is robust across the five repetitions and results in declining contributions within our two-stage voluntary contribution games: on average, stage 2 contributions are 45 percent lower than stage 1 contributions. Thus, deliberate strategic choices made by heterogeneous agents generate a substantial decrease in contributions.

[18] If we instead use actual behavior to estimate the marginal increase in contributions, we find coefficients ranging from 0.14 to 0.22. Alternatively we may look at the average marginal response at each contribution level in each of the five games; doing so we find marginal responses ranging from 0.03 to 0.25. Thus, at best a contributor will break even by contributing in stage 1, and with the exception of the first game her monetary return is always strictly lower from contributing in stage 1 (the largest marginal return is 0.19 in the remaining games).

However, for almost two-thirds of our participants the elicited strategy type falls in different classes in different games. This suggests caution be exercised in interpreting someone who uses a ‘conditionally cooperative strategy’ as a ‘conditionally cooperative person’. We suggest that, as in many other experiments, these changes in behavior over time reflect some type of learning. Participants in our experiment are not able to dictate the outcome of a game, so even if they have well-defined preferences over the outcome of the two-stage game, they may not necessarily know what strategy best serves these interests. The changes in strategy may then reflect the attempts of participants to find the strategies that best work for them. Under this interpretation it is important to point out that transitions from one class of strategy to another are not always in the direction that serves the interests of selfish participants. Thus, participants are not simply learning how to maximize own-payoff.

Indeed, although individuals switch between strategy classes throughout the experiment, the *distribution* of strategy types is quite stable, so experience has a relatively small effect on average contributions. On average, stage 1 contributions fall by about 7 percent each time the two-stage game is repeated.

This direct approach to testing the ‘strategy’ and ‘learning’ hypotheses leads us to conclude that strategic responses have a more pronounced and systematic effect than learning. The decline in contributions is much greater and more reliable across the two stages of a given game than across a given stage of successive games. In extrapolating this conclusion to the causes of declining contributions in other experiments, two further comments are in order.

First, we focus on a game that is simpler in many respects than that used in more conventional public goods experiments. Our game consists of three players (compared with four players in the “small” group treatments of Isaac and Walker); each player decides how to allocate a two-token endowment (compared with at least a 10-token endowment in Isaac and Walker) in each of two stages (compared with 10 stages in Isaac and Walker). These differences between our game and the more conventional one were dictated by our implementation of the strategy method; whether the dynamics in contribution behavior differs according to these variables is an open question.

Second, a possible way to account for our results is that our strategy-method approach encourages participants to think about the two-stage game in a way that they would not if they were playing the two stages in sequence. If this were the case, then our strategy-method results may have little to say about the declining pattern of contributions in other experiments where subjects are not required to submit strategies for playing the entire game. However, our results from a control treatment using a direct-response method do not square with this account. In our direct-response treatment subjects contribute 37 percent of their endowments in stage 1 (compared with 41 percent in the strategy-method treatment) and 23 percent in stage 2 (compared with 23 percent in the strategy-method treatment). The similarity between contributions under the two elicitation methods makes us confident that our results are not an artifact of the strategy method.

Acknowledgments

We thank Nick Bardsley, Rachel Croson, Simon Gächter, Elke Renner, two anonymous referees and participants at the Economic Science Association International Meetings, Pittsburgh, June 2003, for useful comments. This research was supported by the Center on Philanthropy, Indiana University-Purdue University at Indianapolis. We thank them for their generosity. Vesterlund also thanks the NSF for financial support and CEBR at Copenhagen Business School for their hospitality.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2007.09.001.

References

Andreoni, J., 1988. Why free ride? Strategies and learning in public goods experiments. Journal of Public Economics 37, 291–304.

Andreoni, J., 1995. Cooperation in public goods experiments: kindness or confusion? American Economic Review 85, 891–904.

Andreoni, J., Croson, R., in press. Partners versus strangers: random rematching in public goods experiments. In: Plott, C., Smith, V. (Eds.), Handbook of Experimental Economic Results. North Holland, Amsterdam.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Brandts, J., Charness, G., 2000. Hot vs. cold: sequential responses and preference stability in experimental games. Experimental Economics 2, 227–238.

Brandts, J., Charness, G., 2003. Truth or consequences: an experiment. Management Science 49, 116–130.

Burton, A., Loomes, G., Sefton, M., 2005. Communication and efficiency in coordination game experiments. In: Morgan, J. (Ed.), Experimental and Behavioral Economics (Advances in Applied Microeconomics 13). Elsevier, Amsterdam, pp. 63–86.

Croson, R., 1996. Partners and strangers revisited. Economics Letters 53, 25–32.

Fischbacher, U., Gächter, S., 2006. Heterogeneous social preferences and the dynamics of free riding in public goods. CeDEx Discussion Paper 2006-1. University of Nottingham.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters 71, 397–404.

Güth, W., Huck, S., Müller, W., 2001. The relevance of equal splits in ultimatum games. Games and Economic Behavior 37, 161–169.

Houser, D., Kurzban, R., 2002. Revisiting kindness and confusion in public goods experiments. American Economic Review 92, 1062–1069.

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

Keser, C., 2000. Strategically planned behavior in public good experiments. CIRANO Working Paper 2000s-35.

Palfrey, T.R., Prisbrey, J.E., 1996. Altruism, reputation, and noise in linear public goods experiments. Journal of Public Economics 61, 409–427.

Roth, A.E., 1995. Bargaining experiments. In: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 253–348.

Selten, R., 1967. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In: Sauermann, H. (Ed.), Beiträge zur experimentellen Wirtschaftsforschung. J.C.B. Mohr, Tübingen, pp. 136–168.